TOXICOLOGICAL SCIENCES **111(2)**, 233–237 (2009)
doi:10.1093/toxsci/kfp149
Advance Access publication July 17, 2009

FORUM

Are *In Vitro* Tests Suitable for Regulatory Use?

Thomas Hartung,*[,1,2] and George Daston,†[,‡,3]

*Johns Hopkins University, Bloomberg School of Public Health, Department of Environmental Health Sciences, Center for Alternatives to Animal Testing, Doerenkamp-Zbinden-Chair for Evidence-based Toxicology, Baltimore, Maryland 21205; †University of Konstanz, 78464 Konstanz, Germany; and ‡The Procter & Gamble Company, Cincinati, Ohio 45253

Received March 20, 2009; accepted July 7, 2009

"All models are wrong, some models are useful"

George E.P. Box, "Robustness in the strategy of scientific model building", 1979. This quote from George Box was the title of a book chapter on mathematical models—exactly the type of models which we will not address here, although a lot of the general reasoning can as well be translated to them, especially because they depend on the input and thus on the limitations of either *in vitro* or *in vivo* data, when applied for regulatory toxicology. This article was prompted by the 2008 SOT/EuroTox debate (March 2008 in Seattle and October 2008 in Rhodes), which challenged us in an Hegelian approach as thesis and antithesis to present with changing roles on both occasions on the statement "*In vitro* tests are useless for regulatory use." Here, we would like following Hegel a summary of thesis and antithesis, but also try to outline the synthesis.

Key Words: **critical review; animal testing; *in vivo*; cell culture; advances in toxicology.**

Regulatory toxicology is a child of its time. It started in the beginning of the 19th century with the increasing role of chemicals in everyday life. Small animal laboratory research flourished at the same time and has imprinted strongly on the make-up of safety testing approaches. Today, it accounts for about 10% of animal use (i.e., more than 1 million animals in Europe 2005, no comparable data available for other regions). The figure is roughly 25% for regulatory safety testing when vaccine safety control is included, but this is out of the remits of classical toxicology. With the development of life sciences, *in vitro* approaches have gained importance through-out and have contributed largely to the biotech revolution of recent years. Regulatory toxicology has embraced them in part, especially in the field of genotoxicity and more recently for hazard identification in topical toxicity. Furthermore, mechanistic models play an enormous role to support regulatory decisions. Still, the core of regulatory approaches has not changed markedly in the last 40–70 years. One of the reasons for the slow adaptation to advances in science and technology is certainly the fact that internationally harmonized guidelines once agreed on are difficult to change. There are other drivers (e.g., psychological and economical; Bottini and Hartung, 2009) which make it convenient not to question the status quo. The review literature on the limitations of basic toxicological tools is astonishingly scarce (Hartung, 2008a), with the notable exception of the cancer bioassay (Ames and Gold, 1990; Ennever and Lave, 2003; Ennever *et al*., 1987; Fung *et al*., 1995; Gaylor, 2005; Gold *et al*., 1998; Gottmann *et al*., 2001; Haseman *et al*., 1987; Huff, 1999; Lave *et al*., 1988), which has created enormous interests because of public interest in this health effect and the enormous costs of the animal study of up to $1 million per substance. Overall, results for the cancer bioassay suggest maximal 70% correlation between rodent species, about 50% positive findings for all substances tested implying about 80–90% false-positive findings because less than 5–10% of substances are supposed to be really carcinogenic (Fung *et al*., 1995).

In contrast, *in vitro* tests undergo the most extensive evaluation of any model in the life sciences; costs of the validation process easily exceed $500.000 and often tens of men years are spent. The process—although only started for the most promising candidate tests—sorts out about two thirds of the candidates as not (yet) valid. However, most of these die a silent death and failures are not necessarily published. More general reviews of limitations of *in vitro* approaches are also rare (Hartung, 2007a).

What are the limitations and advantages of both approaches in more general terms?

ADVANTAGES OF ANIMAL MODELS

A living being with hundreds of tissues and all physiological reactions and interactions is exposed. Currently, the extensive interactions among cells and tissues cannot be completely duplicated in a nonanimal model. The basic technology is simple and not too expensive. Decades of experience and

[1] EuroTox representative, Baltimore, MD 21205.

[2] To whom correspondence should be addressed. Fax: (410) 614-2871. E-mail: thartung@jhsph.edu.

[3] SOT representative.

international harmonization help interpretation and acceptance of data. There are animal models for the vast majority of toxicant-induced human diseases, indicating that the animal models contain the same molecular targets or pathways as humans, although the identity of those targets or pathways may not be known.

LIMITATIONS OF ANIMAL MODELS

The inevitable killing of animals raises increasing ethical concerns (Hartung, 2008b; Matthews, 2008; Shanks *et al*., 2009). Although the concordance between animal and human responses is generally moderate to good, there are species differences in responses that hinder our ability to use the animal data for predicting human responses. Not only for the cancer bioassay, but also for acute toxicity (Ekwall *et al*., 1998), eye irritation (York and Steiling, 1998), skin irritation (Basketter *et al*., 2004) or reproductive toxicity (Bremer *et al*., 2005), limited predictive capacities of the animal test have been documented. The models require relative large amounts of test substance. The use of inbred-strains does not reflect natural variances (Kacew and Festing, 1996) especially for uptake and biotransformation (Lin, 2007; Lovell, 1993) and only in few routine tests are both sexes or different life-stages used. Data are difficult to interpret because of the complexity of interactions and because the experiments concentrate on pathological outcomes and not the underlying physiological basis of the adverse effect. The exposure scenarios are unrealistic (Rietjens and Alink, 2006) (maximum tolerated doses, no coexposures). The study design is insufficient to permit the investigation of effects at the low end of the dose-response curve. The predictive value of the tests (sensitivity, specificity) and prevalences (i.e., proportion of toxic substances in the real world) allowing to deduce predictive capacity (how much can you rely on a positive or negative finding) are not known (Hoffmann and Hartung, 2006). The tests are typically done only once and often not published (Purchase, 2004), largely because of the expense and time involved in conducting the studies, so reproducibility is unknown. In 2007b, M. Levitt (U.S. Secretary of Health and Human Services) stated: "Currently, nine of ten experimental drugs fail in clinical studies because we cannot accurately predict how they will behave in people based on laboratory and animal studies." (Shanks *et al*., 2009).

ADVANTAGES OF *IN VITRO* MODELS

They originate directly of the current mainstream scientific approach in the life sciences and reflect our current mechanistic understanding of the different health effects. Basic methodologies are widely established. Set-ups are small using little test substance and allowing low costs, high-number of replicates and even miniaturization as well as automation. Novel technologies are quickly emerging which include image technologies as well as the diverse "omic" technologies. The reductionist approach allows access to the test material at all times and eases interpretation. Cell models for practically all tissues or laboratory animal species are available. There are few ethical problems, with the notable exceptions of human tissue donation and embryonic stem cells.

LIMITATIONS OF *IN VITRO* MODELS

The broad use of *in vitro* tests in academic and industrial research in recent years has somewhat obscured the difficulties involved. Good Cell Culture Practices (GCCP) have been proposed (Coecke *et al*., 2005; Hartung, 2007a; Hartung *et al*., 2002) but are rarely applied. Guidance for Good Laboratory Practices to extend to *in vitro* studies for regulatory work have been developed based on GCCP, but again its application in practice is not clear. There are some fundamental problems as to the artificial nonphysiological conditions cells are maintained in (not reflecting the body temperature of animals, the blood electrolyte concentrations of species, the extracellular matrix or the extent of cell contacts, which is maximally 15% of normal in monolayer cultures (which lay like neighboring fried eggs in a pan). Cell densities are less than 1% of the tissue situation. This impairs intracellular signaling. Most cell systems are representing only one cell type (no cell–cell interactions), often monoclonal in origin and further degenerated during maintenance. Culture conditions are not homeostatic (sudden exchange of media, continuous depletion of nutrients and accumulation of waste products) and oxygen supply is not sufficient (dissolved oxygen is typically consumed during the first hours allowing then only diffusion-limited supply, sometimes resulting in anaerobic culture conditions, i.e., glycolysis seen as lactate accumulation by phenol red turning yellow). All growth conditions are usually optimized toward rapid growth of cells to quickly allow the next experiment; however, cell growth and differentiation are just opposite programs (as we well know from cancer). Thus we are driving cells into dedifferentiation and select for mutations and subpopulations which grow faster and do not waste time for differentiated cell functions. The lack of biotransformation capabilities is probably the best-known limitation (Coecke *et al*., 2006), though the lack of defense mechanisms has probably stronger impact on the precision of toxicity estimations (Hartung, 2007a). The cancer origin of many cells commonly used adds to this problem: It has been shown that cancer cells have sometimes ten-thousands of mutations (Frank and Nowak, 2004; Ponten, 2001) including losses of parts or whole chromosomes. The authenticity of cell lines is a further largely neglected problem, with relevant percentages of cell lines even in banks being contaminated or mistaken (Buehring *et al*., 2004; Markovic and Markovic, 1998; MacLeod *et al*., 1999).

CONCLUSIONS

Taken together, all our tools have severe limitations. O.K., but that's what we have, many will say. This reminds somewhat of the joke of the man looking during the night for his key under a lantern; when asked whether he lost them here, he replies: "No, but here I have light."

For the moment, the one-eyed are kings among the blind. But is there a synthesis, is there drive for change? In fact various needs prompt longing for a revision of current practices:

1. Our strong and ever-increasing understanding of the mechanistic underpinnings of toxicity are not adequately reflected in current *in vivo* testing (Hartung and Leist, 2008; Leist *et al*., 2008). There is a need to more routinely incorporate mechanistic investigations into the evaluation of toxicity, either by incorporating mechanistically linked measurements into animal studies, a greater reliance on *in vitro* assays that are mechanistically based, or a combination of the two.

2. New technologies pose new problems in hazard and risk assessment. New products can often not be adequately evaluated with current approaches, for example, half of the new drugs are nowadays human proteins and antibodies, for which only analogues can be tested in animal models. Nanoparticles (Service, 2008), genetically modified food and feed (Kuiper *et al*., 2001), and other new technologies will pose problems in test design and interpretation.

3. There is a need for risk assessment approaches that provide better predictions of human risk. Current approaches are often too conservative because of the large uncertainties that remain after animal testing. Although it was perhaps an easy solution "to close the books" after a set of supposedly meaningful tests had been done to satisfy liability issues, increasingly, the costs of this approach become clear: we lose valuable substances, which never make it to products (e.g., therapies not becoming available). It is frightening to think that aspirin would most probably today not make it to human trials. Now we are using the same tests for existing chemicals: the approach remains precautionary; therefore, there it is likely that more testing will be done in the future to determine whether a certain finding is reproducible or relevant for humans, as is often already the case for pharmaceuticals and pesticides.

4. There is a need to have assay systems that have higher throughput. Toxicology has been shaped by the discovery of drugs in pharmaceutical industry. For good reason a precautionary approach is taken here, because later withdrawals of drugs from the market can be disastrous. For pesticides a similar approach has later been adopted, with the consequence that the toxicology package of a new pesticide amounts nowadays to about $20 million of costs and only about eight new substances/formulations enter the market per year. Interestingly, this is about the same number as for new therapeutic agents. We have adopted this approach to new chemicals, with most testing being dictated by production volume and potential for either high or widespread exposure. It is not feasible to exhaustively test all 200–300 new chemicals introduced per year in Europe.

5. Validation programs have been developed that are rigorous (Abbot, 2005). Formal validation of alternative approaches as carried out by European Centre for the Validation of Alternative Methods, Interagency Coordinating Committee on the Validation of Alternative Methods, and Japanese Center for the Validation of Alternative Methods has created confidence (and expectations) as to the possible replacement of traditional by new approaches. However, there is concern that *in vitro* tests are held to a standard of validation that far exceeds what most *in vivo* tests or *in silico* approaches have undergone (Hartung, 2007b; Hoffmann and Hartung, 2005; Hoffmann *et al*., 2008b).

6. The general public is less and less accepting the use of experimental animals for purposes which do not absolutely require them. This contrasts remarkably with the increasing expectations for increased safety of products and therapies. Politicians (especially in the European Union) have taken up these expectations as voiced by various non-governmental organizations. This has led to both animal welfare legislations (cosmetics) (Hartung, 2008c) and large chemical safety programs (REACH, pesticides) in Europe with impact on the global markets and force change toward nonanimal methods and higher throughput.

SYNTHESIS

The respective advantages and limitations argue much more for a combined approach than either approach alone. Neither represents a gold standard, nor do traditional assays. Acknowledging this is already the first step for a more adequate use of what science has made available so far. In an ideal world this would then call for a systematic validation of current approaches, with the question what our point of reference in the absence of large human datasets could be. A more realistic proposal is a scientifically rigorous retrospective assessment as proposed under the slogan of an evidence-based toxicology (Hoffmann and Hartung, 2006; www.ebtox.org). Structured reviews of existing evidence are a key feature of evidence-based medicine and translating this to toxicology might give us means to judge the quality of our tools without major new trials while remaining as objective as possible.

Integrated testing strategies represent then the first step forward (Hartung, 2009a). Tests can ideally be combined if they are complementary to each other. The tendency to model *in vivo* tests with *in vitro* is exactly the worst we can do to combine them reasonably. The simplest combination is always the one of a sensitive and a specific test. Such combinations when combined with prevalence information can be modeled with regard to the overall performance to see which test to

perform first and whether to control the positive or the negative findings (Hoffmann and Hartung, 2006; Hoffmann *et al*., 2008a). The first step, however, is to determine the prevalences for the various health effects in different chemical universes. But we have to go one step further: we need to understand also which endpoints led to the relevant regulatory decisions. It is fruitless to model all aspects of a two-generation reproductive toxicity study *in vitro* (and will become ridiculous if mating behavior of the rat is addressed). But if, for example, we see that in 80% of cases effects on the testis lead to positive calls, it becomes much more realistic to set up an *in vitro* test battery (Bremer *et al*., 2007).

Testing strategies represent the key opportunity to balance the shortcomings of all approaches. They have a key disadvantage for their users, that is, they require decision points in-between what to do next. The set of tests and thus costs and timing can not be determined upfront. But earlier no-go criteria within such a testing strategy might actually also reduce time and costs, given that, for example the organization, execution and analysis of a two-generation study lasts ca. 5 years.

Given the key characteristics of *in vitro* tests (costs, through-put, animal use) they should play a role in an early phase of a testing strategy, not as the mechanistic tests at the end to further explain *in vivo* findings. However, we must not repeat the mistakes from *in vivo* approaches (maximum tolerated doses whose equivalent are unrealistic high doses *in vitro*, multiple testing adding up false-positives). The result of such translation to *in vitro* we have seen in case of genotoxicity: as nicely shown by Kirkland *et al*., the positive predictive value (i.e., how much we can rely on a positive finding) of the current combination of three genotoxicity tests is 3–20% only (Kirkland *et al*., 2005). This is perhaps prompting more animal tests than saving. Beyond this ethical aspect, it is simply binding test capacities which should be used to make products safer and not to compensate for weaknesses of test approaches.

The last and most probably most important step will then be to fully embrace and integrate the advances of toxicological and other life sciences (Hartung, 2009b). This includes technologies and knowledge, that is, the mechanisms and pathways leading to health effects. The fact that we refer again to health effects and ignore environmental effects has to be attributed to the personal background of the authors not to a difference in opportunities or needs for this part of toxicology. The term of ''systems toxicology'' has been coined for the combination of technology and knowledge by creating information-rich situations mined by bioinformatics (Hartung and Leist, 2008; Leist *et al*., 2008).

In conclusion, neither approach is useless but only fully useful in the right regulatory testing framework. Needs and opportunities for change are increasingly evident. The challenge posed originally by simple *in vitro* tests to already then somewhat outdated *in vivo* approaches has become over the last two decades only more pronounced. Some changes might come as revolutionary rather than as evolutionary changes (Hartung, 2008a). To make them fully useful it will take more than to develop and validate them as individual tests, but to review the current tools in an objective manner and change the regulatory testing framework toward integrated testing strategies.

ACKNOWLEDGMENTS

T.H. wrote this article and made the underlying presentations at SOT and EuroTox as employee of the European Commission. G.D. is employee of The Procter & Gamble Company. The support is gratefully appreciated.

REFERENCES

Abbot, A. (2005). More than a cosmetic change. *Nature* **438,** 144–146.

Ames, B. N., and Gold, L. S. (1990). Chemical carcinogenesis: Too many rodent carcinogens. *Proc. Natl. Acad. Sci. U. S. A.* **87,** 7772–7776.

Basketter, D. A., York, M., McFadden, J. P., and Robinson, M. K. (2004). Determination of skin irritation potential in the human 4-h patch test. *Contact Dermatitis* **51,** 1–4.

Bottini, A. A., and Hartung, T. (2009). Food for thought... on economic mechanisms of animal testing. *Altern. Anim. Exp.* **26,** 3–16.

Bremer, S., Pellizzer, C., Hoffmann, S., Seidle, T., and Hartung, T. (2007). The development of new concepts for assessing reproductive toxicity applicable to large scale toxicological programs. *Curr. Pharm. Des.* **13,** 3047–3058.

Bremer, S., Balduzzi, D., Cortvrindt, R., Daston, G., Eletti, B., Galli, A., Huhtaniemi, I., Laws, S., Lazzari, G., Liminga, U., *et al*. (2005). The effects of chemicals on mammalian fertility. The report and recommendations of ECVAM Workshop 53—The first strategic workshop of the EU ReProTect Project. *Altern. Lab. Anim.* **33,** 391–416.

Buehring, G. C., Eby, E. A., and Eby, M. J. (2004). Cell line cross-contamination: How aware are mammalian cell culturists of the problem and how to monitor it? *In Vitro Cell Dev. Biol. Anim* **40,** 211–215.

Coecke, S., Ahr, H., Blaauboer, B. J., Bremer, S., Casati, S., Castell, J., Combes, R., Corvi, R., Crespi, C. L., Cunningham, M. L., *et al*. (2006). Metabolism: A bottleneck in in vitro toxicological test development. *Altern. Lab. Anim.* **34,** 49–84.

Coecke, S., Balls, M., Bowe, G., Davis, J., Gstraunthaler, G., Hartung, T., Hay, R., Merten, O.-W., Price, A., Schechtman, L., Stacey, G., and Stokes, W. (2005). Guidance on good cell culture practice. *Altern. Lab. Anim.* **33,** 261–287.

Ekwall, B., Clemedson, C., Crafoord, B., Ekwall, B., Hallander, S., Walum, E., and Bondesson, I. (1998). MEIC evaluation of acute systematic toxicity Part V. Rodent and human toxicity data for the 50 reference chemicals. *Altern. Lab. Anim.* **26,** 571–616.

Ennever, F. K., and Lave, L. B. (2003). Implications of the lack of accuracy of the lifetime rodent bioassay for predicting from carcinogenicity. *Regul. Toxicol. Pharmacol.* **38,** 52–57.

Ennever, F. K., Noonan, T. J., and Rosenkranz, H. S. (1987). The predictivity of animal bioassays and short-term genotoxicity tests for carcinogenicity and non-carcinogenicity to humans. *Mutagenesis* **2,** 73–78.

Festing, M. F. (1993). Genetic variation in outbred rats and mice and its implications for toxicological screening. *J. Exp. Anim. Sci.* **35,** 210–220.

Frank, S. A., and Nowak, M. A. (2004). Problems of somatic mutation and cancer. *Bioessays* **26,** 291–299.

Fung, V. A., Barrett, J. C., and Huff, J. (1995). The carcinogenesis bioassay in perspective: Application in identifying human cancer hazards. *Environ. Health Perspect.* **103,** 680–683.

Gaylor, D. W. (2005). Are tumor incidence rates from chronic bioassays telling us what we need to know about carcinogens? *Regul. Toxicol. Pharmacol.* **41,** 128–133.

Gold, L. S., Slone, T. H., and Ames, B. N. (1998). What do animal cancer tests tell us about human cancer risk?: Overview of analyses of the carcinogenic potency database. *Drug Metab. Rev.* **30,** 359–404.

Gottmann, E., Kramer, S., Pfahringer, B., and Helma, C. (2001). Data quality in predictive toxicology: Reproducibility of rodent carcinogenicity experiments. *Environ. Health Perspect.* **109,** 509–514.

Hartung, T. (2009a). Toxicology for the twenty-first century. *Nature* **460,** 208–212.

Hartung, T. (2009b). A toxicology for the 21st century–Mapping the road ahead. *Toxicol. Sci.,* **109,** 18–23.

Hartung, T. (2008a). Towards a new toxicology—Evolution or revolution? *Altern. Lab. Anim.* **36,** 635–639.

Hartung, T. (2008b). Food for thought . . . on animal tests. *Altern. Anim. Exp.* **25,** 3–9.

Hartung, T. (2008c). Food for thought . . . on alternative methods for cosmetics safety testing. *Altern. Anim. Exp.* **25,** 147–162.

Hartung, T., and Leist, M. (2008). Food for thought . . . on the evolution of toxicology and phasing out of animal testing. *Altern. Anim. Exp.* **25,** 91–96.

Hartung, T. (2007a). Food for thought . . . on cell culture. *Altern. Anim. Exp.* **24,** 143–147.

Hartung, T. (2007b). Food for thought . . . on validation. *Altern. Anim. Exp.* **24,** 67–72.

Hartung, T., Balls, M., Bardouille, C., Blanck, O., Coecke, S., Gstraunthaler, G., and Lewis, D. (2002). Report of ECVAM task force on good cell culture practice (GCCP). *Altern. Lab. Anim.* **30,** 407–414.

Haseman, J. K., Huff, J. E., Zeiger, E., and McConnell, E. E. (1987). Comparative results of 327 chemical carcinogenicity studies. *Environ. Health Perspect.* **74,** 229–235.

Hoffmann, S., Edler, L., Gardner, I., Gribaldo, L., Hartung, T., Klein, C., Liebsch, M., Sauerland, S., Schechtman, L., Stammati, A., and Nikolaidis, E. (2008b). Points of reference in validation—The report and recommendations of ECVAM Workshop. *Altern. Lab. Anim.* **36,** 343–352.

Hoffmann, S., Saliner, A. G., Patlewicz, G., Eskes, C., Zuang, V., and Worth, A. P. (2008a). A feasibility study developing an integrated testing strategy assessing skin irritation potential of chemicals. *Toxicol. Lett.* **180,** 9–20.

Hoffmann, S., and Hartung, T. (2006). Towards an evidence-based toxicology. *Hum. Exp. Toxicol.* **25,** 497–513.

Hoffmann, S., and Hartung, T. (2005). Diagnosis: Toxic!—Trying to apply approaches of clinical diagnostics and prevalence in toxicology considerations. *Toxicol. Sci.* **85,** 422–428.

Huff, J. E. (1999). Long-term chemical carcinogenesis bioassays predict human cancer hazards. Issues, controversies, and uncertainties. *Ann. N. Y. Acad. Sci.* **895,** 56–79.

Kacew, S., and Festing, M. F. (1996). Role of rat strain in the differential sensitivity to pharmaceutical agents and naturally occurring substances. *J. Toxicol. Environ. Health* **47,** 1–30.

Kirkland, D., Aardema, M., Henderson, L., and Müller, L. (2005). Evaluation of the ability of a battery of three in vitro genotoxicity tests to discriminate rodent carcinogens and non-carcinogens. I. Sensitivity, specificity and relative predictivity. *Mutat. Res.* **584,** 1–256.

Kuiper, H. A., Kleter, G. A., Noteborn, H. P. J. M., and Kok, E. J. (2001). Assessment of the food safety issues related to genetically modified foods. *Plant J.* **27,** 503–528.

Lave, L. B., Ennever, F. K., Rosenkranz, H., and Omenn, G. S. (1988). Information value of the rodent bioassay. *Nature* **336,** 631–633.

Leist, M., Hartung, T., and Nicotera, P. (2008). The dawning of a new age of toxicology. *Altern. Anim. Exp.* **25,** 103–114.

Lin, J. H. (2007). Pharmacokinetic and pharmacodynamic variability: A daunting challenge in drug therapy. *Curr. Drug Metab.* **8,** 109–136.

Lovell, D. P. (1993). Impact of pharmacogenetics on toxicological studies. Statistical implications. *J. Exp. Anim. Sci.* **35,** 259–281.

MacLeod, R. A. F., Dirks, W. G., Matsuo, Y., Kaufmann, M., Milch, H., and Drexler, H. G. (1999). Widespread intraspecies cross-contamination of human tumor cell lines arising at source. *Int. J. Cancer* **83,** 555–563.

Markovic, O., and Markovic, N. (1998). Cell cross-contamination in cell cultures: The silent and neglected danger. *In Vitro Cell Dev. Biol. Anim.* **34,** 1–8.

Matthews, R. A. J. (2008). Medical progress depends on animal models—Doesn't it? *J. R. Soc. Med.* **101,** 95–98.

Ponten, J. (2001). Cell biology of precancer. *Eur. J. Cancer* **37**(Suppl. 8), 97–113.

Purchase, I. F. H. (2004). Fraud, errors and gamesmanship in experimental toxicology. *Toxicology* **202,** 1–20.

Rietjens, I. M. C. M., and Alink, G. M. (2006). Future of toxicology low-dose toxicology and risk–benefit analysis. *Chem. Res. Toxicol.* **19,** 977–981.

Service, R. F. (2008). Nanotechnology. Can high-speed tests sort out which nanomaterials are safe? *Science* **321,** 1036–1037.

Shanks, N., Greek, R., and Greek, J. (2009). Are animal models predictive for humans? *Philos. Ethics Humanities Med.* **4,** 2 Available from: http://www.peh-med.com/content/4/1/2

York, M., and Steiling, W. (1998). A critical review of the assessment of eye irritation potential using the Draize rabbit eye test. *J. Appl. Toxicol.* **18,** 233–240.