

Alternatives to animal testing: current status and future perspectives

Manfred Liebsch · Barbara Grune · Andrea Seiler · Daniel Butzke · Michael Oelgeschläger · Ralph Pirow · Sarah Adler · Christian Riebeling · Andreas Luch

Published online: 24 May 2011

© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract On the occasion of the 20th anniversary of the Center for Alternative Methods to Animal Experiments (ZEBET), an international symposium was held at the German Federal Institute for Risk Assessment (BfR) in Berlin. At the same time, this symposium was meant to celebrate the 50th anniversary of the publication of the book “The Principles of Humane Experimental Technique” by Russell and Burch in 1959 in which the 3Rs principle (that is, *Replacement, Reduction, and Refinement*) has been coined and introduced to foster the development of alternative methods to animal testing. Another topic addressed by the symposium was the new vision on “Toxicology in the twenty-first Century”, as proposed by the US-National Research Council, which aims at using human cells and tissues for toxicity testing in vitro rather than live animals. An overview of the achievements and current tasks, as well as a vision of the future to be addressed by ZEBET@BfR in the years to come is outlined in the present paper.

Keywords Alternatives to animal testing · ZEBET · 3Rs · Replacement · Reduction · Refinement

Current efforts and future developments at the German Center for Alternatives to Animal Testing—Presented on the occasion of the 20th anniversary of ZEBET at BfR, October 2009, Berlin, Germany.

M. Liebsch · B. Grune · A. Seiler · D. Butzke · M. Oelgeschläger · R. Pirow · S. Adler · C. Riebeling · A. Luch (✉)
German Federal Institute for Risk Assessment (BfR),
Center for Alternative Methods to Animal
Experiments—ZEBET, 12277 Berlin, Germany
e-mail: Andreas.Luch@bfr.bund.de
URL: <http://www.bfr.bund.de/>

On the future of ZEBET

Some general considerations and thoughts

While finalizing the schedule and program of the 20th anniversary symposium, ZEBET staff were highly interested in discussing the possible future direction of the department in order to provide a proposal on how the future of ZEBET might look like, as a milestone not too far away, let us say—maybe in the forthcoming one or two decades.

Before going into what ZEBET’s future might look like, it is worthwhile to assess the *status quo* and to praise what has been achieved during the first 20 years at this German institution. Thus, the initial focus of this section will be the current situation at ZEBET, how it is organized and equipped, and what are its main tasks and commitments. In addition, the current situation of animals used for scientific purposes will be assessed, their numbers, and the trends in alternatives to animal testing.

ZEBET is dedicated to all three facets of the 3Rs, that is, replacement, reduction, and refinement of testing in animals, which have been incorporated into the EU Directive 86/609 (Fig. 1). One of the main tasks of ZEBET is to collect all available information on alternative methods to animal testing and to provide an up-to-date science-based assessment to its stakeholders. ZEBET is also pursuing in-house research and develops modified or new bioassays that would be suitable to replace animal experiments or to reduce animal numbers subjected to particular in vivo assays. At the same time, ZEBET is pushing for validation of alternative methods developed through our own efforts or by any other research group or consortium dedicated to establish alternative methods. Ultimately, ZEBET is interested in the safety evaluation of chemicals or biological products based on alternative methods or testing

strategies, performed by the industry in lieu of “good old” but crude conventional toxicity testing in live animals.

To fulfill all of these tasks, ZEBET is organized in three basic albeit overlapping units dedicated to documentation, evaluation, and research (Fig. 2). Starting at the end of the 1980s as part of the now defunct German Federal Health Authority (BGA), ZEBET is now part of the Federal Institute for Risk Assessment (BfR) for more than 7 years. Being subdivided into a set of different departments, the portfolio of the BfR encompasses the areas of biological safety, food safety, chemicals, and product safety, as well as risk communication. All of these departments have proved important and instrumental in supporting the work and goals of ZEBET ever since the BfR has been established in 2002. Incorporating ZEBET into the larger institution of the BfR holds great promise and offers a range of advantages mainly resulting from the interdepartmental and interdisciplinary interactions. This has occasionally resulted in unexpected progress and success with some of those innumerable activities and engagements initiated by ZEBET. Still, there is further and as yet untapped potential of this structure awaiting to be fully utilized and gainfully applied for the sake of both human health and animal welfare.

Current status of animal experimentation in Europe

A total of about 12.1 million animals have been used in the 25 member states of the EU, according to a report of the European Commission from November 2007, which provides an overview on the numbers of animals used for scientific purposes in Europe in the year 2005 (Fig. 3) (Commission of the European Communities 2007).

For comparison, a USDA/APHIS census estimated that a total of 17–22 million animals were used in research and testing in the US in 1983 (US Congress, Office of

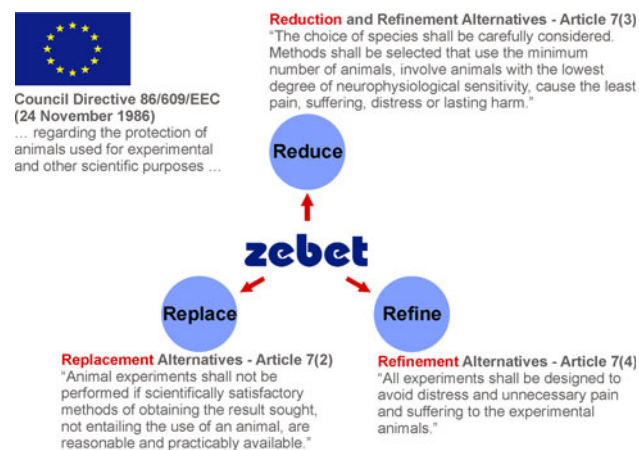


Fig. 1 ZEBET@BfR and the 3R principle

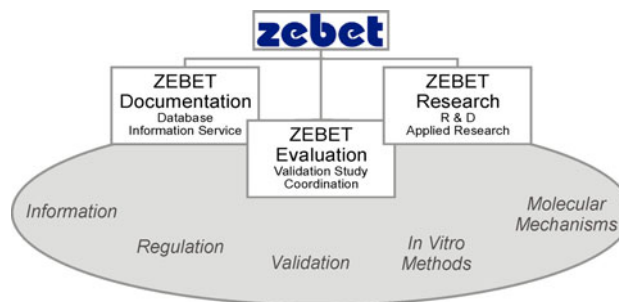


Fig. 2 Organization of ZEBET@BfR

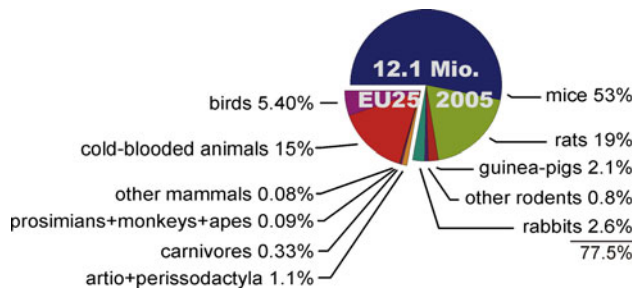


Fig. 3 Animals used for scientific purposes in Europe. The chart shows percentages of animals used by classes and have been taken from the fifth report on the statistics on the number of animals used for experimental and other scientific purposes in the member states of the European Union (Commission of the European Communities 2007)

Technology Assessment 1986). However, a more recent independent estimate suggests up to 80 million animals used, in part due to the advent of transgenic animals (Carbone 2004). The USDA publishes annual reports on animal usage in research; however, those numbers exclude birds, mice of the genus *Mus*, and rats of the genus *Rattus* bred for use in research, according to animal welfare act (AWA) regulations, and are therefore not included in the 1,131,076 animals reported in 2009 (US Department of Agriculture, Animal and Plant Health Inspection Service 2009).

In Europe, of the animals used 79% are mammals, with rodentia, and here mainly mice and rats, making up for 75% of total animals; another 2.6% are rabbits (Fig. 3). The main areas of use are in basic science (33%) and research and development (31%), followed by production and quality control for medical products, substances, or devices (15%).

For toxicology and safety evaluation, about 970,000 animals (8% of total) were used in 2005 (Fig. 4). Testing for medical products, substances, or devices made the largest part in this category with 51%, whereas animal use for the evaluation of products or substances used mainly as animal feed, foods, cosmetics, and household items amount for only 4.3%. Chemicals and pesticides account for 19% of toxicological and safety testing in animals.

From a science perspective, research related to animal and human diseases is responsible for the largest share of 57% of total animals used (this includes the aforementioned research and development and the production and quality control for medical products, substances, or devices), the second is fundamental biology studies (33%) followed by toxicology and safety studies (8%) (Fig. 4). The share of animals used for toxicology and safety evaluation dropped from 9.9% in 2002, when 15 EU countries reported on their animal use, to 8% in 2005, when 25 EU countries reported on their animal use. Despite the increase in reporting countries, this amounts to nearly 40,000 animals less used. However, the total number of animals used increased by 3.1% in the 15 EU countries reporting in both 2002 and 2005 (Fig. 5) (Commission of the European Communities 2007). Similarly, there is a steady increase of animal usage in Germany in the last decade, in part due to rising numbers of transgenic animals and animals killed without prior experimentation for organ and primary cell harvesting (Kretlow et al. 2010).

Animal usage under the REACH legislation

The largest share of animals required for toxicology and safety testing of 42% was used for acute and sub-acute toxicity testing (Commission of the European Communities 2007). A second large fraction of 17.7% was used for testing carcinogenicity, mutagenicity, and toxicity to reproduction (CMR) (Fig. 6) (Commission of the European Communities 2007). In 2003, the EU adopted the Registration, Evaluation, and Authorization of Chemicals (REACH) legislation. It proposes to evaluate 30,000 existing chemicals within a period of 15 years. The number of additional animals required for toxicology and safety testing over this period has been estimated at 9 million (Fig. 7) (Höfer et al. 2004; European Chemicals Agency 2009). However, it has recently been suggested that this number could become as high as 54 million (Rovida and Hartung 2009). In 2005, testing for sub-chronic and chronic toxicity and CMR together accounted for roughly 30% of

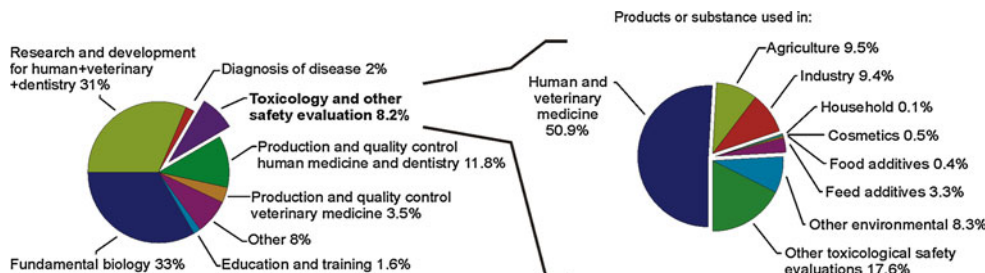


Fig. 4 Areas of animal use for scientific purposes in Europe. The charts show percentages of animal use per area according to the fifth report on the statistics on the number of animals used for

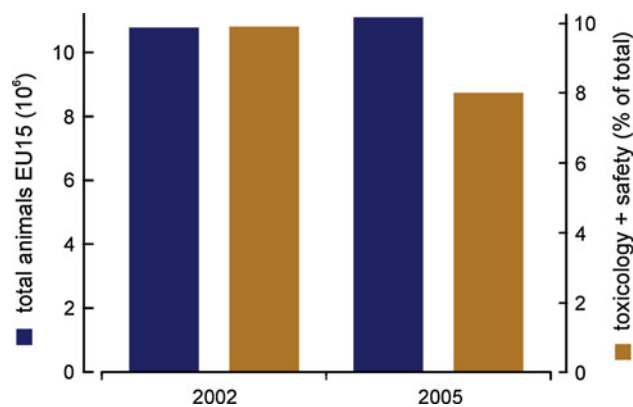


Fig. 5 Comparison of animal numbers used for scientific purposes in Europe. Blue bars depict total animal numbers in million animals for the 15 European member states reporting in 2002 and 2005. Orange bars show percentages of animals used for toxicology and other safety evaluation purposes in the 15 reporting members for 2002 and the 25 reporting states for 2005. Animal numbers are according to the fifth report on the statistics on the number of animals used for experimental and other scientific purposes in the member states of the European Union (Commission of the European Communities 2007)

animals used for toxicology and safety testing. With REACH, the share of this area will increase to about 80% of toxicology and safety testing (Fig. 7) (van der Jagt et al. 2004). These numbers are an economic as well as an ethical call for alternative methods replacing animal testing. ZEBET has been already successfully advocating alternative methods for international acceptance at the Organization for Economic Co-operation and Development (OECD) and continues to do so. These efforts will be described in more detail below.

How to test for CMR and long-term organ damage without animals?

The first toxicological consideration is that of exposure. Every living being has some form of barrier to protect it from environmental stresses and impacts (Fig. 8). For the human body, the main barriers are the skin and the mucous

experimental and other scientific purposes in the member states of the European Union (Commission of the European Communities 2007)

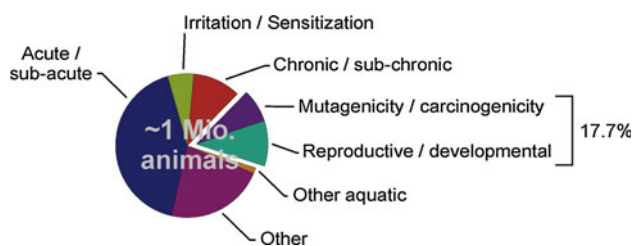


Fig. 6 Areas of animal use for toxicology and safety testing in Europe. The *charts* show percentages of animal use per area in 2005 according to the fifth report on the statistics on the number of animals used for experimental and other scientific purposes in the member states of the European Union (Commission of the European Communities 2007)

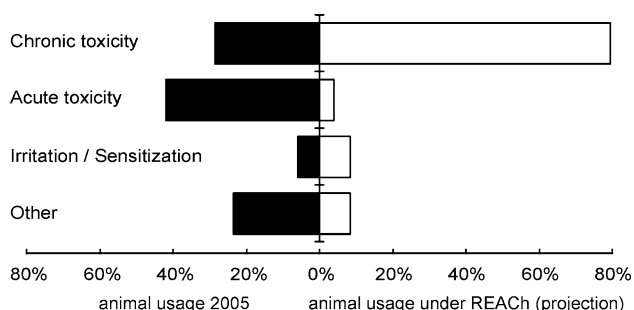


Fig. 7 Projected numbers of animals for testing under REACH in Europe. The *charts* show percentages of animal use per area in 2005 compared to projected total animal usage for REACH according to van der Jagt et al. (2004). Chronic toxicity in this chart comprises testing for chronic, sub-chronic, and reproductive toxicity, as well as for carcinogenicity and mutagenicity

membranes. The first possible routes of exposure are therefore skin penetration, inhalation, and exposure of the respiratory tract, as well as ingestion and exposure of the gastrointestinal tract. A number of *in vitro* models are commercially available as three-dimensional tissue reconstruction models (Pampaloni et al. 2009). Some examples are epidermis, full-thickness skin models, respiratory epithelia, keratinocyte eye cornea, vaginal epithelia, and oral epithelia. Blood-tissue barriers form a second line of defense, the most prominent of which are the blood–brain barrier and the placental barrier; and the blood–brain barrier has been modeled *in vitro* (Stolper et al. 2005; Cecchelli et al. 2007). The advance in three-dimensional models has also allowed for internal organ models (Pampaloni et al. 2007), such as placenta (Mess 2007), lymph node (Giese et al. 2006), and liver (Linke et al. 2007). However, the adult human body presents with about 40 different organs composed of at least 400 cell types (Fig. 9) (Vickaryous and Hall 2006).

To reconstruct all human tissues, *in vitro* would be a formidable task today, and toxicological testing on all separate tissue models would put even the direst estimates of animal use in a temporally and economically favorable light.

External defense barriers Internal defense barriers

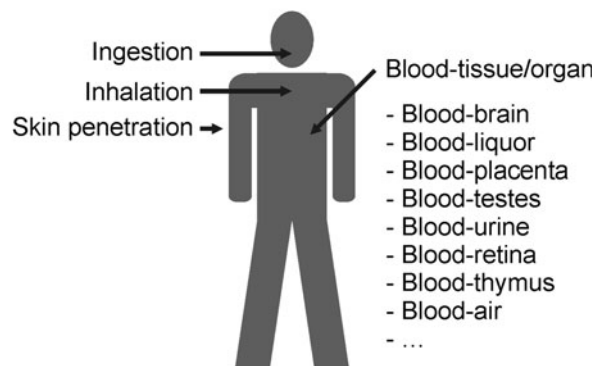


Fig. 8 Barrier functions of the human body

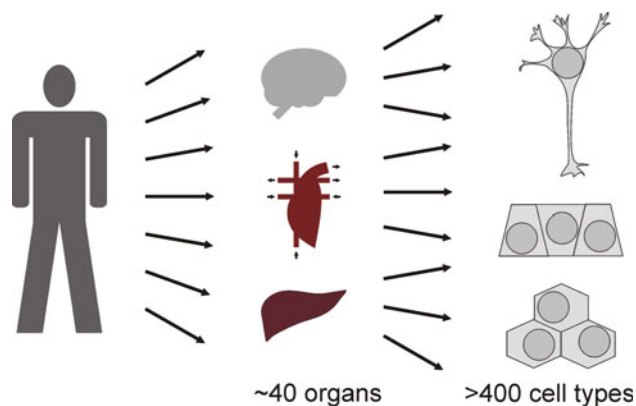
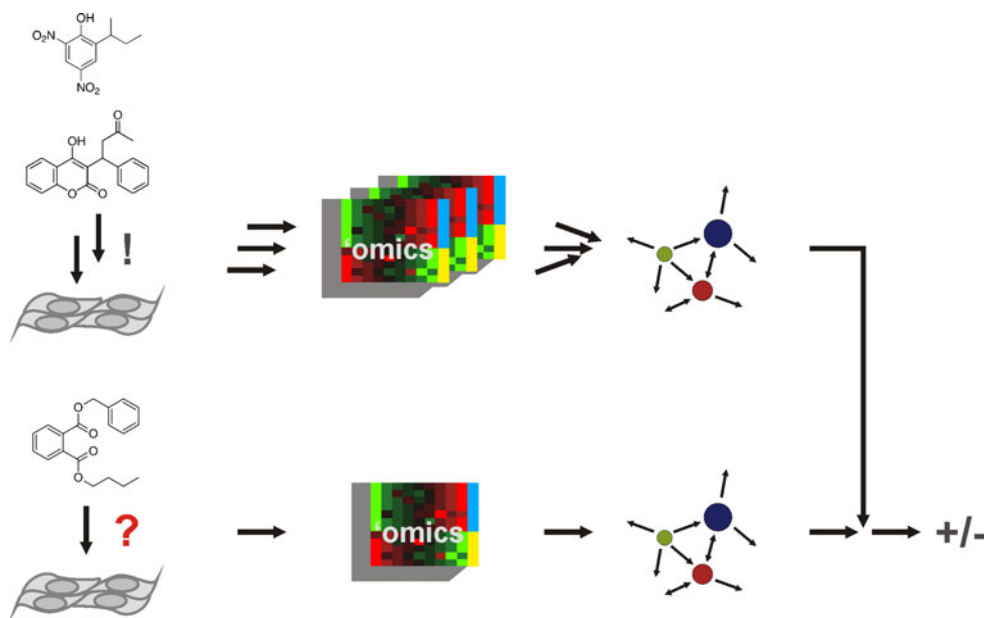


Fig. 9 Composition of the human body

How then can we achieve an idea of the systemic toxicity of substances without animal use? A promising approach is computer-aided modeling of the systemic behavior of a substance, so-called *in silico* methods. The physiology-based pharmacokinetic (PBPK) model integrates information from different *in vitro* and *in silico* quantitative structure–activity relationship [(Q)SAR] approaches and predicts the systemic availability of a compound under certain exposure conditions and hence allows to evaluate the toxicological relevance of an *in vitro* derived toxic concentration (Verwei et al. 2006). An approach to predict the target effects of a substance is the Connectivity Map (Fig. 10) (Lamb 2007). Here, mRNA expression profiles are collected in a database of a range of different cell lines exposed to substances with known mechanisms under a small number of defined conditions. This limitation to ubiquitous cellular pathways is a necessity since expression profiles themselves are already large datasets. The expression profile in response to a new substance is then compared to the signature database, and by similarity a mode of action, including off-target effects, can be predicted (Fig. 10).

We are still far away from a pure *in vitro*—*in silico* approach. However, the described methods are already

Fig. 10 Schematic representation of the principle of the connectivity map. Data on the cellular response to known substances are collected using several profiling methods such as proteomics and transcriptomics. The results are combined in an interaction map, the connectivity map. The cellular response profile of an unknown substance is compared to the connectivity map and a mechanism of action deduced from similarities to profiles of known substances



used today by the pharmaceutical industry to evaluate drug candidates during research and development and thereby reduce animal use in that area. A number of *in vitro* methods have been accepted by the OECD or are under review for acceptance. Computational approaches have the same potential to drastically reduce animal use under the REACH legislature.

Current efforts at ZEBET

Toward international acceptance of alternative methods

Already in 1990, an international committee under participation of ZEBET convened to develop a scientific concept for the experimental validation of toxicological testing methods (Balls et al. 1990). Followed by two similar workshops (Balls et al. 1995), and two OECD workshops in 1996 and 2002 (Organisation for Economic Co-Operation, Development 1996, 2002), this ultimately led to the OECD Guidance Document No. 34 finally ratified in 2005. This process not only contributed to the successful validation and international recognition of alternative methods for toxicological hazard identification tests but also established that the agreed principles of validation hold for any new and updated test method, whether it is an *in vivo* or *in vitro* method, and both for ecotoxicity tests as well as for human health tests. In addition, this exemplifies the long process of introducing new OECD guidelines, which is in part due to the fact that a unanimous consensus has to be found. ZEBET scientists are involved in the development of OECD Test Guidelines (TG) and Guidance Documents

(GD), which has led to an improvement in the welfare of laboratory animals (Table 1). An internationally harmonized testing helps companies and their products to be competitive in the global marketplace in addition to preventing unnecessary repeated testing on live animals or *in vitro*.

The experimental validation of any new or updated toxicological test method requires examining both its intra- and inter-laboratory reproducibility as well as its performance in the prediction of toxic properties to humans. The translation of test method readouts into human health predictions requires prediction models which are often based on biostatistical methods. Thus, prediction models are crucial components of test methods as they must guarantee correct predictions about new substances for the purpose of human health protection. The incorporation of the “prediction model concept” into GD 34 was a major concern of ZEBET and ECVAM (Archer et al. 1997), and probably the most important contribution to this document through a special OECD expert consultation meeting on “Data Interpretation Procedures” in 2004.

Over the years, it has been understood that competent authorities, of which ZEBET itself is a part, and stakeholders need to be involved throughout the entire validation and acceptance process. In addition to their participation in the national and later international consolidation processes (e.g., OECD, ICH, ISO), they should be involved already in earlier steps. The definition of their information needs and identification of suitable readouts and endpoints of the new method, the selection of suitable tests and test chemicals, and subsequently the peer review of the method all benefit from their involvement, and

Table 1 OECD test methods that have been improved in respect to animal welfare under participation of ZEBET

<i>Complete replacement of the animal experiment</i>	
TG 428	Skin absorption: in vitro method
TG 430	In vitro skin corrosion: transcutaneous electrical resistance (TER)
TG 431	In vitro skin corrosion: human skin model test
TG 432	In vitro 3T3 NRU phototoxicity test
TG 437	Bovine corneal opacity and permeability test method for identifying ocular corrosives and severe irritants
TG 438	Isolated chicken eye test method for identifying ocular corrosives and severe irritants
TG 439	In vitro skin irritation: reconstructed human epidermis (RhE) test method
<i>Reduction in the number of animals and stress of the laboratory animals</i>	
TG 420	Acute oral toxicity—fixed dose procedure
TG 423	Acute oral toxicity—acute toxic class method
TG 425	Acute oral toxicity—up-and-down procedure
TG 429	Skin sensitization—local lymph node assay
TG 436	Acute inhalation toxicity—acute toxic class method
<i>The following OECD guidance documents have been developed or improved from the animal welfare perspective</i>	
GD 14	Detailed review document on classification systems for eye irritation/corrosion in OECD member countries
GD 16	Detailed review document on classification systems for skin irritation/corrosion in OECD member countries
GD 19	Guidance document on the recognition, assessment, and use of clinical signs as humane endpoints
GD 24	Guidance document on acute oral toxicity testing
GD 28	Guidance document for the conduct of skin absorption studies
GD 34	Guidance document on the validation and international acceptance of new or updated test methods for hazard assessment
GD 39	Guidance document on acute inhalation toxicity testing
GD 69	Guidance Document on the validation of (quantitative) structure–activity relationship [(Q)SAR] models
GD 105	Report on biostatistical performance assessment of the draft TG 436 acute toxic class testing method for acute inhalation toxicity

importantly help to avoid the development of tests not suitable to support regulatory decisions. In later stages, regulators should be involved in the definition of performance standards and the definition of special studies, e.g., to enlarge the applicability domain of a new test method which can also mean the extension of regulatory acceptance into new areas not yet validated.

This lesson was learned during the evaluation of the Corrositex assay for skin corrosion (Interagency Coordinating Committee on the Validation of Alternative Methods 1999). The test producer used a panel of chemicals during development that mostly covered pH values <3 or >11, substances which would be suspected as corrosive by that fact alone. Indeed, chemical categories with a majority of chemicals providing extreme pH values have been finally accepted to be the validated applicability domain of the Corrositex assay, and it is therefore now instrumental for hazard sub-categorization of corrosive chemicals and products in the context of UN regulations for transportation of dangerous goods.

Similarly, at the heart of the idea of performance standards for test guidelines was the acute problem of loss of suppliers for two human reconstituted skin models during their early validation for testing for skin corrosivity (Balls 1997). Performance standards define global criteria that a

test method is expected to fulfill, independent of a specific test setup. To emphasize the importance of performance standards, we may look at the achievements in the automobile industry. When renting a car of unknown make at an airport nowadays, due to standardization, we are familiar with the instruments and are able to safely drive within a few minutes without the need of a manual and further advice. Most importantly, we can trust that in a dangerous situation, the car will behave similar to other cars; for instance, due to internationally agreed performance standards, braking distances of today's cars are in the range of $40 \text{ m} \pm 7\%$ when coming to a halt from a speed of 100 km/h. Translated into the area of new toxicological test guidelines, this means that definition of performance standards in the test guidelines will allow for so-called catch-up validation. For instance, the OECD TG 431 defines general and functional conditions that an in vitro human reconstructed (dermal or epidermal) model must meet before it can be used routinely for skin corrosion testing. In addition, the guideline requires correct prediction of twelve reference chemicals as well as assessment of intra- and inter-laboratory variability. This allows new developments or very similar assay systems to be rapidly incorporated into guidelines. For example, the SkinEthic RHE in vitro corrosion test was accepted by ECVAM little

more than half a year after publication of its catch-up validation study (Kandarova et al. 2006; European Center for the Validation of Alternative Methods 2006). The more recent TG 439 comprises three elements for performance standards, essential test method components, a minimum list of twenty reference chemicals, and defined reliability and accuracy values. The latter values represent the sensitivity, specificity, and accuracy of current methods, and any new method has to perform equal or better.

ZEBET is also in the unique position to be able to perform “horizontal” method re-evaluation. A study investigating the necessity for the use of one rodent and one non-rodent animal species, usually rats and dogs, in parallel for repeated dose toxicity of pesticides revealed that treatment of dogs for longer than 90 days provides no additional information indispensable for risk assessment (Spielmann and Gerbracht 2001; Box and Spielmann 2005). This was possible since ZEBET had access to regulatory data covering 40 years of authorization of pesticides from companies supplying Germany through the competent regulatory authority (BGA and later BgVV). At ZEBET, such data, which is in most cases not published and confidential, can be analyzed anonymously.

Another example of the scientific use of proprietary data for the development of an *in silico* prediction tool is the development of the “decision support system” for skin irritation and corrosion. Based on the confidential EU New Chemicals Database, a set of exclusion rules derived from physicochemical properties and structural alerts derived from SAR models were developed at the BGA (and later BgVV) in cooperation with the Dutch Ministry of Health (RIVM) and the US-Environmental Protection Agency (US-EPA) for identifying chemicals that are unlikely or likely to cause skin irritation or corrosion (Gerner et al. 2004; Hulzebos et al. 2005; Walker et al. 2004, 2005). This *in silico* method is recommended to be used within the new EU regulation of chemicals (REACH) and is currently integrated in the OECD (Q)SAR Application Toolbox.

The use of embryonic stem cells in developmental toxicity testing

Congenital abnormalities represent perhaps the most severe side effects a chemical can have, and their prevention is an essential goal in toxicological safety assessment of chemicals and drugs. To evaluate adverse effects on reproduction and embryonic development, mandatory OECD test guidelines, or so-called segment studies, encompassing three crucial periods of pre- and postnatal development and fertility have been established (International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use 2005a). These

guidelines specify time-consuming and expensive *in vivo* experiments mostly performed on mammalian species such as rats or rabbits (Fig. 6).

Over the past 30 years, a wide spectrum of *in vitro* models has been developed to detect teratogenic effects of chemicals. These test systems either utilize whole embryos or dissociated cells from embryos (Piersma 2006). In recent years, stem cells have been the subject of increasing scientific interest because they represent an important new tool for developing unique *in vitro* model systems. They also have great potential to predict or anticipate compound-triggered toxicity *in vivo*. Stem cells are capable of self-renewal, that is, they can be continuously cultured in an undifferentiated stage, giving rise to more specialized cells such as heart, liver, bone marrow, blood vessels, pancreatic islets, and nerve cells upon addition or removal of certain growth factors [Fig. 11, reviewed in (Hoffman and Merrill 2007)]. In 1997, ZEBET developed an *in vitro* model for screening of embryotoxicity based on blastocyst-derived embryonic stem cells from mice (mESC). This so-called embryonic stem cell test [EST; (Spielmann et al. 1997; Seiler et al. 2006a)] is based on the assessment of three toxicological endpoints: (1) the morphological analysis of beating cardiomyocytes in embryoid body outgrowths compared to cytotoxic effects on, (2) mESC, and (3) 3T3 fibroblasts (Fig. 11). As an *in vitro* system which mirrors both proliferation and differentiation, the EST proved in an international European Centre for the Validation of Alternative Methods (ECVAM) validation study to be a reliable assay for the prediction of embryotoxicity *in vivo* (Genschow et al. 2004). Using a set of 20 reference compounds with different embryotoxic potencies (non-embryotoxic, weakly embryotoxic, and strongly embryotoxic), the EST was shown to provide a correct judgment in 78% of all experiments. Remarkably, a predictivity of 100% was obtained for strong embryotoxicants. As a consequence, the validated EST has been accepted and successfully introduced by many pharmaceutical companies as a tool for testing the developmental toxicity of lead compounds at an early stage in R&D toward new drug candidates (Paquette et al. 2008; Whitlow et al. 2007).

One prominent drawback of the classical EST, which is shared by other *in vitro* assays for embryotoxicity such as the whole embryo culture test and the rat limb bud micromass test, is its reliance on a morphological endpoint. Although all of these tests systems offered 100% predictivity for strong embryotoxicants as verified in the ECVAM validation trial, it is of great concern that they all rely on experienced laboratory personnel to produce high-quality data (Spielmann et al. 2006). In order to improve the accuracy of these assays, we and other research groups are now focused on the identification of novel molecular

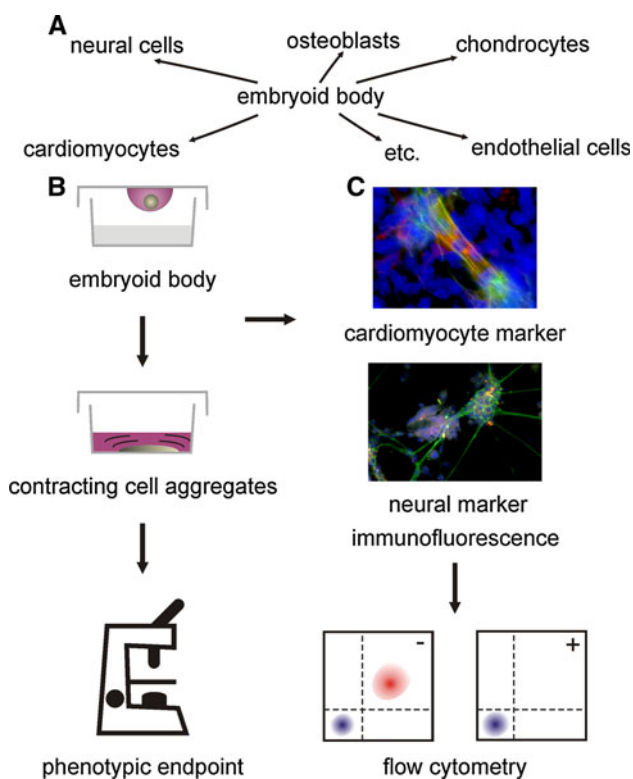


Fig. 11 Using embryonic stem cells for embryotoxicity testing. **a** Embryonic stem cells have the potency to differentiate into all cell types of the body. **b** Schematic representation of the validated embryonic stem cell test (EST) for embryotoxicity testing. **c** Modification of the EST to broaden its applicability domain using molecular endpoints

endpoints to be assessed by more objective and quantitative means, such as gene expression analysis based on real-time RT-PCR and flow cytometry (Seiler et al. 2004; zur Nieden et al. 2004). In a recent study, we worked out a new molecular approach based on the analysis of the expression of certain marker proteins specific for developing heart tissue (i.e., sarcomeric MHC and α -actinin) using quantitative flow cytometry analyses (Fig. 11). The molecular FACS-EST offered the same sensitivity compared to the validated EST protocol but the test duration could be significantly reduced. Due to these improvements, this new molecular method holds promise as a sensitive, more rapid and reproducible screen highly suited to predict developmental toxicity in vivo from in vitro data (Buesen et al. 2009). Recent studies on glycol ether alkoxy acid metabolites (de Jong et al. 2009) and valproic acid derivatives (Riebeling et al. 2011) nicely demonstrated that both the standard EST and the molecular FACS-EST can be reliably used as a tool to assess structure-dependent teratogenicity.

Currently, the stem cell research group at ZEBET is committed to exploring and developing additional stem cell-based approaches, searching for novel predictive biomarkers of developmental toxicity, and extending the

experimental approach to other cellular systems for the prediction of developmental neuro- and osteotoxicity. Promising differentiation protocols for certain cell types most susceptible to chemical-mediated toxicity during early development like cardiac, bone, and neural cells have been successfully developed. For neural cell development, we could demonstrate that mouse embryonic stem cells can be efficiently differentiated in vitro into cell types present in the nervous system like mature neurons, astrocytes, and oligodendrocytes (Seiler et al. 2006b). On the basis of this approach, new rapid and predictive in vitro screens for developmental neurotoxicity testing have been developed. Currently, investigations are underway to explore the use of pluripotent stem cell lines derived from primate blastocysts in assessing developmental osteotoxicity. Due to the close evolutionary relationship, these cells might help to improve the predictivity for human toxicity.

Recently, new exciting avenues of research on the role of microRNA (miRNA) in toxicogenomics and the possibility of epigenetic effects on gene expression were identified. Therefore, miRNA profiling opens the possibility to discover new molecular endpoints that might contribute to a further understanding of chemical-mediated developmental toxicity. Current investigations are aimed at studying miRNA expression in differentiating mouse embryonic stem cells and their dysregulation upon exposure to embryotoxic compounds. Furthermore, in line with the report from the National Research Council on toxicity testing in the twenty-first century, which has proposed fundamentally new directions for toxicity testing in light of advances in understanding biological responses to chemical stressors (Krewski et al. 2010), new research projects at ZEBET involve the mapping of toxicity pathways in differentiating mouse and human stem cells as well as pluripotent stem cell lines and the identification of critical pathway perturbations that either correlate with or directly represent molecular initiation events for adverse effects during human embryonic development.

Implementing the 3Rs principle by funding the development of alternative methods

An increasing demand for health risk assessment due to the REACH program of the EU, as well as changes in the legislature exemplified by the 7th Amendment of the Cosmetics Directive that prohibits the use of animal experiments for the toxicological evaluation of cosmetic ingredients, has stimulated the search for novel experimental approaches that have the potential to reduce or replace animal experiments.

The EU responded to these needs and joined with the cosmetic industry to establish a €50 million research program as part of the 7th EU framework program for the

development of alternative test methods in the field of repeated dose systemic toxicity. A survey conducted in 2006–2007 estimated the total funding of 16 participating European countries at about €17 million (Devolder et al. 2008). The total funding in Germany was estimated at €4.6 million, mainly due to the “Alternatives to Animal Testing” program of the Federal Ministry of Education and Research (BMBF) that has been already established in 1980 (Bundesministerium für Bildung und Forschung 2001). In Germany, this is the only program that provides sufficient financial support for the cost-intensive and time-consuming process of validation. In addition, several federal states, including Baden-Wuerttemberg (<http://www.mlr.baden-wuerttemberg.de>), Hessen (<http://www.hmulv.hessen.de>), and Rheinland-Pfalz (<http://www.mufv.rlp.de>), as well as the private Foundation for the Promotion of Alternate and Complementary Methods to Reduce Animal Testing (SET; <http://www.tierversuche-ersatz.de>), and ZEBET@BfR support such research projects, although with much more limited resources (Fig. 12).

On the other hand, the US National Academy of Sciences published its vision of a modern toxicology in the twenty-first century in 2007 (Krewski et al. 2010). The US-EPA, the NIEHS National Toxicology Program, and the National Institutes of Health Chemical Genomics Center joined forces to follow this ambitious proposal to develop new toxicity testing strategies, with an emphasis on high throughput technologies to establish toxicological

“fingerprints” or reveal toxicological pathways of chemicals, complex mixtures, and pharmaceuticals (Krewski et al. 2009, 2010; Dix et al. 2007; Andersen and Krewski 2009).

These more recent activities reflect a completely changed perception of alternative and *in vitro* methods as a tool for toxicological risk assessment. Following the implementation of the EU Directive 86/609 and the revision of the German animal welfare act, promoting the development of alternative methods became one of the main missions of ZEBET. Since its inception, ZEBET supports research projects throughout Germany which have the potential to provide novel and innovative experimental approaches to reduce or replace animal experiments with a unique funding program (Fig. 12).

The financial resources of ZEBET@BfR allow for the support of approximately ten projects at a time, amounting to over 100 funded research projects over the last 20 years (Fig. 12). The emphasis is given to projects that lack the experimental evidence that is necessary to seek financial support from larger funding agencies but target important regulatory needs and appear promising to contribute substantially to animal welfare. In this respect, the integration of ZEBET into the BGA first and the BfR later proved instrumental, since the close proximity to regulators facilitated the exchange of information concerning regulatory needs and thus the selection of meaningful projects at the right time.

Over a period of 2–3 years, the investigator is provided with sufficient funds to gather experimental evidence in proof-of-concept studies. For example, the establishment of recombinant Chinese hamster V79 cell lines ectopically expressing human cytochrome P450 enzymes allowing the analysis of drug metabolism and toxicity led to subsequent funding by the BMBF and EU and the successful foundation of a company (Döhmer 2001). In addition, the development of bioreactors enabled scientists to avoid the use of ascites as a source for monoclonal antibodies and are now used to simulate various organs, including the human lymph node, to evaluate the sensitization potential of chemicals or pharmaceutical substances (Giese et al. 2006). Furthermore, initial work on the development of the monocyte activation test as an alternative to pyrogen testing in rabbits was supported by the ZEBET and will be adopted by the European Pharmacopoeia (Hartung et al. 2001).

ZEBET@BfR also funded the development of various *in silico* methods based on structural similarities that gained increasing importance over the last years in the evaluation of potential adverse effects, but also for the analysis of *in vitro* data. For example, the lazy structure–activity relationships (LAZAR) program is used to predict genotoxic activities of chemicals based on structural similarities to

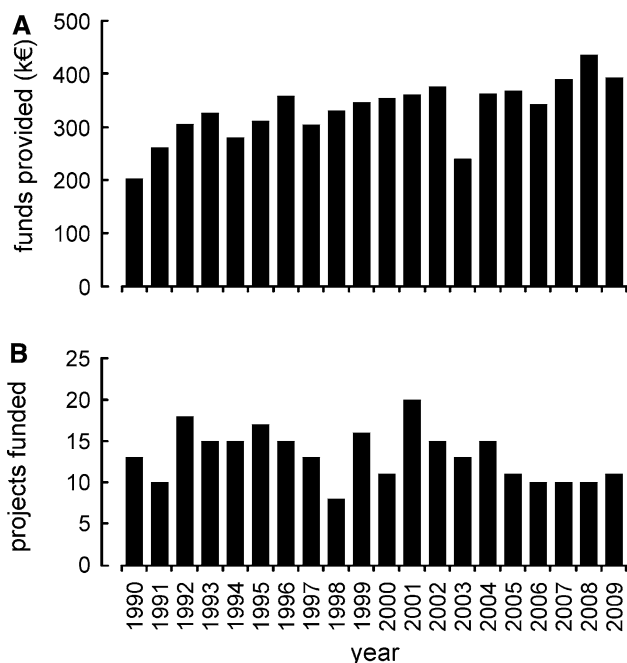


Fig. 12 Research funding at ZEBET@BfR. Charts show **a** the funds available to ZEBET@BfR for funding research and **b** the number of projects funded by ZEBET@BfR since the inception of ZEBET

chemicals with known *in vivo* toxicity data (Helma 2006). Similarly, ZEBET provided the financial support to allow the development of software required for the prediction of the phototoxic potential of chemicals based on the results obtained by the *in vitro* 3T3 NRU phototoxicity test (OECD TG 432), and the biometric evaluation of the Acute Toxic Class (ATC) method for acute inhalation toxicity (OECD TG 436). These projects were essential to achieve international acceptance of these methods as OECD test guidelines (Holzhütter 1997; Holzhütter et al. 2003). Further, the establishment of the so-called Registry of Cytotoxicity that is based on published data from hundreds of *in vitro* cytotoxicity assays provides strong evidence that the *in vivo* toxicity of chemicals can be predicted from *in vitro* data (Halle 2003). This in turn can have a significant effect on the number of animals used for *in vivo* toxicity testing, e.g., by facilitating the calculation of the starting dose for acute oral toxicity studies. Finally, *in silico* models that provide important information concerning the permeability of the human skin have been developed that correlate well with results obtained using *in vitro* experimental skin models (Hansen et al. 2008).

A large number of cell-based, organotypic or *ex vivo* approaches were supported over the years, including models for specific toxicological effects on liver, heart, lymph node, ovary, ear, central nervous system, cornea, or the skin (Fig. 9). In particular, the use of human skin models was promoted by ZEBET. As an *in vitro* alternative to the rabbit test for skin irritation and corrosion (OECD TG 404), this model is now accepted worldwide. In addition, these models are very useful for the prediction of skin adsorption and penetration of chemicals and pharmaceuticals (Schäfer-Korting et al. 2008).

In summary, the support by ZEBET of a broad spectrum of activities and ideas was very successful in promoting the development of *in vitro* systems that already reduce or replace animal experiments or bear the promise to contribute to the 3Rs concept in the near future. The success of the ZEBET@BfR funding program is also reflected by the various national and international awards received by the funded scientist for their work and contributions to the 3Rs as summarized in the brochure published by the BfR on the occasion of the 20th anniversary of ZEBET (<http://www.bfr.bund.de/cd/30995>).

ZEBET's competence in searching for literature and information on alternative methods to animal experiments

Acknowledging that “man has a moral obligation to respect animals and to have due consideration for their capacity for suffering” (Council of Europe 2005), the EU has stipulated legislation to protect animals that are used

for experimental and other scientific purposes (Council of Europe 2005).

As a consequence, in addition to morality, there is a legal obligation to identify and use appropriate methods to replace, reduce, or refine experimental animal use. This obligation considers the internationally accepted 3Rs concept that was laid out by Russel and Burch in 1959. Concerning replacement, an “experiment shall not be performed if another scientifically satisfactory method of obtaining the result sought, not entailing the use of an animal, is reasonable and practicably available” (Council of Europe 2005). Furthermore, as to reduction and refinement experiments which “use the minimum number of animals, involve animals with the lowest degree of neurophysiological sensitivity, cause the least pain, suffering, distress, or lasting harm” should be selected (Council of Europe 2005). Therefore, the search for relevant information about methods compliant with the 3Rs is a key issue of authorization procedures for animal experiments in Europe.

Scientists in the European countries planning to conduct animal experiments are obliged to undertake a valid “indispensability search” prior to applying for an authorization of the experiment at the national competent authority. The aim of an indispensability search is to prove the lack of the presence of (1) a suitable alternative method according to the 3Rs concept that can be applied instead, (2) usable results from comparable previous animal experiments, and (3) results from other research suited to anticipate the outcome of the planned experiment. Only when the availability of a suitable alternative or of usable scientific results have been excluded based upon the current state of knowledge, an animal experiment may be authorized.

Scientifically relevant databases like PubMed provide the opportunity to search an ever-growing number of documents simultaneously via simple and usually general keywords. After retrieving the hit list, extensive efforts are usually required for sorting out irrelevant literature. Moreover, the typical curriculum of scientists completely lacks courses in information retrieval. Most scientists may only be capable of applying ordinary searcher skills. In the field of patent affairs, where the “novelty search” has a similar significance, it is estimated that the cost of duplicate research due to irrelevant information retrieval amounts to about €20 billion a year in Europe alone (European Patent Office 2009).

In order to support the implementation of animal protection obligations in the sciences, ZEBET advises individual scientists and authorities how to obtain, evaluate, and apply information on suitable alternative methods.

ZEBET follows a three-fold strategy to improve information dissemination on alternative methods to animal

experiments that consists of (1) capturing of and supply with information, (2) education in reliable search procedures, and (3) research in retrieval technology.

AnimAlt-ZEBET: all the essential information in a nutshell

Evaluation of value-added databases represents the most reasonable starting point for any structured search for information (Emmerich 2009). At the forefront of these essential resources is the AnimAlt-ZEBET database hosted by the German Institute of Medical Documentation and Information (DIMDI), accessible online free of charge (Fig. 13; <http://www.dimdi.de/static/en/db/dbinfo/zt00.htm>). The documents in the database are written in a structured and standardized manner and focus on the most essential facts of the listed alternative methods. The added value of these documents is constituted through expert selection of the incorporated literature, committed to the most substantial and reliable information only, and a subsequent stringent peer review process including scientists with complementing *in vivo* and *in vitro* expertise. Composition and phrasing of the method summaries are standardized to enable immediate orientation and easy comprehension, and to provide the possible feedstock for up-to-date text mining applications. To obtain a general idea on the method at hand, the reader simply can consult the meta-data fields “Title”, “Evaluation”, “Status”, and “Regulation”, in combination with the abstract section “3R relevance”. Here, users will find highlighted facts to consider, the specific objective that can be addressed by employing the method at hand, the state of development, its acceptance in the scientific and regulatory communities, its application domain, and its contribution to the 3Rs concept. The main text of the method summary aims to cover all aspects necessary for understanding the suitability and applicability of the method with a focus on practical considerations. The section “Expert Panel Opinions” provides a detailed picture on the acceptance of the particular approach and possible objections by official bodies.

In the basic sciences, the impact of a method is judged by bibliometric analyses, i.e., citation analyses and the impact factor of the publishing journal. For a more informed evaluation, it is planned to conduct expert consultations whenever a sufficient body of alternative methods addressing a defined topic in the basic sciences has been compiled. The primary objective of the consultations will be to reach consensus on the relevance and the foreseen application domains of the given methods.

Currently, the emphasis of the database is on methods in toxicity and potency testing, e.g., alternatives in skin sensitization, eye irritation, or Botulinum neurotoxin potency testing. In the future, the focus will be extended to cover

alternatives in the basic sciences more comprehensively. Accordingly, the latest method portraits feature alternatives to animal models in neurodegenerative processes, such as traumatic brain injuries, and Parkinson’s and Alzheimer’s disease.

Education in reliable search procedures

Surveying the relevant literature is part of the daily business in basic science and medicine; however, courses on information retrieval are not obligatory in scientific education. In cooperation with the universities and the Regional Authority for Health and Social Matters of Berlin (LAGeSo), ZEBET@BfR contributes to training courses on “Laboratory Animals, Animal Experiments and Alternatives”. This course is certified by the German Society for Laboratory Animal Science (GV-SOLAS) and is attended by some 200–300 scientists involved in animal experimentation per year. The ZEBET part covers topics of information retrieval like structured searching, choosing the most relevant information resources, conceiving of search terms, using operators and wildcards, index-term-based searching, and semantic search engines. A main goal of the course is to instruct participants in index-term-based (classification) searching strategies (Motschall and Falck-Ytter 2005), and 3R-relevant terms.

In addition, ZEBET cooperates with ECVAM of the European Commission’s (EC) Joint Research Centre (JRC) in developing a search guide primarily to support scientists, regulators, and ethical committees involved in the planning, ethical review, authorization, and conduct of animal experiments.

Research in retrieval technology

Search engines that integrate human expert domain knowledge are a subgroup of “semantic search engines”. They aim to gather the meaning of natural language documents or phrases from the occurrence and co-occurrence of certain terms and their synonyms within the text of a document and thus assist scientists in retrieving and sorting relevant domain literature.

The Go3R tool aids in retrieving 3R-relevant literature from PubMed (Fig. 14) (Sauer et al. 2009). It is the worldwide first tool of its kind specially equipped with expert domain knowledge from the area of the 3Rs. This knowledge is captured within a so-called “ontology”, i.e., an extensive and detailed network of “concepts”, terms that are unambiguous identifiers of a scientific domain such as “gene chip” in the field of genetic high-throughput screening, or “humane endpoints” in the field of the 3Rs. When a user performs a search query with Go3R, the search engine compares the concepts of the ontology with

Fig. 13 The AnimAlt database hosted at DIMDI. AnimAlt-ZEBET is a full text database of evaluated alternative methods to animal experiments in biomedicine and related fields developed by ZEBET

Database Key:	ZT00
Short Description:	Alternative Methods to Animal Experiments
Coverage:	2002 to present
File Size:	138 (Status 03/2010)
Update Cycle:	Irregular
Type:	Full text database
Language:	English
Producer:	<input type="checkbox"/> Bundesinstitut für Risikobewertung (BfR) - ZEBET <input type="checkbox"/> http://www.bfr.bund.de

the vocabulary used in the retrieved documents (Sauer et al. 2009). It highlights relevant terms or their synonyms and uses them to arrange the documents within an “intelligent” directory of contents. The Go3R tool has been developed by scientists from the Technical University of Dresden and the Transinsight GmbH (Dresden, Germany) in cooperation with ZEBET, the chemical company BASF SE (Ludwigshafen, Germany), and the Scientific Consultancy—Animal Welfare (Neubiberg, Germany). The beta-Version is available online free of charge (Fig. 14; <http://www.Go3R.org>). To date, it contains about 17,000 concepts structured in 26 branches with biomedical headings, such as “Cell Culture Technology”, and “3Rs Methods in the Life Sciences”. A current effort to improve the Go3R tool by re-engineering and expansion is funded by the BMBF. An essential extension will be the inclusion of scientific resources other than PubMed. The ultimate aim is to open up resources that right now are non-accessible to an indispensability search due to their non-indexed and/or unstructured nature. Go3R thus may be used to assign 3R-relevant index-terms to unstructured natural language documents in retrospect automatically.

Advisory services for public authorities, ministries, and scientists

Over the last 20 years, ZEBET has developed an impressive expert knowledge on alternative methods. In line with the 3Rs principle, ZEBET’s main priority is to reduce the number of laboratory animals to the absolutely necessary

WHAT - the ontology intelligent table of contents used for rapid semantic sorting
 WHO - find out about the authors
 WHERE - find out about the locations
 WHEN - find out about the time

Fig. 14 The Go3R tool for retrieving 3R-relevant literature from PubMed. The Go3R tool has been developed by scientists from the Technical University Dresden and the Transinsight GmbH in cooperation with ZEBET, the chemical company BASF SE and the Scientific Consultancy—Animal Welfare and is funded by the German Federal Ministry of Education and Research (BMBF)

minimum, to refine experimental procedures and housing of animals, to minimize pain and suffering and, if possible, to replace animal experiments with alternative methods. ZEBET staff members share their knowledge with scientists, competent authority staff, animal welfarists, and other interested persons in a number of ways. Research institutes, competent authorities, and individuals can submit direct inquiries about alternative methods according to the 3Rs principle to ZEBET in addition to the aforementioned

AnimAlt-ZEBET database. ZEBET@BfR receives on average 450–600 individual inquiries per year. A profile of ZEBET's customers for the year 2008 is shown in Fig. 15.

The Federal Ministry of Food, Agriculture and Consumer Protection (BMELV) is responsible for consumer protection and animal welfare in Germany. ZEBET as part of the BfR, which is a higher federal authority, is directly reporting to the BMELV. Consequently, ZEBET advises the ministry's animal welfare division on all scientific questions of animal welfare in the context of laboratory animals. For instance, ZEBET experts repeatedly advised the BMELV in conjunction with the amendment of EU Directive 86/609 which regulates the handling of laboratory animals in the EU.

Furthermore, ZEBET scientists examine on request of the competent authorities of the federal states ("Länder") whether an animal experiment for which an application has been submitted is indispensable pursuant to the animal welfare act. For this assessment, more than 250 scientists at the BfR and also scientists at other federal institutions can be consulted in the evaluation. ZEBET scientists investigate whether the application reflects the latest scientific findings, whether alternative methods exist that can be used instead of the proposed animal experiment, whether the experimental design is statistically sound and, at the same time, whether the number of animals reduced to a minimum without compromising the objectives of the project. Additionally, scientists who develop or wish to establish new alternative methods in research institutes, universities, or industry also frequently approach ZEBET. Because of their many years of experience in the development, validation, and regulatory recognition of alternative methods, ZEBET scientists are able to judge whether a new alternative method is likely to be a suitable replacement for an internationally established animal experiment and how this goal might be achieved. ZEBET scientists are also sought as experts on the international level in research support programs for the development of alternative methods and for the judging of research prizes.

Furthermore, ZEBET scientists are very popular interview partners for the German media because of their comprehensive expert knowledge.

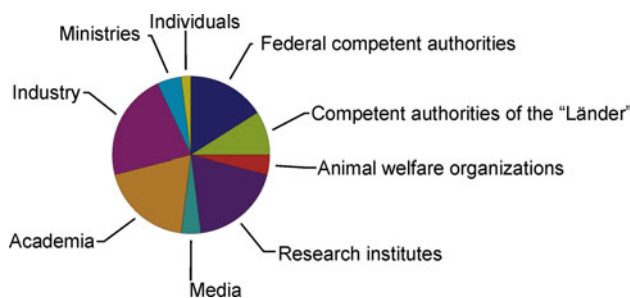


Fig. 15 Client profile of ZEBET@BfR in 2008

Regulatory challenges to be urgently addressed:
The example of Botulinum neurotoxin

Botulinum neurotoxin (BoNT) is a biological product released by the anaerobic, gram-positive bacterium *Clostridium botulinum*. There are seven major and serologically distinct serotypes of BoNT, characterized by a complex mode of action which results in a blockade of acetylcholine release at neuro-muscular nerve endings leading to flaccid paralysis (Fig. 16). This mechanism of toxicity comprises at least three stages of binding, internalization, and proteolytic cleavage of a substrate, which leads to inhibition of synaptic exocytotoxic transmitter release. Different serotypes of BoNT act on different substrates.

Currently, the serotypes BoNT-A and BoNT-B are used as active ingredients in licensed drugs for the treatment of a variety of medical disorders such as cervical dystonia, blepharospasm, spastic conditions, and hyperhidrosis. However, BoNT is also used in so-called "esthetic medicine" to temporarily treat facial asymmetries or reduce facial lines. BoNT's pharmacological activity is extraordinarily high which, due to its biological origin, varies from batch to batch. Therefore, it is not only a question of potency but also of drug safety that the biological activity of BoNT needs to be determined as accurately as possible. The monograph "Botulinum Toxin Type A for Injection" of the European Pharmacopoeia 7.2 (Council of Europe 2011) states that every production lot of BoNT has to be tested in an LD₅₀ potency test in mice (Fig. 16), where the final dilution series results in the highest dosing killing 90% of animals and the lowest dosing in at least 90% survival.

In the case of BoNT LD₅₀ potency testing, the dosing of animals is associated with severe suffering. Death is generally secondary to respiratory failure due to paralysis of the respiratory muscles. Consequently, the introduction of alternative methods is urgently required. A number of alternative tests have been developed in the past decades aiming to replace the mouse bioassay. Additionally, promising methods are under development, which may alone or in combination with other assays meet the rigid requirements of potency testing.

The basis for successful validation studies for alternative bioassays is the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) guideline Q2 (Validation of analytical procedures) (International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use 2005b). Since the LD₅₀ tests of the different BoNT-producing companies vary and are product specific (McLellan et al. 1996; Sesardic et al. 2003), the specific alternative methods have to be validated for each individual medicinal product. In this context, no

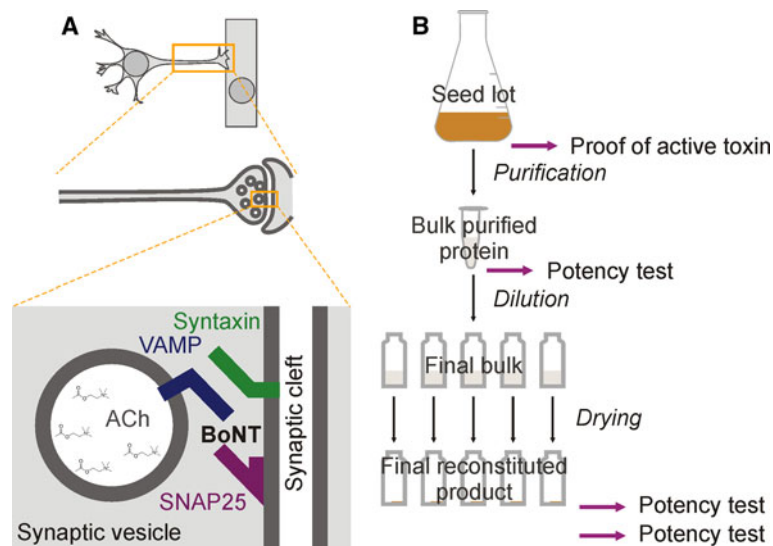


Fig. 16 Botulinum neurotoxin testing. **a** Botulinum neurotoxin acts at the molecular level by proteolysis of factors involved in vesicle docking at the neuromuscular junction. ACh, acetylcholine; BoNT, Botulinum neurotoxin; Syntaxin, VAMP, and SNAP25 are proteins involved in vesicle tethering at the plasma membrane which are the

targets of different Botulinum neurotoxins. **b** As described in the regarding monograph of the current European Pharmacopoeia, BoNT produced for medical applications has to be tested for potency at several steps during production

standard validation program can be applied, but individual validation plans have to be developed.

Ethical concerns have been raised in both Europe and the US about the animal suffering in BoNT potency testing, especially in the context of BoNT's cosmetic applications. As a result, the BMELV commissioned ZEBET to assess the status of the different existing alternative methods to the BoNT LD₅₀ potency test and the most promising approaches for their validation. In the following, an Expert Meeting on the “Current Scientific and Legal Status of Alternative Methods to the LD₅₀ Test for Botulinum Neurotoxin (BoNT) Potency Testing” was held at the BfR on April 27–28, 2009 (Adler et al. 2010). Experts from industry, regulatory authorities, German ministries, academia, national and international validation centers, and animal welfare organizations were invited to actively participate in the meeting.

During the meeting it became clear that guidance on product-specific validation of alternative methods to the LD₅₀ potency test needs to be given by the regulatory authorities in close communication with the manufacturers before and during the validation process. To facilitate validation efforts, international harmonization and mutual acceptance criteria of regulatory authorities are necessary. Importantly, funding should be made available and coordinated to develop and validate alternative assays for BoNT potency testing according to the 3Rs principles. Especially the development of a replacement alternative for BoNT testing should have priority.

At the expert meeting, researchers and industry could demonstrate some progress in implementing reduction and

refinement in BoNT potency testing, and in the development of alternative potency assays. Still, the majority of participants expressed the wish that a “BoNT Expert Working Group” (BoNT EWG) should be established in order to provide advice and guidance on validation requirements for proposed alternative methods and to define minimum standards in order to implement the 3Rs in BoNT potency testing. Furthermore, the BoNT EWG should promote awareness and transparency between the stakeholders and regulatory authorities. By request of the participants of the meeting, ZEBET@BfR in collaboration with the Federal Institute for Drugs and Medical Devices (BfArM) accepted to function as chairs and coordinators of this working group.

The BoNT EWG comprises experts from European regulatory authorities, 3R-related and validation institutions, manufacturers, and scientists. Additionally, experts from overseas are invited as observers. It will meet regularly several times per year during a proposed time frame of 4 years. The BoNT EWG has been formally established and their statutes phrased and adopted in three meetings so far. However, the outcomes of the meetings remain confidential unless otherwise expressly agreed by all members.

Future perspectives

Although much has been achieved in the past 20 years, and in the 50 years since Russell and Burch's seminal book phrasing the 3Rs (Russell and Burch 1959), many areas involving experimental animals have been barely touched

Table 2 Current and future activities of ZEBET@BfR

<i>Current experimental activity at ZEBET</i>
Inhalation toxicity testing in vitro (CULTEX)
Go3R semantic search engine
Metabolic capacity of skin tissue models
Embryotoxicity assays and metabolically competent systems
In vitro developmental cardiotoxicity
In vitro developmental neurotoxicity
In vitro developmental osteotoxicity
<i>Future projects</i>
Increased research efforts regarding chronic toxicity endpoints
Boosting collaborations, national and international networking
Intensifying efforts at the OECD
Severity classification of procedures on animals
Establishment of a Refinement Center at ZEBET@BfR
Establishment of a National Reference Center for Alternatives to Animal Testing

by ZEBET's work, and new areas have developed with the advancement of science. Therefore, our future efforts have to go beyond improving and extending what is already established, both regulatorily and experimentally (cf. above; Table 2).

With the realization that testing and experimentation on animals will be indispensable for some time to come grows the importance of the aspect of refinement (Fig. 1; Table 3). BfR's capacity in animal experimentation should be employed to establish a "refinement center" at ZEBET@BfR. Research into refinement is still on the sideline, but it has been shown already that enriched environments for laboratory animals improve the reliability of test results (Cao et al. 2010; Glaser and Kiecolt-Glaser 2005). Again, ZEBET's ability in funding and fostering collaborations can play a pivotal role in enhancing this area. Most importantly, it would allow for training of veterinary surgeons and staff involved in animal experimentation, which will ideally become a mandatory part of their curriculum. In this respect, a severity classification of procedures on animals could be established in accordance to the new European Directive 2010/63/EU of the European Parliament and of the Council on the protection of animals used for scientific purposes which would function as a guide for people involved in all stages of planning, writing, and execution of animal experiments. In September 2010, the European Parliament adopted the revised Directive on the protection of animals used for scientific purposes. The new Directive will take effect on January 1, 2013 (Directive 2010). Until then, the Member States have time to implement the Directive 2010/63/EU into national legislation (Directorate General for the Environment 2010).

However, building on our strengths, such as boosting collaborations and networking in the national as well as the

Table 3 : The 3Rs in the field of transgenic animals

Implementation and enforcement of the 3Rs principle in the field of transgenic animals
Best practice conditions and animal welfare standards
Defined pure genetic backgrounds
Shared and cryo-archived models
Standardized characterization in Mouse Clinics
Phenotyping results in databases
Standardizing and connecting European databases
Phenotyping of existing models
Mouse line generation by most competent staff
Improved ethical review and biostatistics
Reduced distress in identification/genotyping methods
Indication of 3R relevant technical information in publications

international arena, and intensifying our efforts at the OECD level, will remain a central future aim. Also, the existing expertise in in vitro methodologies as alternatives to animal testing needs to be fairly expanded, and research efforts especially regarding chronic toxicity endpoints increased. By noting this, it becomes clear that without any doubt ZEBET's main tasks and focus currently lie and will remain to lie in the *replacement* and *reduction* of animal testing. In particular with regard to chronic toxicity endpoints, such research efforts are still ill funded and scattered throughout Germany and to increase exchange and collaboration a *National Reference Center for Alternatives to Animal Testing* should be established which could be hosted at ZEBET@BfR.

Acknowledgments The authors are grateful to all of the contributors and attendees of the 20th anniversary of ZEBET celebrated at BfR, Berlin, in October 2009.

Conflict of interest The authors declare that there are no conflicts of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Adler S, Bicker G, Bigalke H, Bishop C, Blümel J, Dressler D, Fitzgerald J, Gessler F, Heuschen H, Kegel B, Luch A, Milne C, Pickett A, Ratsch H, Ruhdel I, Sesardic D, Stephens M, Stiens G, Thornton PD, Thürmer R, Vey M, Spielmann H, Grune B, Liebsch M (2010) The current scientific and legal status of alternative methods to the LD50 test for botulinum neurotoxin potency testing. The report and recommendations of a ZEBET expert meeting. *Altern Lab Anim* 38:315–330
- Andersen ME, Krewski D (2009) Toxicity testing in the 21st century: bringing the vision to life. *Toxicol Sci* 107:324–330

- Archer G, Balls M, Bruner LH, Curren RD, Fentem JH, Holzhtutter HG, Liebsch M, Lovell DP, Southee JA (1997) The validation of toxicological prediction models. *Altern Lab Anim* 25:505–516
- Balls M (1997) Defined structural and performance criteria would facilitate the validation and acceptance of alternative test procedures. *Altern Lab Anim* 25:483–484
- Balls M, Blaauboer B, Brusick D, Frazier J, Lamb D, Pemberton M, Reinhardt C, Roberfroid M, Rosenkranz H, Schmid B, Spielmann H, Stammati A-L, Walum E (1990) Report and recommendations of the CAAT/ERGATT workshop on the validation of toxicity test procedures. Amden report. *Altern Lab Anim* 18:313–337
- Balls M, Blaauboer BJ, Fentem JH, Bruner L, Combes RD, Ekwall B, Fielder RJ, Guillouzo A, Lewis RW, Lovell DP, Reinhardt CA, 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. *Altern Lab Anim* 23:129–147
- Box RJ, Spielmann H (2005) Use of the dog as non-rodent test species in the safety testing schedule associated with the registration of crop and plant protection products (pesticides): present status. *Arch Toxicol* 79:615–626
- Buesen R, Genschow E, Slawik B, Visan A, Spielmann H, Luch A, Seiler A (2009) Embryonic stem cell test remastered: comparison between the validated EST and the new molecular FACS-EST for assessing developmental toxicity in vitro. *Toxicol Sci* 108:389–400
- Bundesministerium für Bildung und Forschung (2001) Hightech statt Tiere: Ersatz- und Ergänzungsmethoden zu Tierversuchen. http://www.bmbf.de/pub/hightech_statt_tiere.pdf
- Cao L, Liu X, Lin EJ, Wang C, Choi EY, Riban V, Lin B, During MJ (2010) Environmental and genetic activation of a brain-adipocyte BDNF/leptin axis causes cancer remission and inhibition. *Cell* 142:52–64
- Carbone L (2004) What animals want; expertise and advocacy in laboratory animal welfare policy. Oxford University Press, New York
- Cecchelli R, Berezowski V, Lundquist S, Culot M, Renftel M, Dehouck MP, Fenart L (2007) Modelling of the blood-brain barrier in drug discovery and development. *Nat Rev Drug Discov* 6:650–661
- Commission of the European Communities (2007) Fifth report on the statistics on the number of animals used for experimental and other scientific purposes in the member states of the European Union, SEC(2007)1455. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:52007DC0675:EN:NOT>
- Council of Europe (2005) European convention for the protection of vertebrate animals used for experimental and other scientific purposes; CETS no.: 123. Council of Europe. <http://conventions.coe.int/Treaty/Commun/QueVoulezVous.asp?NT=123&CM=1&DF=20/10/2010&CL=ENG>
- Council of Europe (2011) Botulinum toxin type A for injection. Stationary Office Books, Strasbourg
- de Jong E, Louisse J, Verwei M, Blaauboer BJ, van de Sandt JJ, Woutersen RA, Rietjens IM, Piersma AH (2009) Relative developmental toxicity of glycol ether alkoxy acid metabolites in the embryonic stem cell test as compared with the in vivo potency of their parent compounds. *Toxicol Sci* 110:117–124
- Devolder T, Reid K, Rogiers V, Webb S, Wilkins D (2008) A review of national public funding programmes in European countries. *ALTEX* 25:233–242
- Directive 2010/63/EU of the European parliament and of the council of 22 September 2010 on the protection of animals used for scientific purposes. Official Journal of the European Union. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:276:0033:0079:EN:PDF>
- Directorate General for the Environment (2010) Revision of directive 86/609/EEC on the protection of animals used for experimental and other scientific purposes. http://ec.europa.eu/environment/chemicals/lab_animals/nextsteps_en.htm
- Dix DJ, Houck KA, Martin MT, Richard AM, Setzer RW, Kavlock RJ (2007) The ToxCast program for prioritizing toxicity testing of environmental chemicals. *Toxicol Sci* 95:5–12
- Döhmer J (2001) Modern drug development by molecular- and cell-biological methods. *ALTEX* 18:9–12
- Emmerich C (2009) Comparing first level patent data with value-added patent information: a case study in the pharmaceutical field. *World Pat Inf* 31:117–122
- European Center for the Validation of Alternative Methods (2006) Statement on the application of the Skinethic™ human skin model. http://ecvam.jrc.it/publication/ESAC25_statement_SKINETHIC_correction_on181206_C.pdf
- European Chemicals Agency (2009) New study inaccurate on the number of test animals for REACH. http://echa.europa.eu/doc/press/pr_09_11_animal_testing_20090828.pdf
- European Patent Office (2009) Business use of patent information. European Patent Office. <http://www.epo.org/patents/patent-information/business.html>
- Genschow E, Spielmann H, Scholz G, Pohl I, Seiler A, Clemann N, Bremer S, Becker K (2004) Validation of the embryonic stem cell test in the international ECVAM validation study on three in vitro embryotoxicity tests. *Altern Lab Anim* 32:209–244
- Gerner I, Schlegel K, Walker JD, Hulzebos E (2004) Use of physicochemical property limits to develop rules for identifying chemical substances with no skin irritation or corrosion potential. *QSAR Comb Sci* 23:726–733
- Giese C, Demmler CD, Ammer R, Hartmann S, Lubitz A, Miller L, Müller R, Marx U (2006) A human lymph node in vitro—challenges and progress. *Artif Organs* 30:803–808
- Glaser R, Kiecolt-Glaser JK (2005) Stress-induced immune dysfunction: implications for health. *Nat Rev Immunol* 5:243–251
- Halle W (2003) The registry of cytotoxicity: toxicity testing in cell cultures to predict acute toxicity (LD₅₀) and to reduce testing in animals. *Altern Lab Anim* 31:89–198
- Hansen S, Henning A, Naegel A, Heisig M, Wittum G, Neumann D, Kostka KH, Zbytovska J, Lehr CM, Schaefer UF (2008) *In-silico* model of skin penetration based on experimentally determined input parameters. Part I: experimental determination of partition and diffusion coefficients. *Eur J Pharm Biopharm* 68:352–367
- Hartung T, Aaberge I, Berthold S, Carlin G, Charton E, Coecke S, Fennrich S, Fischer M, Gommer M, Halder M, Haslov K, Jahnke M, Montag-Lessing T, Poole S, Schechtman L, Wendel A, Werner-Felmayer G (2001) Novel pyrogen tests based on the human fever reaction. The report and recommendations of ECVAM workshop #43. European Centre for the Validation of Alternative Methods. *Altern Lab Anim* 29:99–123
- Helma C (2006) Lazy structure-activity relationships (LAZAR) for the prediction of rodent carcinogenicity and Salmonella mutagenicity. *Mol Divers* 10:147–158
- Höfer T, Gerner I, Gundert-Remy U, Liebsch M, Schulte A, Spielmann H, Vogel R, Wettig K (2004) Animal testing and alternative approaches for the human health risk assessment under the proposed new European chemicals regulation. *Arch Toxicol* 78:549–564
- Hoffman JA, Merrill BJ (2007) New and renewed perspectives on embryonic stem cell pluripotency. *Front Biosci* 12:3321–3332
- Holzhtutter HG (1997) A general measure of in vitro phototoxicity derived from pairs of dose-response curves and its use for predicting the in vivo phototoxicity of chemicals. *Altern Lab Anim* 25:445–462

- Holzhütter HG, Genschow E, Diener W, Schlede E (2003) Dermal and inhalation acute toxic class methods: test procedures and biometric evaluations for the Globally harmonized classification system. *Arch Toxicol* 77:243–254
- Hulzebos E, Walker JD, Gerner I, Schlegel K (2005) Use of structural alerts to develop rules for identifying chemical substances with skin irritation or skin corrosion potential. *QSAR Comb Sci* 24:332–342
- Interagency Coordinating Committee on the Validation of Alternative Methods (1999) Corrositex[®]: an in vitro test method for assessing dermal corrosivity potential of chemicals. The results of an independent peer review evaluation coordinated by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM). National toxicology program. <http://iccvam.niehs.nih.gov/docs/reports/corprep.pdf>
- International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (2005a) ICH harmonised tripartite guideline: detection of toxicity to reproduction for medicinal products and toxicity to male fertility S5(R2). ICH Secretariat. <http://www.ich.org/LOB/media/MEDIA498.pdf>
- International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (2005b) ICH harmonised tripartite guideline: validation of analytical procedures: text and methodology, Q2(R1). ICH Secretariat. <http://www.ich.org/LOB/media/MEDIA417.pdf>
- Kandarova H, Liebsch M, Spielmann H, Genschow E, Schmidt E, Traue D, Guest R, Whittingham A, Warren N, Gamer AO, Remmele M, Kaufmann T, Wittmer E, De Wever B, Rosdy M (2006) Assessment of the human epidermis model SkinEthic RHE for in vitro skin corrosion testing of chemicals according to new OECD TG 431. *Toxicol In Vitro* 20:547–559
- Kretlow A, Butzke D, Götz ME, Grune B, Halder M, Henkler F, Liebsch M, Nobiling R, Oelgeschläger M, Reifenberg M, Schäfer B, Seiler A, Luch A (2010) Implementation and enforcement of the 3Rs principle in the field of transgenic animals used for scientific purposes. Report and recommendations of the BfR expert workshop, May 18–20, 2009, Berlin, Germany. *ALTEX* 27:117–123
- Krewski D, Andersen ME, Mantus E, Zeise L (2009) Toxicity testing in the 21st century: implications for human health risk assessment. *Risk Anal* 29:474–479
- Krewski D, Acosta D Jr, Andersen M, Anderson H, Bailar JCI, Boekelheide K, Brent R, Chamley G, Cheung VG, Green S Jr, Kelsey KT, Kerkvliet NI, Li AA, McCray L, Meyer O, Patterson RD, Pennie W, Scala RA, Solomon GM, Stephens M, Yager J, Zeise L (2010) Toxicity testing in the 21st century: a vision and a strategy. *J Toxicol Environ Health B Crit Rev* 13:51–138
- Lamb J (2007) The connectivity map: a new tool for biomedical research. *Nat Rev Cancer* 7:54–60
- Linke K, Schanz J, Hansmann J, Walles T, Brunner H, Mertsching H (2007) Engineered liver-like tissue on a capillarized matrix for applied research. *Tissue Eng* 13:2699–2707
- McLellan K, Das RE, Ekong TA, Sesardic D (1996) Therapeutic botulinum type A toxin: factors affecting potency. *Toxicol* 34:975–985
- Mess A (2007) The Guinea pig placenta: model of placental growth dynamics. *Placenta* 28:812–815
- Motschall E, Falck-Ytter Y (2005) Searching the MEDLINE literature database through PubMed: a short guide. *Onkologie* 28:517–522
- Organisation for Economic Co-Operation, Development (1996) OECD workshop on harmonisation of validation and acceptance. *Altern Lab Anim* 24:7
- Organisation for Economic Co-Operation, Development (2002) Final report of the OECD workshop on harmonization of validation and acceptance criteria for alternative toxicological test methods. OECD Publications Office, Paris
- Pampaloni F, Reynaud EG, Stelzer EH (2007) The third dimension bridges the gap between cell culture and live tissue. *Nat Rev Mol Cell Biol* 8:839–845
- Pampaloni F, Stelzer EH, Masotti A (2009) Three-dimensional tissue models for drug discovery and toxicology. *Recent Pat Biotechnol* 3:103–117
- Paquette JA, Kumpf SW, Streck RD, Thomson JJ, Chapin RE, Stedman DB (2008) Assessment of the embryonic stem cell test and application and use in the pharmaceutical industry. *Birth Defects Res B Dev Reprod Toxicol* 83:104–111
- Piersma AH (2006) Alternative methods for developmental toxicity testing. *Basic Clin Pharmacol Toxicol* 98:427–431
- Riebeling C, Pirow R, Becker K, Buesen R, Eikel D, Kaltenhäuser J, Meyer F, Nau H, Slawik B, Visan A, Volland J, Spielmann H, Luch A, Seiler A (2011) The embryonic stem cell test as tool to assess structure-dependent teratogenicity: the case of valproic acid. *Toxicol Sci* 120:360–370
- Rovida C, Hartung T (2009) Re-evaluation of animal numbers and costs for in vivo tests to accomplish REACH legislation requirements for chemicals—a report by the transatlantic think tank for toxicology (t(4)). *ALTEX* 26:187–208
- Russell WMS, Burch RL (1959) The principles of humane experimental technique. Methuen & Co. Ltd., London
- Sauer UG, Wächter T, Grune B, Doms A, Alvers MR, Spielmann H, Schroeder M (2009) Go3Rsemantic internet search engine for alternative methods to animal testing. *ALTEX* 26:17–31
- Schäfer-Korting M, Mahmoud A, Lombardi BS, Brüggener B, Kleuser B, Schreiber S, Mehnert W (2008) Reconstructed epidermis and full-thickness skin for absorption testing: influence of the vehicles used on steroid permeation. *Altern Lab Anim* 36:441–452
- Seiler A, Visan A, Buesen R, Genschow E, Spielmann H (2004) Improvement of an in vitro stem cell assay for developmental toxicity: the use of molecular endpoints in the embryonic stem cell test. *Reprod Toxicol* 18:231–240
- Seiler AE, Buesen R, Visan A, Spielmann H (2006a) Use of murine embryonic stem cells in embryotoxicity assays: the embryonic stem cell test. *Methods Mol Biol* 329:371–395
- Seiler AE, Buesen R, Hayess K, Schlechter K, Visan A, Genschow E, Slawik B, Spielmann H (2006b) Current status of the embryonic stem cell test: the use of recent advances in the field of stem cell technology and gene expression analysis. *ALTEX* 22:346–352
- Sesardic D, Leung T, Gaines DR (2003) Role for standards in assays of botulinum toxins: international collaborative study of three preparations of botulinum type A toxin. *Biologicals* 31:265–276
- Spielmann H, Gerbracht U (2001) The use of dogs as second species in regulatory testing of pesticides. Part II: subacute, subchronic and chronic studies in the dog. *Arch Toxicol* 75:1–21
- Spielmann H, Pohl I, Döring B, Liebsch M, Moldenhauer F (1997) The Embryonic stem cell test, an in vitro embryotoxicity test using two permanent mouse cell lines: 3T3 fibroblasts and embryonic stem cells. *In Vitro Mol Toxicol J Basic Appl Res* 10:119–127
- Spielmann H, Seiler A, Bremer S, Hareng L, Hartung T, Ahr H, Faustman E, Haas U, Moffat GJ, Nau H, Vanparys P, Piersma A, Sintes JR, Stuart J (2006) The practical application of three validated in vitro embryotoxicity tests. The report and recommendations of an ECVAM/ZEBET workshop (ECVAM workshop #57). *Altern Lab Anim* 34:527–538
- Stolper G, Klausner M, Sheasgreen J, Hayden P (2005) Development of an in vitro blood-brain barrier model for brain disposition screening of pharmaceuticals. *Toxicologist* 84:257

- US Congress, Office of Technology Assessment (1986) Alternatives to animal use in research, testing, and education (OTA-BA-273). US Government Printing Office, Washington
- US Department of Agriculture, Animal and Plant Health Inspection Service (2009) Annual report animal usage by fiscal year. US Department of Agriculture. http://www.aphis.usda.gov/animal_welfare/efoia/downloads/2009_Animals_Used_In_Research.pdf
- van der Jagt K, Munn S, Tørsløv J, de Bruijn J (2004) Alternative approaches can reduce the use of test animals under REACH. http://ecb.jrc.ec.europa.eu/documents/REACH/PUBLICATIONS/Reducing_the_use_of_test_animals_under_REACH_IHCP_report.pdf
- Verwei M, van Burgsteden JA, Krul CA, van de Sandt JJ, Freidig AP (2006) Prediction of in vivo embryotoxic effect levels with a combination of in vitro studies and PBPK modelling. *Toxicol Lett* 165:79–87
- Vickaryous MK, Hall BK (2006) Human cell type diversity, evolution, development, and classification with special reference to cells derived from the neural crest. *Biol Rev Camb Philos Soc* 81:425–455
- Walker JD, Gerner I, Hulzebos E, Schlegel K (2004) (Q)SARs for predicting skin irritation and corrosion: mechanisms, transparency and applicability of predictions. *QSAR Comb Sci* 23:721–725
- Walker JD, Gerner I, Hulzebos E, Schlegel K (2005) The skin irritation corrosion rules estimation tool (SICRET). *QSAR Comb Sci* 24:378–384
- Whitlow S, Burgin H, Clemann N (2007) The embryonic stem cell test for the early selection of pharmaceutical compounds. *ALTEX* 24:3–7
- zur Nieden NI, Kempka G, Ahr HJ (2004) Molecular multiple endpoint embryonic stem cell test—a possible approach to test for the teratogenic potential of compounds. *Toxicol Appl Pharmacol* 194:257–269