Toxicology for the twenty-first century

Thomas Hartung

The testing of substances for adverse effects on humans and the environment needs a radical overhaul if we are to meet the challenges of ensuring health and safety.

Synthetic chemicals have been components of consumer products for just over a century. A system for identifying which chemicals pose a danger to individuals and the environment was first put in place about 80 years ago. But after several productive decades, in which a patchwork of testing approaches was formed, fewer and fewer of the latest scientific developments were incorporated. The system of regulatory toxicology fell asleep, much like the fairy-tale character Snow White when she bit into the poisonous apple. In the case of toxicology, the poison was international guidelines. This international harmonization was tempting because it allowed manufacturers and suppliers to use fewer resources, and it overcame barriers to trade in global markets. But implementing these guidelines came at a price: the slow and complicated international consensus process hindered self-criticism and modernization of the field of toxicology.

There is almost no other scientific field in which the core experimental protocols have remained nearly unchanged for more than 40 years. Yet consumers continually increase their expectations about the safety of products. One recent effect of this was the instigation of the largest safety assessment of chemicals that has ever been carried out: the European Union introduced the regulation known as Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) by legislation in 2007. Whereas new chemicals have been systematically evaluated in the European Union and the United States for about a quarter of a century, the safety of any chemicals produced before 1981 (which includes 97% of the major chemicals in use, and more than 99% of chemicals produced by volume) has not necessarily been properly addressed. In fact, it is estimated that data for 86% of the chemicals are lacking, and the REACH process seeks to redress this. The regulation affects 27,000 companies, which are required to provide information on the toxic properties and uses of 30,000 chemicals, after a pre-registration phase in 2008. But REACH might turn out to be like the prince whose kiss awoke Snow White after a long sleep, rousing toxicology at last.

Defining the problem

So what is wrong with the current approach to toxicology testing? An ideal study to understand whether an agent is harmful to humans would require an extremely large number of human subjects who are representative of the diversity of humans and who are unknowingly exposed to the agent under realistic conditions. All possible effects should then be assessed. If there is any deviation from these experimental conditions, which are unrealistic and unethical, the study will provide only an approximation of the real situation — it is a model. The crucial question therefore is how useful are the current models, which are mostly animal models, and how incorrect are they? Given that about €10 billion (US$14 billion) is spent on animal experimentation worldwide every year (about €2 billion of which is for toxicological studies), and given that more than 100 million experimental animals are used[1] and that products worth €5.6 trillion are regulated by such testing, the question is certainly appropriate. It encompasses four main issues.

The first issue is the extent to which animal models reflect human responses. It is clear that the use of animals has limitations[2]: we are not 70 kg rats; we take up substances differently; we metabolize them differently; we live longer (allowing certain diseases to develop and prompting evolutionary adaptations to protect against them); and we are exposed to a multitude of environmental factors. However, few studies have systematically measured the accuracy of animal models. In one example, results from animal models were compared with information from poison centres: comparing the dose of a chemical that is lethal to 50% (LD_{50}) of rats tested and the lethal concentration of the same chemical in the blood of humans showed a rather poor correlation (coefficient of correlation of 0.56; unpublished observations from an international validation study[3]). Similarly, in another study, 40% of the chemicals that irritated the skin of rabbits were found not to be irritants in the skin 'patch test' in humans[4].

Given the overall lack of data, this problem can be considered in more general terms by looking at how one species models for another. In several animal species, similar experiments with the same agents have been carried out, and there is no reason to assume that, for example, mice, rats and rabbits predict each other's response to a lesser extent than they predict that of humans. Typical results from such studies show agreement between animal species for 53–60% of chemicals[5,6].

Similar results have also been obtained for pharmaceuticals (as opposed to chemicals) that have been tested in humans. In one study, 43% of toxic effects in humans were correctly predicted by tests in rodents, and 63% by tests when non-rodent animals were also included[7]. It is clear therefore that many adverse effects are not uncovered by such traditional tests. This is also evident in data from the pharmaceutical industry, showing that 20% of the failure of drug candidates occurs as a result of toxicity only after the drugs have been administered to humans in clinical trials[8]. And it is estimated that 6.7% of hospitalized patients experience unexpected adverse reactions to drugs (1 in 20 of which are fatal)[9], showing the limitations of anticipating toxic effects from preclinical animal studies. To improve the toxicity assessment, tests are often carried out in two animal species: usually substances that show no toxic effect in one species are then tested in another species to improve the likelihood of finding any toxic properties. This increases the sensitivity of testing (that is, it increases the proportion of toxic substances that are found) but at the cost of increasing the number of false positives (when non-toxic chemicals seem to be toxic in the tests carried out).

The second key issue facing animal testing relates to the study design, particularly to the highly precautionary (conservative) approach that is taken at present. To limit costs and animal numbers, toxicity testing is typically carried out with the maximum dose of the chemical that can be tolerated, which has previously been determined. Such doses can be more than 1,000-fold higher than the doses intended for humans (in terms of milligrams per kilogram body weight, for example). This strategy yields many false positives and further diminishes the correlation between findings in animal models and humans[10].

The third issue is the testing of multiple end points, which also contributes to false-positive results. When enough end points are studied, positive associations will always be found. This is elegantly illustrated by a study that searched for correlations between disease and zodiac sign in the health records of 10 million residents of Ontario, Canada[11]: those born under the sign of Leo had a significantly higher probability of developing a gastric haemorrhage than individuals of other zodiac signs, and Sagittarians had far more fractures of the humerus over the period analysed. The explanation for this is simple: a total of 223 medical conditions were studied in a single population, and examining so many variables inevitably results in some extreme clustering of random results. Similarly, in toxicological studies, a large number of end points are measured: about 40 in repeat-dose toxicity studies (long-term studies in which animals are exposed to a chemical for a month to a year, and the effects on many organs are studied); and 80 in reproductive toxicity studies (in which adverse effects on the reproductive system, from fertility to embryonic malformations, are analysed). Unavoidably, some end points will be positive, and the group sizes used are too small to allow statistical correction for this.

Given that the cost of current tests is several hundred thousand euros per substance, large increases in group sizes are unrealistic, in addition to this being undesirable from the perspective of animal welfare. Therefore, all positive results have to be recorded as true positive results. However, this is less undesirable in the risk-assessment process than one might think, because the positive findings are simply used to establish the 'lowest observed effect level' (that is, the smallest amount of a substance that causes an observable change in the organism being studied). But, because the maximum tolerated dose is being used, there is usually a large safety margin (typically a factor of 100), so the substance could still be used even if seems to be toxic at high doses (with the exception of chemicals observed to have tumour-inducing properties, as these effects are generally thought to be relevant at much lower doses). It is thus often not important whether a positive result is an artefact. It is relevant, however, to those who later need to reproduce the presumed organ toxicity to validate an alternative approach, because false-positive results are difficult to reproduce when a different test is used. In addition, whether or not a positive result is false is unlikely to be noticed, because most regulatory tests are carried out only once[12] and because toxicological studies are often not reported publicly[13]. So the self-corrective mechanisms of science are not in place: there is no cross-referencing between similar studies in different laboratories.

The fourth issue concerns the prevalence of chemical effects on health. In other words, how many chemicals actually have hazardous properties[14]? Despite the use of highly precautionary tests, more than 87% of chemicals registered as new chemicals over the last 25 years are not acutely toxic in current tests; 93% of them do not irritate the skin[15]; and only 2–3% impair the reproductive cycle[16].

So toxicological studies search for a rare hazard with imperfect models. What are the consequences of this?

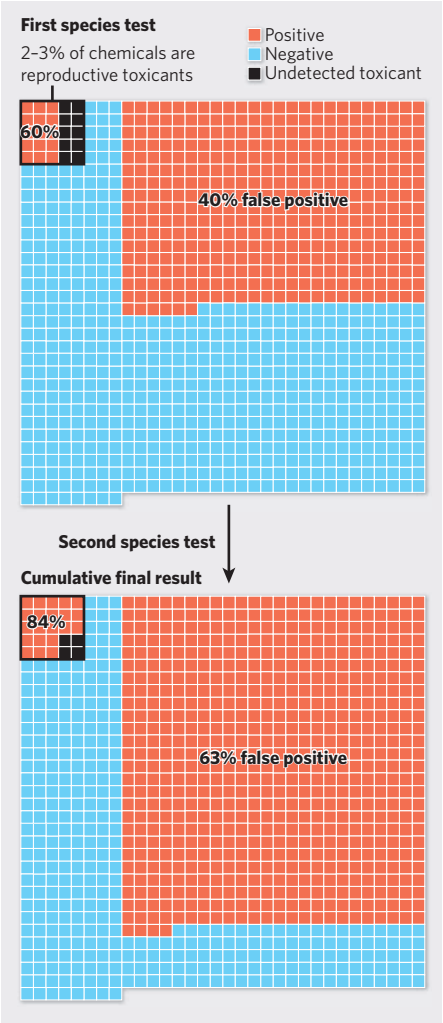

Figure 1 | The consequences of searching for rare hazards using imperfect tests. For reproductive toxicity testing, the concordance between animal species is about 60%. So roughly 40% of non-toxic chemicals will yield false-positive results. This problem is compounded by the standard practice of testing these chemicals that yield negative results in a second species to increase the number of hazardous substances identified.

False-positive issue

Take, for example, reproductive toxicity testing under the REACH legislation. All chemicals that were marketed before 1981 and are produced at more than 100 tonnes per year in the European Union will be subject to testing: about 5,500 of the 30,000 chemicals covered by REACH. It is estimated[16] that about 2.5% of these (138 substances) are true reproductive toxicants in humans (Fig. 1), and the goal of toxicological testing is to identify these. The reported concordance between species is about 60% for reproductive toxicity testing, using the two-generation study in rats (in which toxic effects are followed not only in the offspring of exposed rats but also, after further mating, in the next generation). Between animals and humans, however, this concordance might be even lower, owing to the high-dose, precautionary approach. So, when testing 5,500 chemicals with a test that is 60% accurate, 83 of the 138 reproductive toxicant will be found, but about 2,145 substances (almost 40%) will yield a false-positive result. The standard procedure would then be to test the apparently non-toxic substances in another animal species. Given the same accuracy, in rabbits or mice, 40% of the 3,272 substances that showed negative results in the first test (1,309 chemicals) will test as false positives. At the same time, 60% of the 55 true toxicants (33 chemicals) that were missed in the first test, in rats, will be found.

In total, 116 of the 138 true reproductive toxicants (84%) will be found, and 3,454 non-toxic chemicals will be found to be toxic (a total of 63% false-positive findings). These results might therefore restrict the use of a large number of these substances, which are subject to testing because they are produced in the highest quantities in Europe[17]. This scenario might be difficult to believe, but an analysis of reproductive toxicity studies for chemicals between 1981 and 2007 confirms this[16]: in 27 years, 72 chemicals reached a production volume that triggered reproductive toxicity tests. Of these, 41 (57%) tested positive, as the above calculation (of 63%) would suggest.

There are several caveats though. The above scenario might be too pessimistic because the correlations between species are biased by the inclusion of more chemicals that test positive in at least one species (because, in the past, a second test was often carried out to challenge the result). In addition, triggers others than production volume might have indicated the need for testing a substance: that is, if substances are tested because they are suspected of being toxic to the reproductive cycle, then this biases the number of positive results in the database. Nonetheless, it is unlikely that we can afford to falsely assign a large proportion of high-production-volume chemicals as reproductive toxicants. This will unnecessarily restrict the use of many substances, require large and expensive efforts to replace chemicals that are widely used, and create unnecessary fears in consumers about previous exposure. It might also prompt a situation similar to that for pharmaceuticals: if such results are obtained for a drug that is in the late stages of development (when it is already certain that the drug has financial value), then large amounts of toxicological work are required to determine whether the animal studies are in any way relevant to humans so that a valuable substance can be saved.

Another important issue is that the tests for each chemical require an average of 3,200 animals for a single two-generation test[17] — a total of 17.6 million animals for 5,500 substances — and the current REACH testing guidance

for industry does not include much scope for waivers or alternatives. Even if the use of alternatives to animal studies, such as cell-culture-based testing, were feasible, such methods do not have fewer limitations[18], except for ethical ones. And particularly in the field of reproductive toxicity, alternative methods are only being developed[19], and the cost–benefit ratio of using these for large-scale screening programmes still needs to be established.

Towards a solution

It is unlikely that researchers will suddenly produce new tools and design new methods with great accuracy. The solution to using fewer animals and making better predictions in the mid-term is to design integrated testing strategies. At present, the typical process is to use a default animal test and then, in some cases, to use cell-culture and computer-based methods to define the mode of action of the toxin and to interpret and balance the results further. But the best opportunity to improve regulatory toxicology lies in strategies in which optimal use is first made of all existing information about a substance and structurally similar substances, and then information is gained by approaches that do not involve animal testing, leading to targeted animal testing only if necessary. Such strategies will ideally include decision points that depend on interim results. An example of such a strategy is shown in Fig. 2.

The simplest testing strategy would combine two different approaches, such as a screening approach (a method to identify 'suspicious' substances with less effort and allowing false-positive results) and a confirmatory one (which may be more sophisticated and specifically identifies hazards with higher certainty). All substances that test positive during the screening approach or another prioritization step would enter the confirmatory stage, which would consist of, for example, a battery of mechanistic tests examining relevant pathways of toxicity. Instead of testing a large number of substances that includes few true toxicants by using one definitive test, this new approach would increase the number of true positives entering the confirmatory stage by creating a subset of suspicious substances, offering more evidence about whether a chemical is hazardous than the screening test alone. Alternatively, analysing which end points (for example, which of the up to 80 end points measured in a reproductive toxicity study) actually lead to classification as toxic or non-toxic in a particular animal test might allow researchers to identify the end points for which dedicated tests are required[16].

Despite the advantages of such a change in approach, several difficulties are apparent. It would first require acknowledging and analysing the limitations of the current approach. One central problem here is that the current system is convenient for the key players: namely, the regulators and the regulated industry. At present, a clearly defined set of tests, which make predictable demands on time and costs, is carried out. Then, the books are closed, and industries' liability is minimized. Every integrated testing strategy with decision points in its course will bring this simple procedure to an end, making the uncertainties more evident, as well as the fact that only the probability of a particular hazard is assessed. Any shortcomings identified in current practices will mean that products need to be examined further and will open up liabilities again.

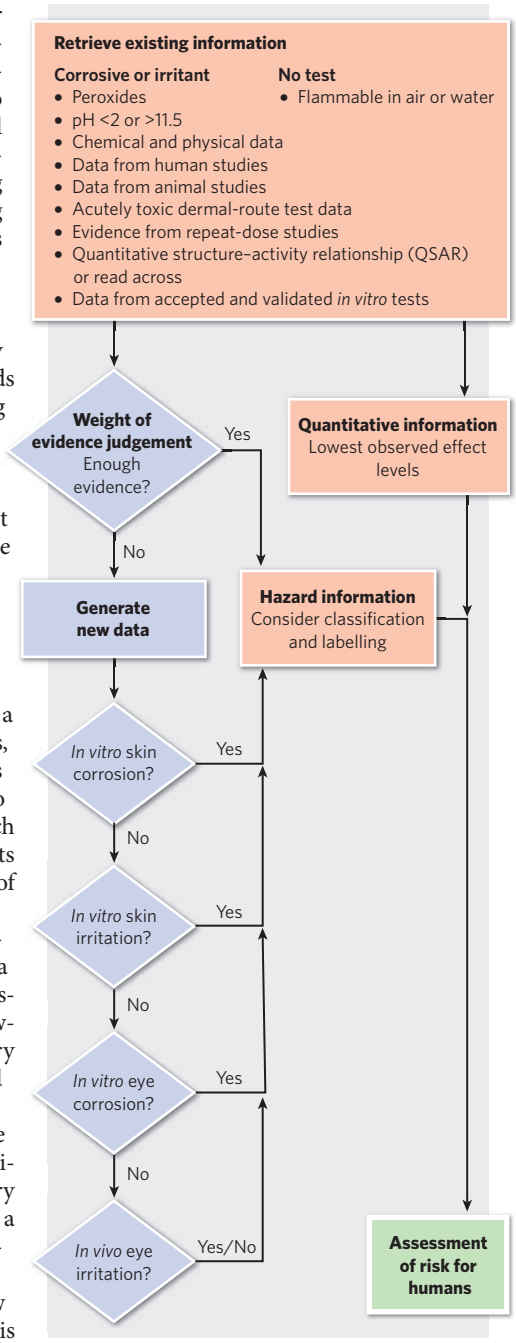

Figure 2 | Integrated testing strategy for eye and skin toxicity. This strategy from the REACH guidelines for industry is one of the first examples of an integrated testing strategy. The sequence includes decision points and involves assessing the existing information and then carrying out various *in vitro* tests, with animal tests being used only as a last resort (*in vitro* tests for eye irritation are being validated at present).

This situation is similar to that for pharmaceuticals. In fact, the current risk-assessment methodologies for chemicals are derived from those for preclinical studies of pharmaceuticals. However, for pharmaceuticals, there are two further steps in the process: clinical trials in humans and post-marketing surveillance (in which data are collected after a drug has been released onto the market). A considerable proportion of drug candidates (8–30%) fail because of safety problems in humans[20], despite having passed the entire toxicological programme of animal testing. Many of these safety issues are minor, for example nausea or a transient increase in the concentration of liver enzymes, but major chronic effects are not assessed at this stage. In addition, biologically active substances such as drugs often produce side effects as a result of their intended actions on human physiology (an effect known as 'excess pharmacology'); this is less of a problem for other areas of chemical use, in which the chemicals are not usually intended to affect the human body. But even though drugs undergo additional trials in human volunteers and patients, in my opinion there is always a need to follow up products after marketing, as illustrated by the anti-inflammatory drug Vioxx. Similarly, the possible hazards of chemicals in consumer products will probably need to be followed up more intensively after marketing.

Today, the pharmaceutical field is again driving changes in safety testing. With human proteins or antibodies (collectively known as biologicals) making up about half of the new drugs entering the market, classical toxicology is largely useless, because these proteins mostly have species-specific actions and animals raise antibodies to them, limiting the value of animal testing. This has created pressure to develop human-cell-based models for these biologicals, and other areas of toxicology will benefit from this. The inadequacy of current methods is also evident for new products such as genetically modified food and animal feed[21], functional food (food with intended health effects), and nanoparticles[22], creating an additional demand for new testing methods. Similarly, current methods are not tailored to assess the risk of acute poisonings associated with chemical accidents, or biological or chemical weapons[23].

REACH will also be a key instigator of change. This is partly because unexpected positive test results for important chemicals will trigger a review of the approaches — it is unlikely that important chemicals with decades of use will be abandoned easily, without raising doubts about the assessment. In addition, the legislation itself already represents a revolution in safety-assessment practices. Over the past three decades, internationally agreed (animal) testing guidelines have set out precisely how data must

be obtained, whereas REACH calls for the integrated use of all methodologies and for the use of animals as a last resort (with certain obstacles in place). So REACH calls for more flexibility and for tailored approaches. In terms of REACH, the test guidance for industry that has been developed in the past three years guides scientists through the combined use of existing data, and *in silico* (computer-based), *in vitro* and *in vivo* approaches. The greatest challenge will be to standardize these approaches in test guidelines and to reach international agreement on them. It is reasonable to assume that at least five times more guidelines will be necessary to accommodate the new approaches, an enormous challenge to the regulatory community.

But the challenge goes one step further: for each new method, test guidelines need to be not only agreed but also implemented. An interesting test case is the local lymph-node assay, which is used to predict whether topical application of a chemical to the skin will induce an allergic response. In 2002, the assay was internationally agreed by the Organisation for Economic Co-operation and Development (OECD) as the preferred animal model for studying skin allergies, but it has been seldom used until recently. Since 2002, less than 10% of new chemicals have been tested in this way, as indicated by notifications to European regulatory bodies. Applying a new method is hindered by, on the one hand, tradition and established practices and, on the other hand, obstacles such as the absence of international agreements with countries in important economic markets (for example, Brazil, Russia and China have not yet necessarily accepted the new OECD approaches).

International companies tend to use the traditional test until the last important market has accepted the new approach. So the banning of the original test method when alternatives become available is the prime opportunity to force a change. The OECD have only banned one test so far, however: the classical LD_{50} test, which required 45 rats for testing each substance, was abandoned in 2000, when three validated alternatives were introduced, requiring only 8 to 15 animals to test one substance. In other cases, the traditional animal tests have not been banned or modified when alternatives were introduced, so the original tests can still be carried out for regulatory purposes if justification is provided. But when a new approach does not suit all needs (that is, it is not appropriate for all chemicals or accepted by all member states), it is difficult to remove the traditional guidelines. The regulators must then urge that the new approach be used, to reinforce its implementation. For this to work, the advantages of the new test or the shortcomings of the old test need to be made evident, and to be credible, this assessment must have a sound and objective basis. The problem is that established practices have become intertwined with scientific insights during the decades in which toxicological tests have been shaped, and political compromises around such tests have been made.

Clinical medicine has a similar problem in that diagnostic and therapeutic approaches need to be objectively appraised so that the best decision can be made for each patient. Here too, new scientific approaches are interwoven with traditions, financial compromises in terms of health care, and so on. In the past couple of decades, the most important development in this area has been the evidence-based health-care movement, steered by the Cochrane collaboration[24]. Using structured reviews, consensus processes and meta-analyses, a series of 5,000 guidance documents has been developed. These provide the best available consolidation of the evidence in a particular field.

It is tempting to translate this evidence-based approach to toxicology[25], and a similar movement has been initiated. A realistic assessment of the methods used in toxicological studies will help to improve these tools and to integrate them into testing strategies. At the same time, it will be important to find ways to combine information from various studies, both systematically and quantitatively. The difficulties entailed are illustrated by the results of 29 independent risk assessments of the industrial solvent trichloroethylene: 6 studies deemed it non-carcinogenic; 10 found it to be carcinogenic in animals but unlikely to be carcinogenic in humans; 9 found it a plausible carcinogen in humans but with negative epidemiological findings; and 4 found it a plausible carcinogen in humans, with positive epidemiology[26].

Future visions

So it is clear that the current system of testing needs to change. Moreover, the individual testing tools have limitations and are inadequate for toxicology in the twenty-first century. To resolve this, I propose a three-step solution (Fig. 3). First, the limitations of the current tools need to be objectively assessed, and a better understanding of their uses is needed (for example, we need to analyse the prevalence of particular hazards because appropriate test strategies depend strongly on whether the hazard is rare or frequent). Second, in the mid-term, the various approaches need to be integrated into testing strategies, making the best use of the existing methods by combining them strategically. And, third, an entirely new system is urgently needed and should be built from scratch, using modern methods.

The basis for such a new system has emerged over the past two decades: advances in cell-culture techniques have enabled biological phenomena to be studied *in vitro*, unlike when toxicological experiments were first designed. In fact, most data generated in the life sciences now originate from studies of *in vitro* systems. This change in experimental approach required not only the accumulation of experience in these new techniques but also the provision of standardized equipment, materials and training. Early cell-culture-based experiments were

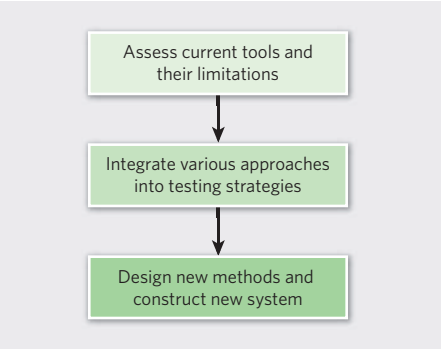

Figure 3 | Towards a new toxicology. The toxicology community needs to take three main steps to arrive at a new system of toxicology.

relatively simple, but they evolved rapidly, with many researchers now using three-dimensional ('organotypic') cultures that resemble organs in structure and function. Even one of the last big challenges in cell culture — the lack of availability of primary human cells (usually only sourced from surgically removed tissues with the notable exception of blood cells) — is now increasingly being overcome by isolating or generating human stem cells, from which most of the cell types in the body can be produced[27,28].

The avenue now opening for designing a new regulatory toxicology originates from the combination of bioinformatics and biotechnological approaches that yield huge amounts of information[29,30]. Three important technologies developed during the past decade have entered the field of toxicology[31,32]: 'omics' technologies (such as genomic and proteomic analyses), imaging techniques and robotized testing platforms. The testing platforms allow high throughput of samples, enabling large numbers of substances to be tested under standardized conditions. Omics technologies and imaging methods compile enormous sets of information about a single compound. Together, the three technologies not only allow researchers to 'fish' for new biological markers of specific toxic effects but also increasingly allow the deduction of patterns (or signatures) that are characteristic of certain toxic effects. By also harnessing advances in bioinformatics and *in silico* modelling, this information can be mined and then integrated with knowledge from other areas of the life sciences[33]. Such integration of information will be particularly important for investigating cellular pathways and should allow the cross-fertilization of ideas between toxicology and basic science[34]. The combination of biochemical knowledge of cellular pathways with genomics, proteomics and metabonomics (the study of metabolic responses to environmental factors, drugs and diseases) is already advancing as systems biology, and systems toxicology is a new subbranch of this field.

Such a systems approach was put forward as a toxicology for the twenty-first century in a 2007 report by the US National Academy of Sciences on behalf of the Environmental Protection

TABLE 1 | TOWARDS A NEW REGULATORY TOXICOLOGY

Scientific developments	Strategic developments
Mapping of pathways of toxicity by combining 'omics' technologies and data mining	Objective assessment of current practices (evidence-based toxicology)
Organotypic cell cultures and human tissues derived from stem cells	Guidance on Good Cell Culture Practice (an initiative for standardizing practices globally), Good Modelling Practices
Modelling of kinetics of substances (especially physiologically based pharmacokinetic modelling) in an organism for extrapolating from effective tissue concentrations to whole-organism doses	Systematic composition of testing strategies (mainly decision theory and sensitivity analysis)
In silico methods such as quantitative structure–activity relationship (QSAR) modelling	Validation of complex methodologies, in the absence of a gold standard
Imaging technologies and automated testing	Change management based on cost–benefit analysis
Integration of technologies	More communication

Agency (EPA)[35]. And this has already led to the formation of a coalition of US agencies[36] and to a revised toxicity testing strategy by the EPA[37], with the EPA's ToxCast programme[38] being closest in terms of vision to the new process required. So, in Europe, the prince who awakened toxicology was politics (with REACH calling for a new safety testing approach on a large scale, which was assisted by the animal testing ban in the seventh amendment of the cosmetics directive)[39]. By contrast, in the United States, it was science responding to the EPA's request for a new vision. What lies ahead, however, must be an entirely scientific process. Furthermore, the political process is necessary to make the funding available so that the political will can be put into practice. The dimensions of the project call for a global programme. The groundwork for such an effort has been laid by increasing awareness of the shortcomings of current methods, as well as emerging technological opportunities and political demands. The opportunity to create a new regulatory toxicology lies in a programme, similar to the Human Genome Project, that analyses the interactions of small molecules with cells. Such a programme will provide the molecular biological tools to switch cellular pathways on and off and to identify 'druggable' targets, and it will uncover the cellular pathways of toxicity, knowledge that is needed for a new way of approaching toxicology.

The main challenge is to design a new system of regulatory toxicology. Toxicology has grown step by step over a century to adapt to increasing and changing demands. Instead of amending the current patched-together system, a new system must be designed. And what is constructed from scratch with today's understanding and technologies will differ markedly from the current regulatory toxicology. We must forgo the approach that has been taken so far, which has been to add a new piece or to replace an old piece of the system: for example, by validating new tests that are each designed to substitute for a particular toxicological tool. This might deliver new ways to handle biologicals, nanoparticles and so on, but it will not solve the inherent shortcomings of an outdated architecture.

The necessary science seems to be available, but are the necessary scientists also available? Regulatory toxicology has not been the most appealing research area in the past. It could hardly compete for the best students with areas such as molecular biology, immunology and stem-cell biology, which have been advancing rapidly. However, science is invigorated when there are sufficient challenges and funding. And there is money available: I recently estimated that, in Europe, fulfilling toxicological regulations costs about €600 million per year for products that are traded at €1.7 trillion[40]. Generating data to comply with the REACH legislation will cost €8.8 billion using today's tests[40]. This is stimulus for a large number of targeted developments so that the process becomes better, quicker and cheaper, if scientists were only aware of this.

The scientific challenge laid down by this new vision of toxicology should appeal to scientists and to the commercial providers of solutions, mostly small-to-medium enterprises that are involved in commercializing new biotechnologies. Some key areas are listed in Table 1. But the challenge itself will not result in a new regulatory system. It will be important to open up regulators to the possibilities of a new system so that they give up on the old system and do not just use the new system as another patch for the old one (or as "useful additional information", as a regulator would say). For this revolutionary change to occur[41], the shortcomings of current methods need to be mapped and considered, and the transition from the old to the new approach needs to be steered[42]. This process needs to include standardization, validation and quality assurance of the new approaches, as well as the systematic integration of these approaches into testing strategies. There is a profusion of new concepts and technologies at present, but what is lacking is communication between stakeholders. Giving direction to the current stakeholders and to those in neighbouring disciplines who are not yet aware of the emerging opportunities, and allowing synergies to develop between approaches, might be even more important than the individual technological developments that are required. Promoting this process could be the real challenge for toxicology today. ■

Thomas Hartung is in the Department of Environmental Health Sciences at the Johns Hopkins University Bloomberg School of Public Health, and is Doerenkamp-Zbinden Professor and Chair for Evidence-based Toxicology. He is director of the Johns Hopkins Center for Alternatives to Animal Testing (CAAT), Baltimore, Maryland, as well as professor of pharmacology and toxicology at University of Konstanz, Germany.
e-mail: thartung@jhsph.edu

1. Taylor, K., Gordon,N., Langley, G. & Higgins, W. *Altern. Lab. Anim.* **36,** 327–342 (2008).
2. Hartung, T. *ALTEX* **25,** 3–9 (2008).
3. http://iccvam.niehs.nih.gov/docs/acutetox_docs/ATpanelrpt06/ ATpanelrpt.pdf
4. Basketter, D. A., York, M., McFadden, J. P. & Robinson, M. K. *Contact Dermatitis* **51,** 1–4 (2004).
5. Gottmann, E., Kramer, S., Pfahringer, B. & Helma, C. *Environ. Health Perspect.* **109,** 509–514 (2001).
6. Schardein, J. L., Schwetz, B. A. & Kenel, M. F. *Environ. Health Perspect.* **61,** 55–67 (1985).
7. Olson, H. et al. Regul. Toxicol. Pharmacol. **32,** 56–67 (2000).
8. Stevens, J. L. *Chem. Res. Toxicol.* **19,** 1393–1401 (2006).
9. Lazarou, J. *J. Am. Med. Assoc.* **279,** 1200–1205 (1998).
10. Rietjens, I. M. C. M. & Alink, G. M. *Chem. Res. Toxicol.* **19,** 977–981 (2006).
11. Austin, P. C., Mamdani, M. M., Juurlink, D. N. & Hux, J. E. *J. Clin. Epidemiol.* **59,** 964–969 (2006).
12. Purchase, I. F. H. *Toxicology* **202,** 1–20 (2004).
13. Wandall, B., Hansson, S. O. & Rudén, C. *Arch. Toxicol.* **81,** 605–617 (2007).
14. Hoffmann, S. & Hartung, T. *Toxicol. Sci.* **85,** 422–428 (2005).
15. Hoffmann, S., Cole, T. & Hartung, T. *Regul. Toxicol. Pharmacol.* **41,** 159–166 (2005).
16. Bremer, S., Pellizzer, C., Hoffmann, S., Seidle, T. & Hartung, T. *Curr. Pharm. Des.* **13,** 3047–3058 (2007).
17. Höfer, T. *et al. Arch. Toxicol.* **78,** 549–564 (2004).
18. Hartung, T. *ALTEX* **24,** 143–147 (2007).
19. Hareng, L., Pellizzer, C., Bremer, S., Schwarz, M. & Hartung, T. *Reprod. Toxicol.* **20,** 441–452 (2005).
20. Kola, I. & Landis, J. *Nature Rev. Drug Discov.* **3,** 711–715 (2004).
21. Kuiper, H. A. *Plant J.* **27,** 503–528 (2001).
22. Service, R. F. *Science* **321,** 1036–1037 (2008).
23. Wood, M., Pichard, A., Gundert-Remy, U., de Rooij, C. & Tissot, S. M. *J. Hazard. Mater.* **133,** 8–15 (2006).
24. http://www.cochrane.org/
25. Hoffmann, S. & Hartung, T. *Hum. Exp. Toxicol.* **25,** 497–513 (2006).
26. Hansson, S. O. & Ruden, C. *Toxicology* **218,** 100–111 (2006).
27. Davila, J. C. *et al. Toxicol. Sci.* **79,** 214–223 (2004).
28. Stummann, T. C. & Bremer, S. *Curr. Stem Cell Res. Ther.* **3,** 117–130 (2008).
29. Hartung, T. & Leist, M. *ALTEX* **25,** 91–96 (2008).
30. Leist, M., Hartung, T. & Nicotera, P. *ALTEX* **25,** 103–114 (2008).
31. MacGregor, J. T. *Toxicol. Sci.* **75,** 236–248 (2003).
32. Pognan, F. *Pharmacogenomics* **5,** 879–893 (2004).
33. Lotti, M. & Nicotera, P. *Nature* **416,** 481 (2002).
34. Orrhenius, S. & Zhitvotovsky, B. *Chem. Res. Toxicol.* **19,** 729–733 (2006).
35. Committee on Toxicity Testing and Assessment of Environmental Agents, Board on Environmental Studies and Toxicology, Institute for Laboratory Animal Research, National Research Council of the National Academies. *Toxicity Testing in the 21st Century: a Vision and a Strategy* (The National Academies Press, 2007).
36. Collins, F. S., Gray, G. M. & Bucher, J. R. *Science* **319,** 906–907 (2008).
37. http://www.epa.gov/osa/spc/toxicitytesting/
38. http://www.epa.gov/ncct/toxcast/
39. Hartung, T. *ALTEX* **25,** 147–162 (2008).
40. Bottini, A. A. & Hartung, T. *ALTEX* **26,** 3–16 (2008).
41. Hartung, T. *Altern. Lab. Anim.* **36,** 635–639 (2008).
42. Hartung, T. *Toxicol. Sci.* **109,** 18–23 (2009).