

Animal Carcinogenicity Studies: Implications for the REACH System

Andrew Knight,¹ Jarrod Bailey² and Jonathan Balcombe³

¹Animal Consultants International, London, UK; ²School of Population and Health Sciences, Faculty of Medical Sciences, University of Newcastle upon Tyne, UK; ³Physicians Committee for Responsible Medicine, Washington DC, USA

Summary — The 2001 European Commission proposal for the Registration, Evaluation and Authorisation of Chemicals (REACH) aims to improve public and environmental health by assessing the toxicity of, and restricting exposure to, potentially toxic chemicals. The greatest benefits are expected to accrue from decreased cancer incidences. Hence the accurate identification of chemical carcinogens must be a top priority for the REACH system. Due to a paucity of human clinical data, the identification of potential human carcinogens has conventionally relied on animal tests. However, our survey of the US Environmental Protection Agency's (EPA's) toxic chemicals database revealed that, for a majority of the chemicals of greatest public health concern (93/160, i.e. 58.1%), the EPA found animal carcinogenicity data to be inadequate to support classifications of probable human carcinogen or non-carcinogen. A wide variety of species were used, with rodents predominating; a wide variety of routes of administration were used; and a particularly wide variety of organ systems were affected. These factors raise serious biological obstacles that render accurate extrapolation to humans profoundly difficult. Furthermore, significantly different International Agency for Research on Cancer assessments of identical chemicals, indicate that the true human predictivity of animal carcinogenicity data is even poorer than is indicated by the EPA figures alone. Consequently, we propose the replacement of animal carcinogenicity bioassays with a tiered combination of non-animal assays, which can be expected to yield a weight-of-evidence characterisation of carcinogenic risk with superior human predictivity. Additional advantages include substantial savings of financial, human and animal resources, and potentially greater insights into mechanisms of carcinogenicity.

Key words: *alternative, animal experiment, animal test, bioassay, cancer prevention, carcinogenicity, chemical classification, chemical safety, computer simulation, REACH, risk assessment.*

Address for correspondence: A. Knight, Animal Consultants International, 91 Vanbrugh Court, Wincott Street, London SE11 4NR, UK.
E-mail: info@animalconsultants.org

Introduction

The 2001 European Commission (EC) proposal for the Registration, Evaluation and Authorisation of Chemicals (REACH) produced or imported in quantities in excess of one tonne has two main objectives: to improve the competitiveness of the European Union (EU) chemical industry internationally; and to provide enhanced protection for human health and the environment from the toxic effects of chemicals.

Some tangible estimates of the potential benefits to human health are provided in a study conducted for the EC's Environment Directorate General (DG) by Postle *et al.* (1). The study examined the likely impacts of the implementation of the REACH system on the occupational health of chemical industry workers and downstream users of chemicals. Specifically, it estimated the likely impacts of the REACH system on five body systems and associated diseases: the skin (eczema, allergic contact dermatitis, irritant contact dermatitis); the respiratory system (asthma, allergic rhinitis, and other

respiratory illnesses); the eyes (conjunctivitis); the central nervous system (CNS disorders); and 16 types of cancer death.

The precise causes of cancer are often difficult to discern, owing to the contributions of lifestyle, environmental, occupational and genetic factors. Nevertheless, some 32,500 cancer deaths, representing around 3.5% of the annual EU total, are considered to arise mainly from occupational exposure to carcinogens. Around 6500 (20%) of these 32,500 deaths were estimated to arise from exposure to unknown chemical carcinogens (with the remainder associated with known or suspected carcinogens, and thus covered by existing legislation). Postle *et al.* (1) estimated that the REACH approach might prevent between one-third and two-thirds of these deaths within the EU annually, i.e. between 2167 cancer deaths (0.23% of the EU total) and 4333 cancer deaths (0.47%).

Compared with most other diseases, the human and economic costs of cancer are very high. Based on cost-benefit analyses carried out by DG Environment (1), the cost of a cancer death was

estimated as being between €1.39 million (lower estimate) and €2.14 million (best estimate). These estimates were initially sourced from the UK Government, and were based on the amounts surveyed individuals stated they would be willing to invest to reduce their personal risks of transport-related fatality. Additionally, the best estimate included some element of medical costs, and lost production and human costs. These estimates were then adjusted to reflect the age of those at risk of cancer related-death (typically, elderly), and the fact that there is usually a period of ill health prior to a cancer death, which the individuals surveyed also stated they would be willing to pay to avoid (a “cancer premium”).

Not included in these estimates were the medical, production and human costs associated with non-fatal cancer cases. Other costs which were excluded or potentially under-estimated, included:

- diagnostic and treatment-related medical costs;
- productivity losses to employers and society;
- administrative, management and legal costs incurred by employers; and
- government expenditure on sick pay and disability benefits.

Based on these conservative estimates of the cost of cancer deaths, and these lower and upper estimates of the number of cancer deaths likely to be prevented through the REACH system, the cancer-related economic benefits of implementing the system over 30 years were estimated by Postle *et al.* (1) to be between about €18 billion and €54 billion (Table 1).

On the other hand, the economic benefits of implementing the REACH system for all non-cancer diseases combined were estimated to be between €23 million and €225 million, despite the over-optimistic assumption that the number of these cases related to unknown chemicals would be effectively reduced to zero.

These estimates do not represent the total economic value of the REACH system, because the positive effects for wider public health and the environment have not been incorporated. However, it seems clear that the prevention of cancer is of greater potential benefit to occupational health than the prevention of all other diseases combined,

when considered from an economic viewpoint alone. The same might also be reasonably expected of the human costs. Consequently, the accurate identification of occupational carcinogens to which chemical and downstream industry workers are exposed must be a top priority for the REACH system.

Due to a paucity of human clinical data, the identification of chemical carcinogens has conventionally relied heavily on information provided by animal tests. However, are animal bioassays truly predictive for human carcinogenicity? The potential savings of thousands of lives annually within the EU, and billions of euros in human and productivity costs, clearly demand the use of carcinogenicity assays that offer the best possible human predictivity.

To examine the human predictivity of animal carcinogenicity data, and its utility in deriving human carcinogenicity classifications for regulatory purposes, we surveyed the chemicals contained within the Integrated Risk Information System (IRIS) toxic chemicals database maintained by the US Environmental Protection Agency (EPA). The EPA is the federal agency most responsible for protecting Americans from environmental contaminants, and its IRIS database contains human carcinogenicity assessments of the chemicals of greatest US public health concern, along with the data from which those assessments are derived.

Methods

We examined the chemicals contained within the IRIS database that lacked significant human exposure data but possessed animal carcinogenicity data (the great majority), and that had received a human carcinogenicity assessment by 1 January 2004. We determined the proportion for which the EPA was able to derive the classifications of probable human carcinogen or non-carcinogen, based primarily on their animal carcinogenicity data. A 95% confidence interval (CI) for this proportion was derived via the modified Wald method, which is said to provide more-accurate results than the so-called “exact” method which is commonly used (2).

To investigate the impact of these factors on the human utility, or otherwise, of the animal carcinogenicity data, we examined the species and routes of administration used, and the organ systems affected.

Table 1: Estimated EU economic benefits over 30 years of reducing annual cancer deaths through the implementation of the REACH system

	Lower predicted benefit (2167 deaths prevented annually)	Upper predicted benefit (4333 deaths prevented annually)
Lower estimate of 30-year benefit	17,592 (€ million)	35,183 (€ million)
Best estimate of 30-year benefit	27,083 (€ million)	54,167 (€ million)

Data source: Postle *et al.* (1).

To assess the reliability of the EPA carcinogenicity assessments, we compared them with those of the World Health Organisation's International Agency for Research on Cancer (IARC). Of 128 chemicals assigned human carcinogenicity classifications by both agencies, 17 were considered by the EPA to possess at least limited human data, while 111 were primarily reliant on animal data for their classifications. The consistency of classifications between the EPA and IARC was examined for each of these two groups, by comparing the carcinogenicity classification proportions within each group via chi-squared tests, and by comparing the individual classifications of the 111 chemicals primarily reliant on animal carcinogenicity data for their classifications.

Chi-squared tests provide a statistical calculation of the probability that two data sets, such as the EPA and IARC human carcinogenicity classifications, are samples from the same underlying data population, and that any observed differences are simply due to random sampling variation. Large chi-squared values (χ^2) reflect increased probabilities that observed differences are due to real differences in underlying data populations (3).

Chi-squared and two-tailed *p* values were derived from the online statistical calculators available at <http://www.graphpad.com/quickcalcs/index.cfm>.

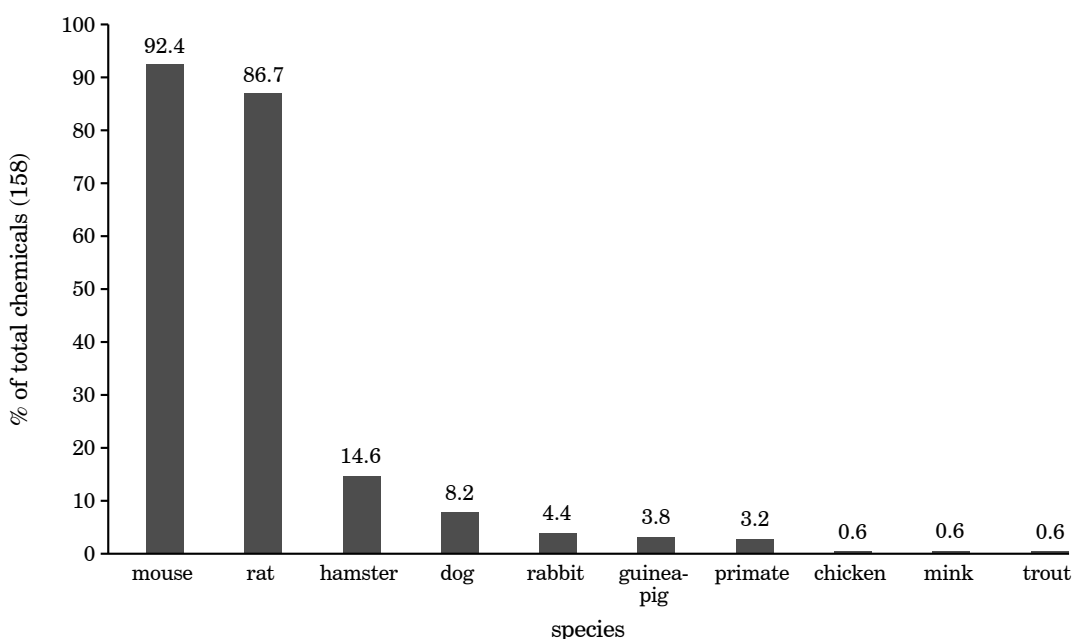
Results

For 93 (58.1%; 95% CI: 50.4–65.5) of the 160 chemicals lacking even limited human data but possessing animal data, and which had received human carcinogenicity assessments, the EPA considered the animal data inadequate to support the substantially useful classifications of probable human carcinogen or probable human non-carcinogen (3).

The species used were available for 158 of these chemicals. At least 10 different species were used, namely: chicken, dog, guinea-pig, hamster, mouse, mink, non-human primate (one macaque, three unspecified "monkey" species, and one unspecified "primate" species), rabbit, rat, and trout. The three species most commonly used (Figure 1; 4) were the mouse (92.4%), rat (86.7%), and hamster (14.6%).

The routes of administration used were available for 156 of the chemicals. Twelve non-oral routes of administration, and a variety of oral routes, not always specified, were used. They were: dermal, inhalation, intramuscular, intraperitoneal, intrapleural, intrarenal, intratesticular, intravenous, oral: food, oral: gavage, oral: water, oral: other (for example, capsule, toothpaste additive), oral: unspecified, subcutaneous, surgical implantation, transplacental, and vaginal painting. Those most commonly used (Figure 2; 4) were food (49.4%), gavage (33.3%), and dermal administration (26.3%).

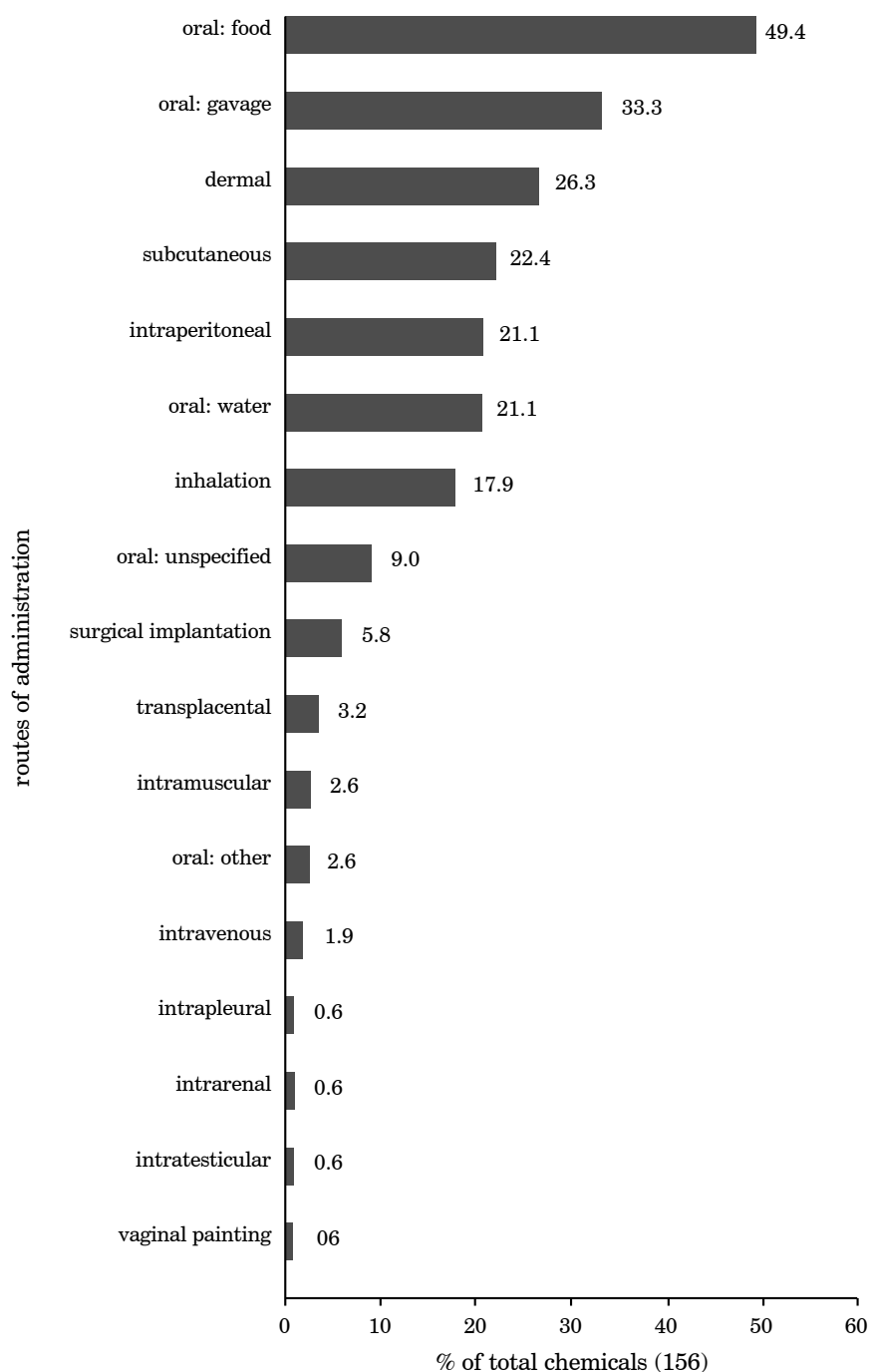
Figure 1: Species used with assessed EPA chemicals lacking significant human data but possessing animal data



Data source: The EPA Integrated Risk Information System database, 1 January 2004.

Reproduced from Knight et al. (4).

Figure 2: Routes of administration for assessed EPA chemicals lacking significant human data but possessing animal data



Data source: The EPA Integrated Risk Information System database, 1 January 2004.

Reproduced from Knight *et al.* (4).

Other routes of major interest were drinking water (21.1%), and inhalation (17.9%).

Chemicals considered probably not carcinogenic to humans were not, of course, known to exhibit significantly neoplastic lesions. For those chemicals

considered unclassifiable, it was frequently difficult to establish whether or not significant treatment-related results occurred. However, for the remaining 104 chemicals, considered probable or possible human carcinogens, up to 43 organs or organ sys-

tems were found to exhibit neoplastic lesions, with those most commonly affected (Figure 3; 4) being the liver (66.3%), the lung (31.7%), and the kidney, skin and stomach (all 17.3%). It should be noted that differentiation between primary and metastatic tumours was often impossible, so all tumours were included, even when, infrequently, they were identified as metastases.

128 chemicals with human or animal data were assessed by both the EPA and the IARC. Of these 128, human carcinogenicity classifications were similar only for those 17 with significant human data ($\chi^2 = 0.291$, 1 df, $p = 0.5896$; Table 2). Note that chi-squared analysis does not allow comparison when one category lacks any data, hence acrylonitrile, assessed as the only possible human carcinogen by the IARC, but as a probable human carcinogen by the EPA, was excluded, yielding a more conservative result.

For the 111 classifications primarily reliant on animal data, the EPA was much more likely to assign carcinogenicity classifications indicative of greater human risk ($\chi^2 = 215.548$, 2 df, $p < 0.0001$; Figure 4; 3). To permit a chi-squared analysis, methacrylate, assessed as unclassifiable by the IARC, but as the only probable human non-carcinogen by the EPA, was excluded, yielding a more conservative result.

The data reveal that the EPA was much more likely than the IARC to assign carcinogenicity classifications indicative of greater human hazard (3). The numbers of chemicals classified by the EPA as probable human carcinogens (60 chemicals) compared to all other categories (51 chemicals) were very significantly different from those predicted by IARC figures of 12 and 99 respectively ($\chi^2 = 215.273$, 1 df, $p < 0.0001$). Similar disparities were found for possible human carcinogens ($\chi^2 = 19.771$, 1 df, $p < 0.0001$) and unclassifiable chemicals ($\chi^2 = 24.378$, 1 df, $p < 0.0001$).

Comparison of the individual classifications of these 111 chemicals revealed that 67 (60.4%) were assigned an EPA carcinogenicity classification

indicative of greater human hazard, 38 (34.2%) were assigned an equivalent classification, and 6 (5.4%) were assigned a classification indicative of lower human hazard, than the corresponding IARC classifications of the same chemicals (3).

Discussion

Based on the EPA figures alone, the predictivity of animal carcinogenicity data for human hazard, and hence its utility in deriving substantially useful human carcinogenicity classifications for regulatory or other purposes, is clearly questionable. Of those 160 IRIS chemicals lacking even limited human data but possessing animal data, the EPA considered the animal data inadequate to support the substantially useful classifications of probable human carcinogen or non-carcinogen, in the majority (93, 58.1%) of cases. Classifications of definite human carcinogen relied on the existence of convincing human data. Classifications of unclassifiable or possible human carcinogen were not considered substantially useful for risk assessment or regulatory purposes, and are excluded from the US National Toxicology Program (NTP) annual *Report on Carcinogens* (5).

A wide variety of species were used in these carcinogenicity bioassays, with rodents predominating; a wide variety of routes of administration were used; and a particularly wide variety of organ systems were affected. The likely causes of the poor human predictivity of these bioassays include: 1) the profound discordance of bioassay results between rodent species, strains and genders, and between rodents and human beings; 2) the substantial stresses caused by handling, restraint and stressful routes of administration, with consequent effects on immunocompetence and predisposition to carcinogenesis; 3) the differences in transport mechanisms and rates of absorption between test routes of administration and other important routes of human exposure;

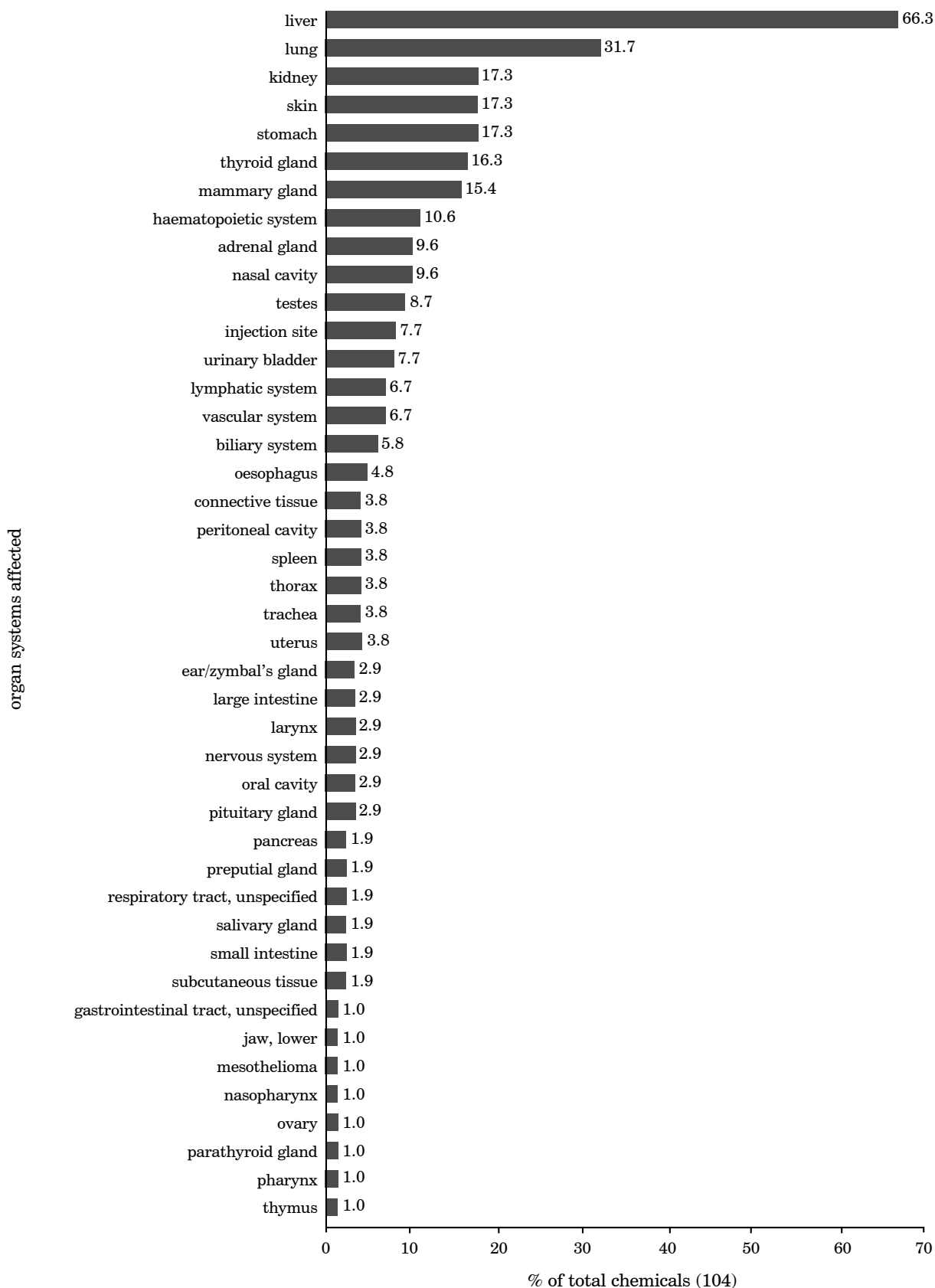
Table 2: IARC classifications of EPA chemicals possessing significant human data (EPA categories A or B1)

Human carcinogenicity classification	EPA	IARC
Human Carcinogen (A)	11	12
Probable Human Carcinogen (B1)	6	4
Possible Human Carcinogen	0	1
Total	17	17

Data sources: The EPA Integrated Risk Information System database, 1 January 2004, and the IARC Monographs Programme on the Evaluation of Carcinogenic Risks to Humans, Volumes 1–82, 1 January 2004.

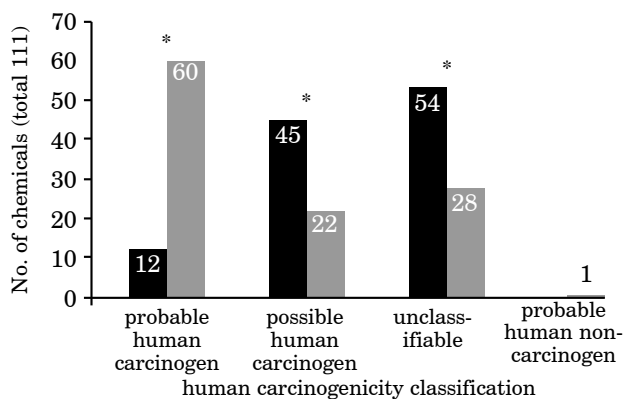
Reproduced from Knight et al. (3).

Figure 3: Organ systems affected by chemicals assessed by the EPA to be probable (B2) or possible (C) human carcinogens



Data source: The EPA Integrated Risk Information System database, 1 January 2004.
 Reproduced from Knight et al. (4).

Figure 4: EPA and IARC human carcinogenicity classifications of chemicals considered by the EPA to lack human data but to possess animal data



■ = IARC; ■ = EPA; * = $p < 0.0001$.

Data sources: The EPA Integrated Risk Information System database, 1 January 2004, and the IARC Monographs Programme on the Evaluation of Carcinogenic Risks to Humans, Volumes 1–82, 1 January 2004.

Reproduced from Knight *et al.* (3).

4) the considerable variability of organ systems in response to carcinogenic insults, within and between species; and 5) the inherent predisposition of chronic high-dose bioassays toward false positive results, due to the overwhelming of physiological defences, and the unnatural elevation of cell division rates during *ad libitum* feeding studies (4). These factors, reviewed in detail in Knight *et al.* (4), render profoundly difficult any attempts to accurately extrapolate human carcinogenic hazard from animal data.

Furthermore, the IARC assessments for the same chemicals indicate that the human predictivity of animal carcinogenicity data is probably even poorer than that indicated by the EPA figures alone. The EPA and IARC carcinogenicity classifications were similar only for chemicals possessing human data. For the remainder (the great majority), the EPA was much more likely than the IARC to assign carcinogenicity classifications indicative of greater human hazard. The IARC is recognised as one of the most authoritative sources of information on potential human carcinogens (6, 7), and it is implausible that IARC assessments would, in general, be inaccurate or based on incomplete data. Consequently, the significant differences in human carcinogenicity classifications of identical chemicals between the IARC and the EPA indicate that: 1) in the absence of significant human data the EPA is

over-reliant on animal carcinogenicity data; 2) as a result, the EPA tends to over-predict carcinogenic risk; and 3) the true predictivity for human carcinogenicity of animal data is even poorer than is indicated by EPA figures alone (3).

EPA human carcinogenicity classifications appear to be less scientifically-based than those of the IARC, due to: 1) the varying depth of EPA assessments, due to resource constraints; 2) the less rigorous standards required of data incorporated into EPA assessments; and 3) EPA public health-protective policy, which errs on the side of caution by assuming that tumours in animals are indicative of human carcinogenicity (3).

Alternatives to the Bioassay

Conventional animal carcinogenicity tests take around three years to design, conduct and interpret (8). They have so far consumed hundreds of millions of dollars (9), millions of skilled personnel hours (10), and millions of animal lives (8, 10). However, several investigations (6, 11–17) have illustrated the poor human specificity, and hence, poor predictivity, of animal carcinogenicity data. Clearly, more-predictive alternatives to the conventional rodent bioassay are required, particularly for use in large-scale testing programs such as the REACH system.

Proposed bioassay modifications have included: the elimination of mice; the use of genetically-altered mice that exhibit the altered expression of genes mechanistically relevant to carcinogenesis; neonatal mice; decreased timeframes; initiation-promotion models, which employ non-carcinogenic promoters to speed the effects of carcinogenic initiators, or vice-versa; greater incorporation of toxicokinetic and toxicodynamic assessments; computerised quantitative structure-activity relationship (QSAR) systems, which predict and quantify carcinogenic effects based on the presence of electrophilic molecular substructures or other chemical moieties; enhanced *in vitro* assays; cDNA microarrays containing hundreds or thousands of complementary DNA transcripts of mRNA templates, which can be used to detect genetic expression changes caused by carcinogens (toxicogenomics); limited human clinical trials; and epidemiological research (18).

Based on our detailed review of these bioassay alternatives in Knight *et al.* (18), we propose the replacement of the conventional carcinogenicity bioassay with the following protocol, based on a tiered combination of alternative assays.

1. Before any assay is conducted, all existing information about the test compound should be collated and reviewed in a critical and unbiased fashion, to determine which tests are scientifically justified.

2. Initial screens should include computerised QSAR systems, cell or tissue cultures, and cDNA microarrays, where possible. QSAR systems should be used to identify and estimate the toxic effects of specific chemical groups. Appropriate *in vitro* screening assays, such as the Ames *Salmonella*, Syrian Hamster Embryo cell transformation, *Saccharomyces* GreenScreen, and human basal and target organ cell or tissue culture assays, should be fully employed to seek evidence of cytotoxicity, mutagenicity and genotoxicity. Carefully chosen and well-conducted cDNA microarray assays for genotoxicity and non-genotoxicity should be used for the detection of changes in genetic expression.
3. Following these initial screens, human toxicological studies with barrier models and biological simulations, and microdosing and non-invasive biomarkers, should be appropriately selected to model toxicokinetics and estimate target organ concentrations.
4. In the case of human pharmaceuticals, and non-pharmaceuticals for which a human carcinogenicity assessment is also considered of high importance, and for which human carcinogenicity or other toxicity is not already suggested on the basis of data acquired through the application of all the other methods specified in stages 1–3, above, limited human trials involving fully informed and consenting volunteers (phase I and II human clinical trials in the case of pharmaceuticals) might be conducted, albeit with considerable caution, commencing with microdoses.

The further development, validation and implementation of some of these alternative assays will no doubt require a redistribution of funding. However, the proper collation and examination of the more-targeted data obtained through such a testing protocol for evidence of carcinogenic risk factors such as genotoxicity, immunosuppression, hormonal activity or chronic irritation/inflammation, is likely to yield a weight-of-evidence characterisation of superior human predictivity when compared to that offered by the conventional rodent bioassay. Additional advantages include the likelihood of greater insights into mechanisms of carcinogenicity, and substantial saving of financial, human and animal resources (18).

The impending demands of the REACH chemicals testing system are unparalleled in EU history. Consequently, the further development, validation and implementation of non-animal carcinogenicity assays must be accorded the highest priority by both chemical regulators and the chemical industry.

Acknowledgement

This research was partly funded by the Physicians Committee for Responsible Medicine, Washington DC, USA.

References

1. Postle, M., Vernon, J., Zarogiannis, P. & Salado, R. (2003). *Assessment of the Impact of the New Chemicals Policy on Occupational Health: Final Report: prepared for European Commission — Environment Directorate-General*. 96pp. Lodon, Norfolk, UK: Risk & Policy Analysts Limited.
2. Agresti, A. & Coull, B.A. (1998). Approximate is better than “exact” for interval estimation of binomial proportions. *The American Statistician* **52**, 119–126. The modified Wald method is also described at <http://www.graphpad.com/quickcalcs/ConfInterval2.cfm>.
3. Knight, A., Bailey, J. & Balcombe, J. (2006). Animal carcinogenicity studies: 1. Poor human predictivity. *ATLA* **34**, 19–27.
4. Knight, A., Bailey, J. & Balcombe, J. (2006). Animal carcinogenicity studies: 2. Obstacles to extrapolation of data to humans. *ATLA* **34**, 29–38.
5. National Toxicology Program (2002). *National Toxicology Program Report on Carcinogens, 10th Edition*. Website <http://ntp.niehs.nih.gov/ntpweb/index.cfm?objectid=72016262-BDB7-CEBA-FA60E922B18C2540> (Accessed 13.01.05).
6. Tomatis, L. & Wilbourn, J. (1993). Evaluation of carcinogenic risk to humans: the experience of IARC. In *New Frontiers in Cancer Causation* (ed. O.H. Iversen), pp. 371–387. Washington, DC, USA: Taylor and Francis.
7. International Agency for Research on Cancer. (undated). *IARC Monographs Programme on the Evaluation of Carcinogenic Risks to Humans*. Website <http://monographs.iarc.fr> (Accessed 01.01.04).
8. Monro, A.M. & MacDonald, J.S. (1998). Evaluation of the carcinogenic potential of pharmaceuticals. Opportunities arising from the International Conference on Harmonisation. *Drug Safety* **18**, 309–319.
9. Ashby, J. (1996). Alternatives to the 2-species bioassay for the identification of potential human carcinogens. *Human and Experimental Toxicology* **15**, 183–202.
10. Gold, L.S., Manley, N.B., Slone, T.H. & Rohrbach, L. (1999). Supplement to the Carcinogenic Potency Database (CPDB): Results of animal bioassays published in the general literature in 1993 to 1994 and by the National Toxicology Program in 1995 to 1996. *Environmental Health Perspectives* **107**, Suppl. 4, 527–600.
11. Hogan, K.A. (2000). *Characterization of Data Variability and Uncertainty: Health Effects Assessments in the Integrated Risk Information System (IRIS)*. EPA/635/R-00/005A. 27pp. National Center for Environmental Assessment, Washington, DC, USA: Environmental Protection Agency, Office of Research and Development, Environmental Protection Agency.
12. Haseman, K. (2000). Using the NTP database to assess the value of rodent carcinogenicity studies for determining human cancer risk. *Drug Metabolism Reviews* **32**, 169–186.

13. Gold, L.S., Slone, T.H. & Ames, B.N. (1998). What do animal cancer tests tell us about human cancer risk? Overview of analyses of the carcinogenic potency database. *Drug Metabolism Reviews* **30**, 359–404.
14. Rall, D.P. (2000). Laboratory animal tests and human cancer. *Drug Metabolism Reviews* **32**, 119–128.
15. Ashby, J. & Purchase, I.F. (1993). Will all chemicals be carcinogenic to rodents when adequately evaluated? *Carcinogenesis* **8**, 489–495.
16. Fung, V., Barrett, J. & Huff, J. (1995). The carcinogenesis bioassay in perspective: Application in identifying human hazards. *Environmental Health Perspectives* **103**, 680–683.
17. Ennever, F.K. & Lave, L.B. (2003). Implications of the lack of accuracy of the lifetime rodent bioassay for predicting human carcinogenicity. *Regulatory Toxicology and Pharmacology* **38**, 52–57.
18. Knight, A., Bailey, J. & Balcombe, J. (2006). Animal carcinogenicity studies: 3. Alternatives to the bioassay. *ATLA* **34**, 39–48.