

TestSmart and Toxic Ignorance

Sidney Green¹ and Alan M. Goldberg²

¹Department of Pharmacology, Howard University College of Medicine, Washington, DC, USA;

²Johns Hopkins University Center for Alternatives to Animal Testing, Baltimore, MD, USA

Summary — There are a number of national and international efforts designed to screen chemicals for toxicity. Although the emphasis in terms of the specific chemicals is different, e.g. endocrine disruptors, children's health, High Production Volume (HPV), the European Registration, Evaluation and Authorisation of Chemicals (REACH) programme, the purpose is the same. Each is intended to evaluate the potential toxicity of chemicals to humans and, in some cases, to the environment. How best can these tasks be accomplished? The first need is to provide a realistic prioritisation of which chemicals need to be evaluated. Once identified, a defined decision-tree approach with an emphasis on short-term *in vitro* assays and new genomic technologies offers the greatest promise. The more practical matter of screening the chemicals would be by using a tiered decision-tree approach. Common features of the approach would be the use of three tiers. The first tier would be a screening/prioritisation tier, the second would provide an initial characterisation of toxicity, and the third would discern mode of action/biological activity. The intent of this approach is to provide a concept that will allow decisions to be made as to which chemicals need to be tested, provide some idea as to their toxicity and finally mode of action, and at the same time, taking into account the Three Rs, *reduction*, *refinement* and *replacement*. Thus, specific batteries of tests are not discussed, as these would need to be tailored to the specific chemicals of concern, e.g. endocrine disruptors, HPV. Neither are regulatory requirements factored into the concept, but the data that would be gathered should consider the possibility of eventual submission of the data obtained by *in vitro* and other non-traditional approaches by regulatory authorities.

Key words: *alternatives, endocrine disruptors, High Production Volume (HPV), REACH, TestSmart, tier testing, toxicity screening.*

Address for correspondence: A.M. Goldberg, Johns Hopkins University Center for Alternatives to Animal Testing, 111 Market Place, Suite 840, Baltimore, MD 21202, USA.
E-mail: Goldberg@jhsph.edu

Introduction

There are a number of national and international efforts designed to screen chemicals for toxicity. These efforts are directed toward providing basic toxicity information on commercial chemicals in use worldwide. An important objective of these initiatives is to make such data available to scientists, policy makers, industry and the public (see www.epa.gov/chemrtk/index.htm). Although the emphasis in terms of the specific chemicals is different, e.g. endocrine disruptors, children's health, High Production Volume (HPV), the European Registration, Evaluation and Authorisation of Chemicals (REACH) programme, the purpose is the same. Each is intended to evaluate the potential toxicity of chemicals to humans, and in some cases, to the environment. Issues of whether to use animals or *in vitro* tests, what to test, how to prioritise those chemicals to be tested and determining types and amounts of information sufficient to allow use all complicate how preliminary safety/hazard determinations will be accomplished for these chemicals.

This paper is intended to offer a concept or a blueprint for discussion and development of a

master plan. It is not intended to provide a protocol for the testing of chemicals. From the outset, we have made a number of assumptions in our approach. We accept that the US Environmental Protection Agency (EPA) and the European Commission are correct in their conclusion that it is necessary to provide data on these compounds to the public, but there are few currently validated (within the guidelines of the Interagency Coordinating Committee for the Validation of Alternative Methods [ICCVAM] or the European Centre for the Validation of Alternative Methods [ECVAM]) *in vitro* or *in vivo* methods that could be employed in testing these chemicals. However, there are many methods considered as "valid". For example, the Screening Information Data Set (SIDS) guidance of the OECD. Valid, in this sense, means a well-established and accepted test method whose outcome is interpretable. Lastly, there is a basic set of data available on most of these compounds, e.g. structure–activity relationship (SAR). We recognise that the concept outlined may just be one of many ways of addressing these issues. We offer this proposal in order to stimulate discussion and highlight questions we think need to be addressed.

Table 1: Basics of the tier approach**Guiding principles**

Most humane science

Cost-effective, rapid screening

The process

Literature search

Structure-activity relationships

ADME and pilot studies

The approach

Tier one: prioritisation of chemicals, with negatives being eliminated

Tier two: prioritisation of tier one chemicals and initial characterisation of toxicity/safety (all *in vitro*)Tier three: testing of individual chemicals, determining mode of action and biological activity by use of *in vitro* systems and if necessary using non-invasive techniques in animals*ADME = absorption, distribution, metabolism and excretion.***Methods/Discussion**

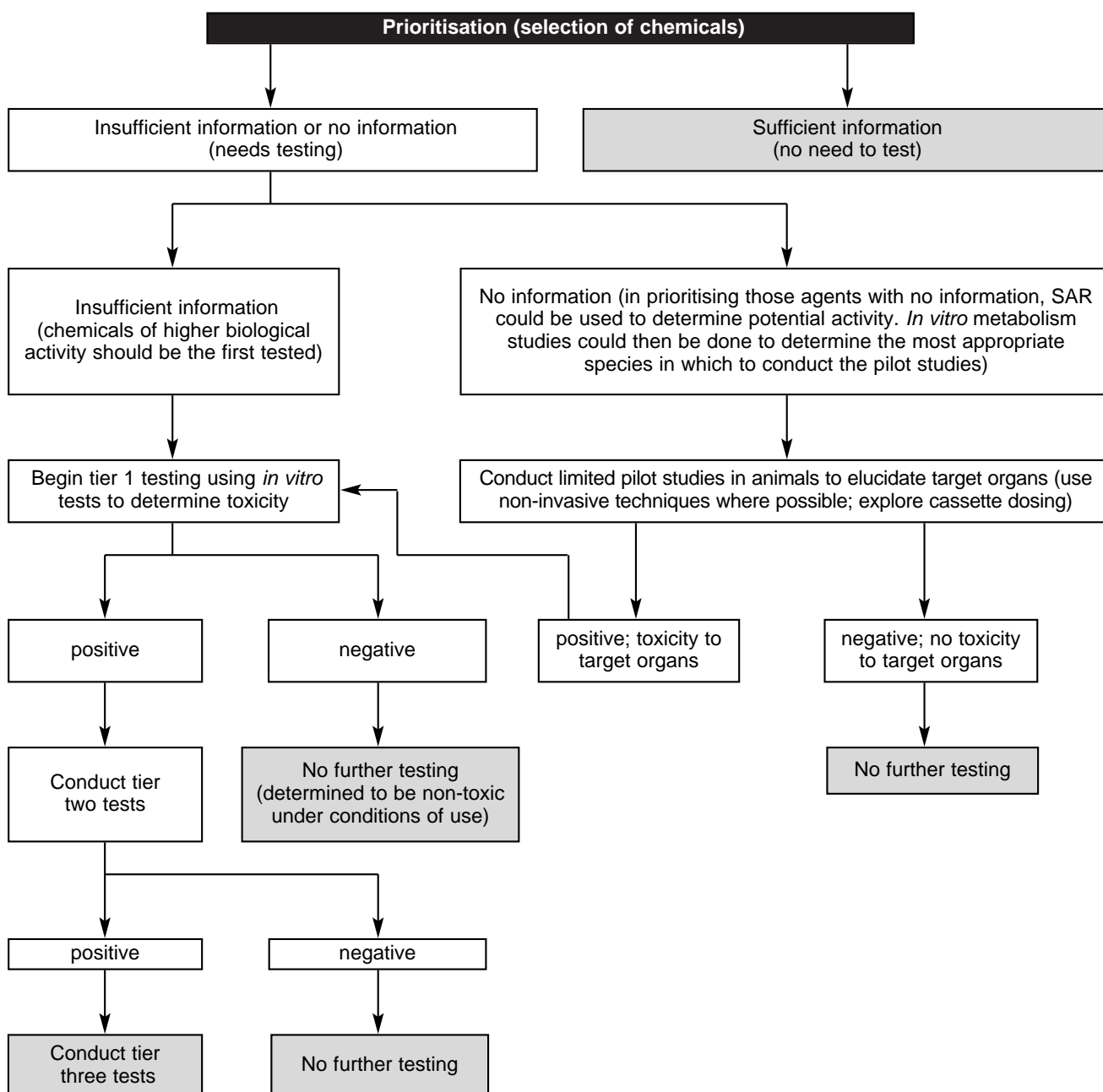
The concept involves the use of a tiered-testing scheme (Table 1). This scheme employs *in vitro* methods primarily and *in vivo* methods only in situations where their use is unavoidable. There are many approaches to prioritising groups of chemicals for testing and evaluation. Some are based on production volume, exposure or other rationales. With respect to production volume and exposure, the HPV chemicals and pesticides used in food crops can be cited as examples. The tiered approach suggested here prioritises within a group and, in some cases, will provide approaches to mode of action. A first requirement is a review of the available literature, followed by determination of SAR. Chemicals with sufficient information to determine hazard or risk would not be tested. Those with insufficient or no information (no information in this case would be not enough to even initiate *in vitro* testing) would enter a first tier of testing. In the latter situation (no information), and contrary to current thinking, limited *in vivo* pilot studies might be conducted to obtain information about possible target organs. Such studies should use, where possible, non-invasive methods and would always incorporate humane approaches and humane endpoints. This information would allow *in vitro* studies of the proper endpoints and cell types to be conducted. Chemicals negative at tier one would require no further testing at this time, while those positive would be tested at tier two. Likewise, those negative at tier two would require no further testing, while those positive would be tested at tier three (Figure 1).

Specifically, the concept embodies the following. An assessment of potential biological activity using SAR could be used to prioritise chemicals with insufficient information. Those chemicals considered to possess a higher degree of biological activity would be the first tested. Once identified, these chemicals would be tested in tier one. As far as those chemicals with no information are concerned, SAR could also be used to establish potential biological activity, along with *in vitro* metabolism, to determine a target species for limited pilot studies. The metabolism studies could elucidate (in terms of activity/toxicity) the most appropriate species, along with the organ systems to be investigated. This information could then be used to select appropriate *in vitro* systems for tier one testing. It is highly recommended that non-invasive techniques and, perhaps, cassette dosing be used, to the extent possible, in the *in vivo* studies. Those chemicals demonstrating no activity/toxicity would be placed in the lowest priority for testing or not tested. Criteria would need to be established to define "no activity/toxicity". It may be advisable to consider additional testing of certain negative chemicals, depending on a level of exposure that would be considered significant or high. It would need to be determined what constitutes significant or high exposure. This would ensure that high exposure chemicals are tested at least in tiers one and two.

Tier one

The purpose of this tier is to screen in a qualitative manner for *potential* toxicity to humans and thereby prioritise within a group. Criteria for tests in this tier should be: a) allowance of a certain false positive rate, but very low or no false negatives, b) ability to detect a broad range of chemical categories of agents, c) efficiency in terms of time between testing and obtaining results, and d) cost-effective, given that more chemicals would be expected to be tested at this level than at other levels or tiers.

Tier one test systems could be composed of *in vitro* liver, vascular, cardiac, respiratory, reproductive, autonomic nervous system, kidney and mitochondrial cells. Hepatic metabolism has been investigated in isolated hepatocytes (1). Liver homogenates, subcellular fractions and liver slices have been used for hepatotoxicity (2). Although the vascular system (blood vessels, blood vessel walls) is mentioned in the context of tier one systems, to our knowledge very little information is available on *in vitro* approaches that have been used in screening chemicals for effects on the vascular system. It is mentioned to emphasise a possible research need (3). Tissue slices of cardiac tissue have been used in evaluating chemicals for cardiotoxicity (4), as well as single cell suspensions of embryonic or neonatal heart cells (5, 6). Potential effects on the respira-

Figure 1: Tier scheme for TestSmart and toxic ignorance

tory system can be determined by conducting *in vitro* studies of human cell lines, such as CCL-30 (7) and A-549 (8), which are cells of the nasal septum and lower alveolar region, respectively. As a first tier method for reproductive toxicity, sperm motility (9) could possibly provide preliminary information, albeit for the male only. Methods that could be considered relevant to the female and useful for tier one are fully lacking. For examining potential toxicity to the kidney, isolated cells, kidney fragments, precision-cut-slices and cell lines have been used. A review of the application of kidney cells to toxicity testing is given by Pfaller & Gstraunthaler (10).

Inclusion of cells of mitochondria is intended to emphasise a research need. To our knowledge, there are virtually no established test methods that are routinely used to screen chemicals for toxicity to mitochondria, although the MTT assay may provide indirect information. While these systems would provide primarily information relative to potential effects on organ systems, it may also be possible to determine chromosomal effects, which could point to follow-up genetic toxicology studies. Otherwise, stand-alone clastogenicity studies would need to be done. Gene mutational studies, e.g. the *Salmonella* microsome assay, would also need to be a part of

tier one. For the assessment of potential developmental toxicity, whole embryo or rodent limb bud cultures may need to be a part of this tier. With the exception of tests for genetic toxicity, there may not be any other validated or valid tests for this tier.

What configuration of results should lead to tier two testing? Should the decision be based on a positive result in one organ system, or should some combination of results be the requirement? Scientific judgment will need to be exercised, for the decision should not be reduced to an arbitrary number of positive results. Is there a rationale that could explain what is seen? One needs to determine that the results make good scientific sense. Only then will the testing in tier two yield meaningful information. Chemicals that are negative are eliminated from further testing when used as expected or accorded the lowest priority for further testing.

Tier two

The purpose of this level of testing is to prioritise those chemicals from tier one on the basis of biological activity and to provide an initial characterisation of the toxicity of a chemical. The tests to be selected and performed in tier two depend, in part, on results from SAR and information from the literature, as well as on results in tier one; for example, the target organs and toxicity endpoints affected. Thus, the number and kinds of tests for any one chemical may vary. Ideally, this level of testing should exclusively use *in vitro* methods. Test methods similar to those in tier one may be used in this tier, but the endpoints would be decidedly different.

There are certain criteria that should apply to tests, particularly *in vitro* tests, used in this tier. The first is that tests at this level should possess an identifiable mechanism that is involved in the animal or human biological response. They should have good sensitivity and specificity and be relatively easy to perform. Such tests would need to be identified and the extent of validation determined. This may be the best level to use "omic" information. Omic studies, such as genomics, proteomics and metabonomics (11–13) using microarrays, tissues and body fluids, respectively, associated with the target organs identified in animals, could involve exposure to a toxicant in parallel with those developed or obtained from similar human tissue to better understand the complementarity of the animal model to humans. In this regard, patterns of up-regulation and down-regulation of specific genes, patterns of metabolites, and changes in protein constitution, could be examined and a determination made as to the correlation between the animal and human data. This could provide greater confidence with respect to extrapolation that would be a part of the third tier.

Tier three

The purpose of this tier is to characterise the mode of action and biological activity of individual chemicals positive at tier two. The criterion for tests at this tier is that they have an explicit relevance to humans in terms of pathway of toxicity. In all likelihood, tests at this level will use *in vitro* systems, as well as animals. As such, use of non-invasive approaches in animals should be sought.

This concept, as presented, attempts to combine the best science and the most recent techniques with the most humane science possible. It is intended to be a discovery or decision-making approach and not one based necessarily on needs of the regulatory community. However, the assembly of methods or test systems for each tier should be done with full knowledge of regulatory requirements and, as we learn which assays provide reliable and relevant information, it may be possible to use these data in regulatory submissions in the future. In all cases, good science and humane science must be coequal partners.

References

1. Cross, D.M. & Bayliss, M.K. (2000). A commentary on the use of hepatocytes in drug metabolism studies during drug discovery and development. *Drug Metabolism Reviews* **32**, 219–240.
2. Waring, J.F. & Ulrich, R.G. (2000). The impact of genomics-based technologies on drug safety evaluation. *Annual Review of Pharmacology and Toxicology* **40**, 335–352.
3. Green, S. & Goldberg, A.M. (2001). TestSmart-High Production Volume Chemicals: an approach to implementing alternatives into regulatory toxicology. *Toxicological Sciences* **63**, 6–14.
4. Gandolphi, A.L., Brendel, K., Fisher, R.L. & Michaud, J.P. (1995). Use of tissue slices in chemical mixture toxicology and interspecies investigations. *Toxicology* **105**, 285–290.
5. Chan, E.M., Thomas, M.J., Bandy, B. & Tibbits, G.F. (1996). Effects of doxorubicin, 4-epirubicin, and antioxidant enzymes on the contractility of isolated cardiomyocytes. *Canadian Journal of Physiology and Pharmacology* **74**, 904–910.
6. Horwitz, L.D., Wallner, J.S., Decker, D.E. & Buxser, S.E. (1996). Efficacy of lipid soluble, membrane-protective agents against hydrogen peroxide cytotoxicity in cardiac myocytes. *Free Radical Biology and Medicine* **21**, 743–753.
7. Poliquin, J.F. & Crepeau, J. (1985). Immune defense mechanisms of the nasal mucosa. *Journal of Otolaryngology* **14**, 80–84.
8. Hukkanen, J., Lassila, A., Paivarinta, K., Valanne, S., Sarpo, S., Jakkola, J.K., Pelkonen, O. & Raunio, H. (2000). Induction and regulation of xenobiotic-metabolizing cytochrome P450s in the human A549 lung adenocarcinoma cell line. *American Journal of Respiratory Cell and Molecular Biology* **22**, 360–366.
9. Chapin, R.E., Sloane, R.A. & Haseman, J.K. (1998). Reproductive endpoints in general toxicity studies: are they predictive? *Reproductive Toxicology* **12**,

- 489–494.
10. Pfaller, W. & Gstraunthaler (1998). Nephrotoxicity testing *in vitro* — what we know and what we need to know. *Environmental Health Perspectives* **106**, (Suppl. 2), 559–569.
 11. Hamadeh, H.K., Bushel, P.R., Jayadev, S., DiSorbo, O., Bennett, L., Li, L., Tennant, R., Stoll, R., Barrett, J.C., Paules, R.S., Blanchard, K. & Afshari, C.A. (2002). Prediction of compound signature using high density gene expression profiling. *Toxicological Sciences* **67**, 232–240.
 12. Man, W.J., White, J.R., Bryant, D., Bugelski, P., Camilleri, P., Cutler, P., Heald, G., Lork, P.G. & Kramer, K. (2002). Protein expression analysis of drug-mediated hepatotoxicity in the Sprague-Dawley rat. *Proteomics* **2**, 1577–1585.
 13. Nicholson, J.K., Connelly, J., Lindon, J.C. & Holmes, E. (2002). Metabonomics: a platform for studying drug toxicity and gene function. *Nature Reviews Drug Discovery* **1**, 153–161.