The Knowledge Engineering Review, Vol. 14:4, 1999, 307–317
Printed in the United Kingdom

Application of artificial intelligence and computer-based methods to predicting chemical toxicity

ANN M. RICHARD

Maildrop 68, National Health & Environmental Effects Research Laboratory, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711, USA (email: richard.ann@epa.gov)

Abstract

The toxicity prediction problem lies squarely at the interface of biology, chemistry and computational domains and will require the integrated application of knowledge and approaches from each domain for its solution. It is not a single modeling problem, but rather multiple levels of modeling compartments derived from the goals of toxicity risk assessment. These compartments include different categories and characteristics of toxicity (e.g., cancer vs. non-cancer, acute vs. delayed) and, therein, different levels of biofunctional organization and multiple mechanisms of toxicity extending to the level of individual chemical structures. A toxicity prediction model should strive to resolve the global toxicity prediction problem to local modeling compartments that reflect coherent biofunctional mechanisms and common modes of action while retaining some level of useful generalizations. This paper attempts to use these general concepts to frame the challenges and opportunities for application of Artificial Intelligence (AI) and computer-based methods to the goal of chemical toxicity prediction.

1 Introduction

E. O. Wilson asserts that the most difficult yet pressing challenges in science today lie not within established academic disciplines (e.g., chemistry, toxicology, computer science) but rather at the boundaries and intersections of different disciplines, each of which is associated with its own language, tools, and modes of thought (Wilson, 1998). ''Toxicity prediction'' could easily be the poster child of this thesis. However, this all-encompassing term is too conceptually vague to serve as a reasonable starting point for discussion. To consider the opportunities and the most appropriate means for tackling the challenge of toxicity prediction by Artificial Intelligence (AI), or any other computerized approaches, we must first clearly delineate the nature of the problem. Secondly, we must strive to apply objective computerized approaches to developing scientifically credible models that will build upon and reflect the chemical and biological underpinnings of the problem.

2 The problem

Diverse types of information relative to toxicity risk characterization of chemicals (Johannsen, 1990) can be represented schematically as layers of an onion, with the chemical of concern at the center of the onion and the human-health risk assessment at the outermost layer (Figure 1). The intervening layers range from calculated or measured physical properties, to mechanistically defined biochemical interactions, to *in vitro* cell-culture bioassays, to *in vivo* responses in whole animals and populations. The radius of each onion layer increases according to the level of biological complexity and proximity to human-health risk assessment, whereas the boundaries between the layers separate physically and conceptually distinct types of information. Structure-activity relationships (SAR) are models that attempt to extrapolate from the center of the onion, i.e., the chemical structure, to different types of biological endpoints (Figure 2). Activity-Activity Relationships (AAR) are models

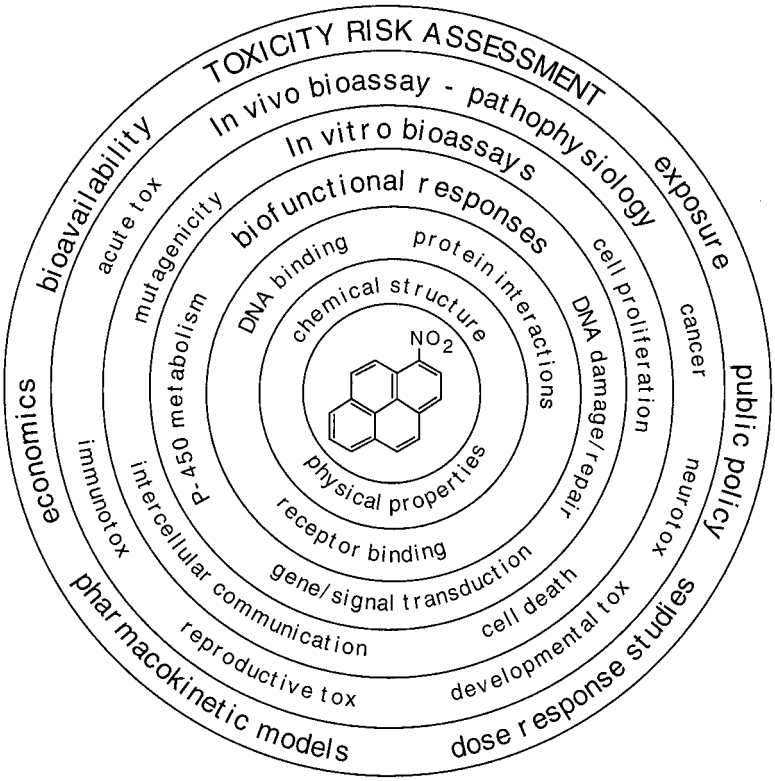

Figure 1 Layers of chemical and biological organization relevant to the toxicity prediction problem.

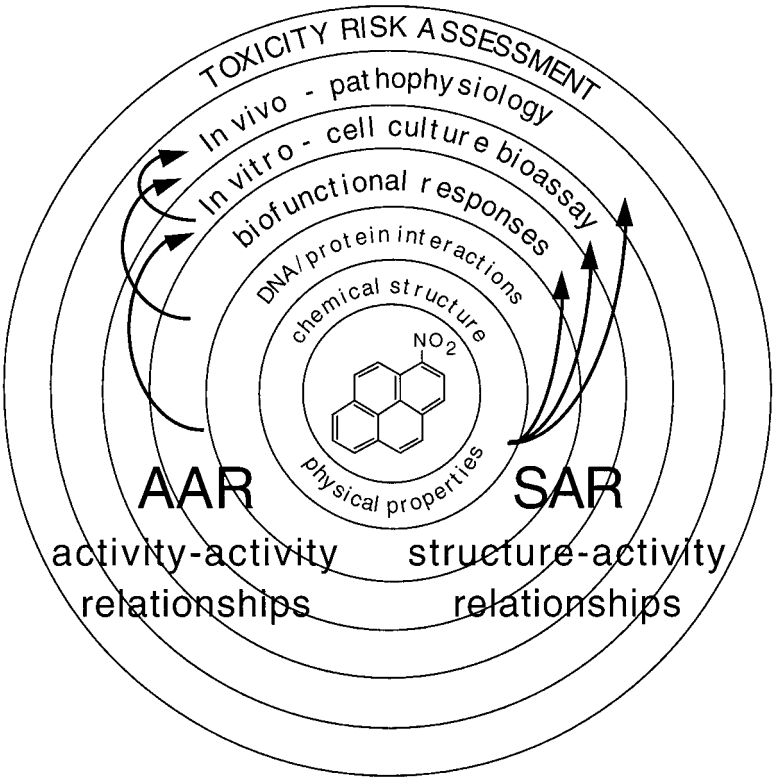

Figure 2 Types of information being related in SAR and AAR models.

that attempt to extrapolate from one level of biological information to another (Figure 2). SAR methods utilize only the chemical structure and associated properties of an isolated molecule to construct predictive models and, hence, are most useful for hazard identification and screening of chemicals for which bioassay data is unavailable. AAR methods, in contrast, utilize bioassay data that incorporate information on the interaction of the chemical within the biological system as the basis for predictive models. Hybrid methods attempt to combine both structure/property data and bioassay information in toxicity prediction models. In general, the more intervening layers there are in Figure 2 relative to the SAR or AAR prediction model, the greater the extrapolation distance across levels of biological organization, and the greater uncertainty in model prediction. Bioassay data also tend to increase in cost and decrease in availability as one moves to more integrated whole animal systems and higher levels of biological complexity. In addition, note that the simplistic representation of the toxicity prediction problem in Figure 1 becomes considerably more complex if one attempts to represent species-to-species extrapolation (e.g., rodent to human) or to consider confounding factors, such as chemical mixtures, sensitive subpopulations, etc.

The goal of a global toxicity prediction model is to be able to predict any type of potential toxicity for any type of chemical. In support of a toxicity risk assessment, the even larger goal is to consider a toxicity prediction along with factors such as exposure, bioavailability, and biodegradation that have the potential to modulate the toxicity risk. These ambitious goals can be broken down into modeling compartments for the factors that modulate toxicity risk, as well as for the major categories of toxicity (Figure 3). Within each category of toxicity, models can be further compartmentalized according to bioassay, toxicity mechanism, chemical class, etc. (Figure 4). Hence, the toxicity modeling challenge has two major components: (1) the development of distinct models within toxicity compartments corresponding to common biological mechanisms and effects; and (2) the appropriate linkage of these compartments to support an overall toxicity assessment. These compartments and linkages are more than a modeling convenience; they represent attempts to delineate and relate scientifically coherent information and to resolve the modeling challenge to levels of distinct biofunctional processes. Given the context and potential use of a toxicity modeling compartment in the larger toxicity risk assessment scheme of Figure 1, the remainder of this paper

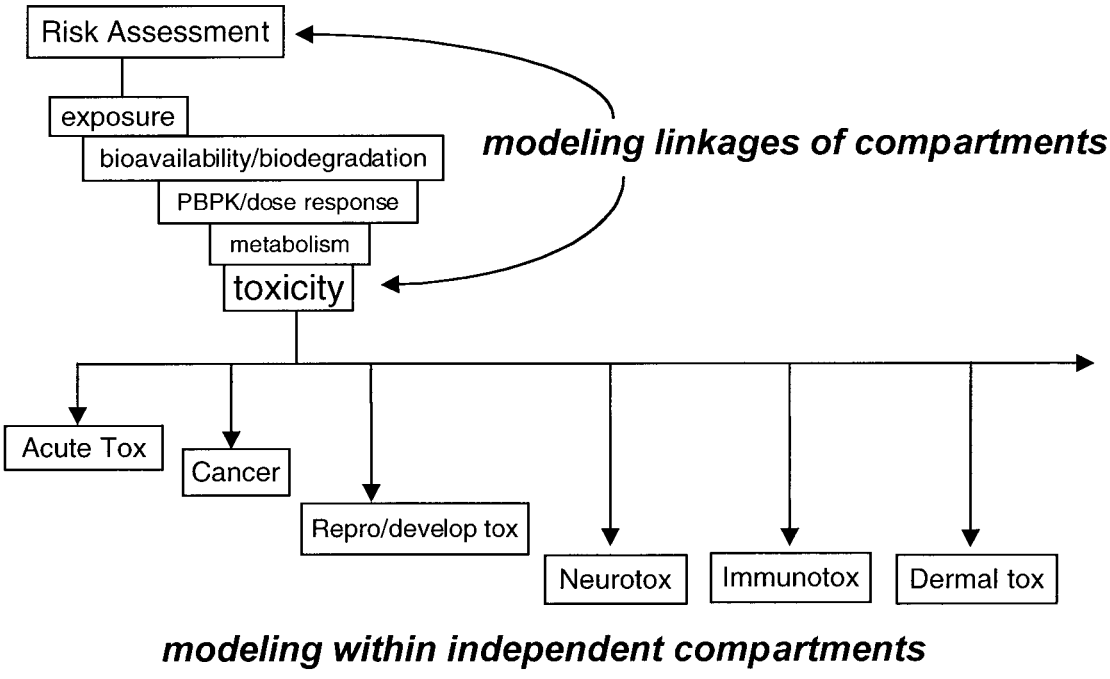

Figure 3 Schematic of modeling compartments, including major categories of toxicity, and linkages relative to the toxicity prediction problem.

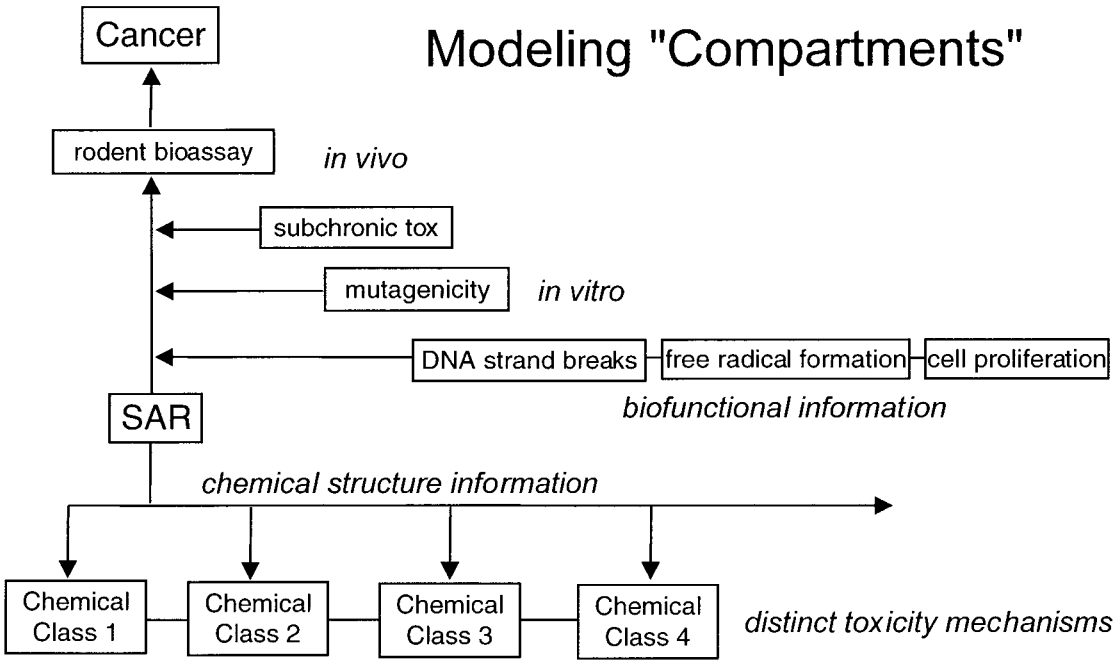

Figure 4 Schematic of modeling compartments within a sample category of toxicity: cancer.

will focus more specifically on problems and challenges associated with modeling efforts within distinct toxicity compartments.

Each category of toxicity presents a different, largely independent modeling challenge. Most major toxicity categories are represented by multiple assay systems, corresponding to endpoints that are mechanistically ill-defined and associated with limited data and knowledge bases for use in modeling efforts. Within some toxicity categories, e.g., neurotoxicity and immunotoxicity, the problems may extend even further to a lack of consensus and standardization of bioassays and toxicity measures among the toxicology domain experts, measures that are necessary before useful data bases can be constructed for modeling purposes.

Cancer offers an example of a toxicity endpoint, of prominent concern from a regulatory and risk assessment standpoint, for which relatively large, standardized databases have been made publicly available for use in modeling efforts (see, e.g., the Carcinogenicity Potency Data Base of Gold et al. (1991), available at `http://potency.berkeley.edu/hybrid.html`). The general ability of a chemical to act as an electrophile or alkylating agent in reaction with DNA provides a conceptually useful justification for grouping chemically diverse structures in efforts to model the cancer endpoint (Singer, 1983). However, there are many possible mechanisms for forming DNA reactive intermediates, for producing DNA damage or mutations, or for repairing genomic damage. In addition, a number of non-DNA reactive, epigenetic mechanisms of tumor formation have been postulated (e.g., involving cell proliferation, changes in apoptosis, alterations in gap junction communication, disruption of hormone signaling pathways, etc.) (e.g., see Trosko, 1997). Cancer is, in truth, neither a single endpoint nor a single disease but, rather, a multitude of causes, modulating factors, mechanisms, and disease outcomes relating to tumor formation (Arcos, 1995). Hence, the cancer modeling compartment can be considered poorly resolved from a mechanistic biofunctional point of view.

Furthermore, the cancer endpoint may be subject to interpretation and politics when attempts are made to extrapolate from a battery of rodent bioassays to humans. For example, a number of cancer prediction models, including the CASE (Klopman & Rosenkranz, 1994) and TOPKAT (Enslein et al., 1994) commercial SAR models (discussed below), have been constructed based on the consensus "calls" for the results of the National Toxicological Program (NTP) rodent cancer bioassay. To reach a consensus, an expert NTP panel evaluates the 2-year rodent bioassay data and catagorizes a

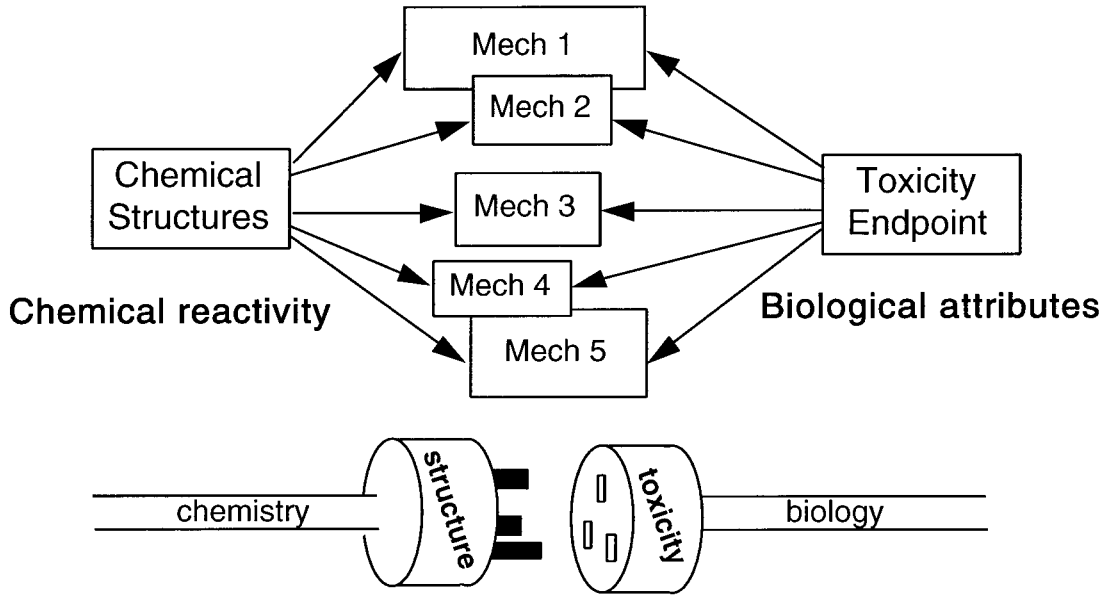

Figure 5 Resolution of toxicity prediction SAR problem to discreet mechanistic compartments, enabling linkage of underlying chemical and biological elements.

chemical as a likely carcinogen if it causes tumors, by statistically significant measure, at a single site in any one of four rodent species/sex groups. Recently a new TOPKAT SAR model has been developed to predict qualitative (+/–) chemical carcinogenicity based on a more stringent FDA definition that labels a chemical as a carcinogen only if it causes tumors at multiple organ sites or in multiple rodent species (Enslein, 1999). These two definitions of a carcinogen, when used as the basis of model development, yield two distinct models for prediction. Hence, as we move closer to risk assessment concerns, the toxicity measure used in modeling may incorporate subjective, weight-of-evidence considerations in its definition. This further distances the toxicity that we are attempting to predict (e.g., human carcinogen) from the underlying chemical or biological basis for the toxicity that has been actually measured (e.g., male rat carcinogen).

Further limitations on efforts to predict toxicity endpoints from chemical and biological attributes are due to the inherent complexity and indeterminacy of chemical-biological interactions. Because toxicity endpoints frequently encompass many possible biofunctional mechanisms, different chemicals often act through different mechanisms to produce the same apparent toxicity outcome. In contrast, the basic assumption underlying an SAR study is that differences in biological activity can be mapped to differences in chemical structure provided that a common mechanism or rate-determining step applies. *Hence, it follows that the SAR problem must be resolved to the level of toxicity mechanisms, or common modes of action, to enable the mapping of chemical structure to toxicity endpoint*. This concept is represented schematically in Figure 5, and is a crucial point that should serve to constrain and guide modeling strategies for toxicity prediction. Most importantly, this is the level of problem resolution where a mechanism-based rationale can be built, where the chemical and biological functional elements of toxicity can be potentially linked, and where scientific credibility of these linkages can be established.

3 Model development

The first task in development of a toxicity prediction model is clear definition of the problem. What is the nature of the toxicity endpoint to be modeled? What is the quantity and quality of the available data relative to that endpoint? What is known of the biological mechanisms, chemical interactions, and structural basis for the activity? And what is the ultimate goal of the predictive

modeling effort, i.e., to support risk assessment, to prioritize chemicals for testing, to generate insight, for basic research, etc.?

Each modeling effort can be divided into three stages – data input, model development, and data output. At each stage, the modeler has opportunities to incorporate biological and chemical knowledge to shape and inform the modeling effort. Such information can be used to classify chemicals according to function prior to analysis (e.g., DNA reactive, requiring oxidative metabolism, forming free radicals, causing cell proliferation, etc.). Modelers can choose mechanistically informative molecular descriptors or biological attributes that can potentially relate to and reveal the structural basis for an activity (e.g., size and shape parameters affecting bioavailability or ligand binding, electronic properties of a reactive site, energies of formation of reactive intermediates, enzyme substrate affinities, etc.). Finally, models can be designed to generate predictions that can be traced back to the original model assumptions and that potentially lend themselves to interpretation and mechanistic rationalization (e.g., referenced to clearly stated rules, properties, or putative molecular analogues within the modeled toxicity database).

The modeling effort is also defined by the nature of available data and the corresponding goal of developing qualitative (+/–) or quantitative (relative potency) toxicity prediction models (see Figure 6). A global classification model is qualitative and attempts to discern chemical or biological attributes that provide gross separation of "active" chemicals from "inactive" chemicals. Methods such as linear discriminant analysis or Bayesian statistics can be used to separate data into two activity classes, whereas clustering techniques can potentially distinguish multiple "islands" of activity from the "sea" of inactive molecules. Quantitative models (e.g., linear regression), in contrast, are local in nature and are used to model relative potencies or differences in activity among chemicals belonging to the same activity class, i.e., where all molecules are assumed to satisfy the same basic criteria for activity according to a common mechanism or mode of action. Application of global qualitative classification models should, therefore, logically precede application of local quantitative models since the latter presume that all molecules belong to the same activity class (Benigni et al., 1998).

A fundamental challenge in modeling major categories of toxicity, such as cancer, is in reconciling

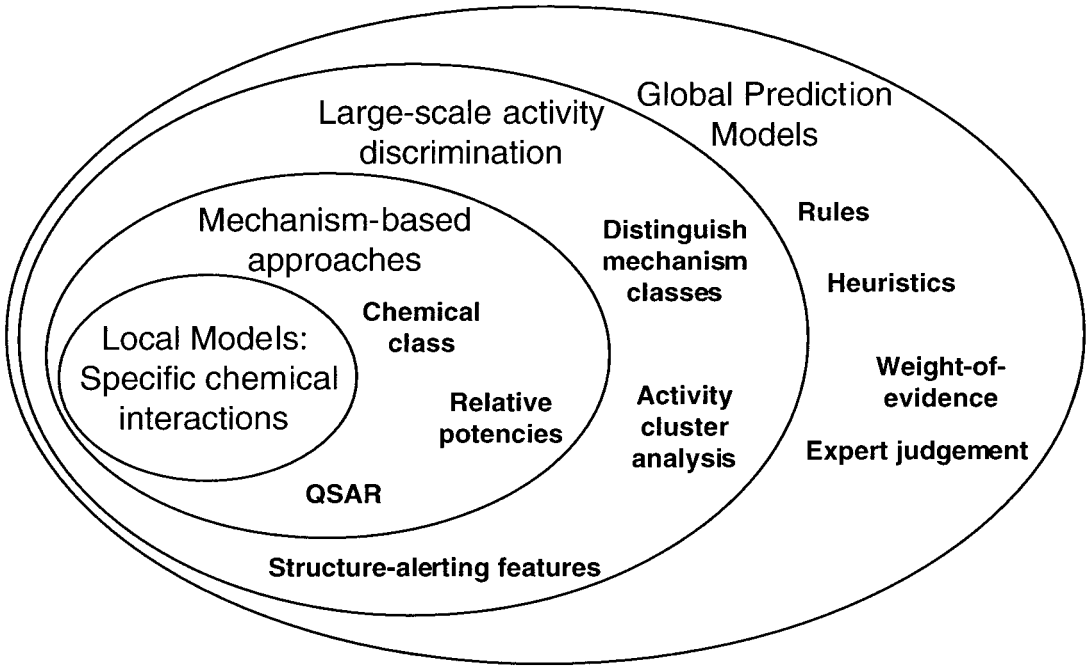

Figure 6 Schematic representation of different types of toxicity prediction models, ranging from global, qualitative models to local, quantitative models.

the global and local modeling viewpoints. For many reasons – data limitations, the goal of general prediction, lack of knowledge of mechanism-based classes – a number of toxicity prediction models have been built from structurally diverse (i.e., non-congeneric) data sets that are likely to encompass a wide variety of toxicity mechanisms. Such modeling efforts must either attempt to extract, by unbiased means, or must presume *a priori* a mechanism-based organization of the data, e.g., through the use of biofunctional information, structural "alerts" (i.e., common structural fragments associated with activity), and/or concepts of functional classes derived from organic chemistry. If efforts to resolve the global prediction problem to the level of mechanistically-distinct classes are unsuccessful, then efforts to uncover the finer elements distinguishing activity and potency within and across mechanism-based classes are likely to fail.

4 Commercial systems

Existing commercial software programs have taken two distinct approaches to the problem of toxicity prediction. The MultiCASE/CASE (Klopman & Rosenkranz, 1994) and TOPKAT (Enslein, 1994) approaches rely upon statistical or automated algorithms for extracting SAR associations from existing data, with little or no prior application of expert judgement or organization of data according to mechanism or chemical class. The DEREK (Sanderson & Earnshaw, 1991; Greene, 1996), ONCOLOGIC (Woo et al., 1995) and HazardExpert (Darvas et al., 1999) approaches, in contrast, are rule-based systems built almost entirely upon prior knowledge, heuristics, expert judgement, and chemical and biological mechanism considerations. The CASE and TOPKAT systems have developed separate models for many different types of toxicity endpoints (e.g., SAL mutagenicity, rodent carcinogenicity, skin sensitization, acute tox, etc.). DEREK and HazardExpert have rules for multiple toxicity endpoints that can be applied simultaneously to prediction. HazardExpert has additional capability to estimate bioavailability and bioaccumulation. Finally, OncoLogic deals exclusively with carcinogenicity prediction within pre-defined chemical classes, and is the only program that has some ability to consider biological attributes in prediction. These systems differ considerably according to: the toxicity endpoints modeled, the information used in model development and prediction, the means for representing and generalizing chemical structure information, the feedback provided in support of a prediction, and user accessibility to the data or knowledge base used in model development (for further discussion, see Combes & Judson, 1995; Benfenati & Gini, 1997; Dearden et al., 1997; Taningher et al., 1997; Benigni & Richard, 1998; Richard, 1998a, 1998b; and references therein).

One is struck by the seemingly complementary nature of these commercial prediction systems and, yet, the almost complete independence of model application and development efforts. The automated statistical approaches incorporate little or no expert judgement of chemistry and toxicology domain experts and, to date, the expert systems have made little use of objective data analysis methods such as rule-induction. (A partial exception is HazardExpert, which has employed quantitative models for estimating bioavailability and bioaccumulation, and fuzzy logic to modulate toxicity structural alerts in prediction.) There have also been virtually no efforts to compare, combine and cross validate various commercial approaches for particular types of toxicity endpoints. Although there are legitimate reasons for the lack of such efforts, it engenders confusion among the potential user community concerning the relative merits and limitations of the various commercial systems. The general consensus is that a purely statistical association is insufficiently compelling to yield a confident toxicity prediction in the absence of mechanistic justification and supporting information (Richard, 1998a). Each individual toxicity prediction must be judged based not only on the statistical performance of the model (e.g., overall model sensitivity, specificity, accuracy), but also on the strength of the mechanistic or molecular analogy argument supporting the prediction. EPA's Revised Cancer Risk Assessment Guidelines, for example, specifically encourages the use of SAR information provided that it strengthens an argument based on "biological plausibility" and "mode of action consistent with generally agreed upon principles and understanding of carcinogenicity". In summary, commercial prediction programs at present offer

limited predictive capability pertaining to some categories of toxicity, subject to many caveats and restrictions.

On the plus side, commercial toxicity prediction systems are exploring different algorithms of toxicity prediction and laying the groundwork for continuing model development. Somewhat under-appreciated is the considerable effort that has been expended by commercial system developers to gather and scrutinize toxicity data and expertise relative to various modeled toxicity endpoints, and to link chemical structures and properties to these data. The convenient access to toxicity data, structures, and/or expertise that these programs provide offers means for identifying chemical analogues and generating potentially useful SAR hypotheses. Another value of such systems is for exploring global issues related to the larger structure-activity information content of the toxicity data bases, e.g., to enable comparisons between assay systems, different categories of toxicity, etc. (e.g., see Liu et al., 1996).

In addition to the above-mentioned efforts, other investigators and a few additional commercial systems are exploring the use of machine-learning techniques, rule-induction, and inductive logic reasoning for toxicity prediction (e.g., see King & Srinivasan, 1996; Lee et al., 1996; Gini & Katritzky, 1999; and references therein). Most of these efforts are focusing on relatively few data sets and toxicity endpoints. A prominent example is provided by the Predictive Toxicity Evaluation challenge (PTE-2), sponsored by the National Toxicology Program in which carcinogenicity predictions have been solicited from modelers for 30 chemicals currently undergoing testing in the two-year rodent bioassay. A data set containing information on the 30 chemicals and an array of structural and biological attributes for these, as well as for a training set of approximately 250 chemicals previously tested for carcinogenicity, have been made available to encourage and assist AI modelers in participating in the challenge (Bristol, 1996). It remains to be seen, however, whether this training set of chemicals is adequate for the task at hand, i.e., whether there are sufficient numbers of chemicals represented within mechanism-based classes to enable characterization and extrapolation from these classes. Benigni (1997), in summarizing the results of an earlier NTP prospective predictive challenge for the carcinogenicity of 44 chemicals, noted that all the predictive methods did reasonably well at determining gross structural alerting features conferring carcinogenicity, but most overestimated the number of positives, i.e., they failed to account adequately for refined features modulating (or eliminating) activity within the ''structural-alert classes''.

Also worthy of mention are efforts to develop general models for biofunctional processes, such as metabolism or bioavailability. Such models have the potential to be important components of toxicity prediction models for many categories of toxicity. The META (Klopman, 1994; Talafous et al., 1994) and MetabolExpert (Darvas, 1999) programs are AI, rule-based expert systems designed to predict the possible metabolic transformations of a chemical in a particular (e.g., rat) or generic (mammalian) biological organism (see also Benigni & Richard, 1998). Such programs provide information on a wide range of metabolic activation pathways but, at present, are limited by their inability to identify a few most likely metabolic pathways from among the many possible pathways. MetabolExpert has been incorporated into HazardExpert and META has been used in conjunction with the CASE/MultiCASE systems to evaluate the potential toxicity of chemicals along with their complement of possible metabolites (e.g., see Cunningham et al., 1996; Lewis et al., 1998). The COMPACT approach is another type of metabolism prediction system that uses quantitative structure-based modeling to predict relative substrate binding affinities of chemicals with cytochrome P-450 enzymes. This system has been used in conjunction with the HazardExpert system for refining carcinogenicity predictions based on structural alerting features (Lewis et al., 1998).

5 For the future

Accurate chemical toxicity prediction from AI and computerized approaches is a possible goal for the future. However, the need for tools to aid chemical toxicity screening and risk assessment is now, in the present, based on whatever limited knowledge and data resources are available to us (Richard, 1998b). So let us consider what steps modelers can take at present, in the absence of full knowledge

of toxicity mechanisms and within current limitations of data and understanding, to attempt to meet the toxicity prediction challenge. First, we can acknowledge the formidable problems and fundamental limits on solutions when attempting to model complex and variable biological responses, i.e., we can accept that absolute certainty in toxicity prediction of a novel chemical structure is unlikely to ever be an achievable goal. We can attempt to improve our definition of each toxicity prediction problem in terms of appropriate modeling compartments, decision elements, and biofunctional mechanisms to reflect more accurately the biological and chemical underpinnings of the problem. We can concentrate new method development and "proof of concept" of computer-based methods on the most well characterized toxicity endpoints, i.e., go where the data are and begin to evaluate objectively and compare the results of different methods. We can steer away from use of "computationally convenient" molecular descriptors devoid of chemical meaning (such as graph-theoretical indices), and modeling approaches that obscure rather than illuminate underlying mechanisms (such as some neural net applications); the goal should be to foster, not inhibit cross-talk between chemists, biologists and AI experts. We can begin to develop hybrid models that attempt to incorporate both objective analysis and expert judgement towards the goal of prediction, i.e., all relevant information should ultimately be brought to bear on the problem of prediction. We can consider every modeled toxicity prediction a hypothesis awaiting articulation, and every toxicity test result an opportunity to validate and refine prediction models. In other words, we can attempt to communicate more effectively the underlying assumptions and basis for a model prediction in a language that can be understood by AI experts, chemists, and toxicologists, and we can ask whether a test result supports or refutes model assumptions and hypotheses. Finally, we can view efforts to predict a particular bioassay result, from chemical structure and/or bioassay information, in the context of the larger toxicity risk assessment problem and begin to consider means for linking and weighing different types of information (such as multiple toxicity estimates, exposure estimates, bioavailability estimates, etc.) relevant to a toxicity risk assessment.

As for the future, new technologies likely to have the greatest impact on toxicity prediction efforts will come, not from new modeling algorithms or new molecular descriptors but, rather, from better resolution and description of toxicity endpoints and better linkages of chemical and biological processes. No amount of computer processing can compensate for fundamental data and knowledge limitations. An example of an emerging technology that is poised to radically alter the field of toxicology is that of cDNA micro arrays (Service, 1998). This technology, although still in the early stages of development and reproducibility, holds out the promise of monitoring the simultaneous expression of hundreds or thousands of genes that may be associated with different types and modes of toxicity. These signaling patterns are the molecular biological equivalent of a neural net, potentially capturing the true complexity and coherence of a toxicological effect in a functioning biological organism. From a toxicity modeling standpoint, the technology may ultimately offer unprecedented ability to organize biological response data according to function and to classify chemicals according to toxicity mechanisms. The challenge, however, will be to derive useful and reproducible generalizations from these complex patterns of genetic expression.

Other toxicity modeling opportunities on the horizon relate to large scale toxicity testing efforts. Two recent examples are the proposed toxicity testing battery for approximately 2500 High Production Volume (HPV) chemicals, and the proposed large scale screening of chemicals in commerce for endocrine disruption potential. For this latter effort, plans are to screen 8000–15,000 chemicals over a two year period in high-throughput estrogen receptor-binding assays. Pharmaceutical companies have already implemented some high-throughput toxicity screening assays in conjunction with drug development efforts. The future is likely to see more and more of these biofunctionally informative, high-throughput assays being used as initial screens or as surrogates for long-term whole animal testing. A major challenge is to find ways to overcome industry confidentiality concerns and to encourage industry to make such toxicity data available in usable formats to the larger scientific community. AI methods have the potential to play an important role in extracting generalized rules for use in toxicity prediction, possibly circumventing such data confidentiality concerns (Greene, 1996).

6 Conclusion

AI and statistical concepts lend necessary structure and objectivity to a toxicity modeling effort, whereas chemical and biological concepts help to establish the essential scientific rationale for a toxicity prediction. Practical, real-world applications of predictive toxicology efforts in industry and government, coupled with current limitations in existing knowledge pertaining to chemical toxicity mechanisms, emphasize the need for toxicity prediction efforts to pass the ''biological plausibility'' test. In addition, biological and chemical mechanisms are the language spoken by the ultimate clients of toxicity prediction models, i.e., the toxicologists and risk assessors. This reinforces the need for AI and computer modelers to move towards meaningful collaborations with chemists, biologists and toxicologists, to define, characterize, and meld the conceptually diverse elements of the toxicity prediction problem.

Acknowledgements

This paper was the subject of an invited presentation at the American Association of Artificial Intelligence 1999 Spring Symposium, Predictive Toxicology of Chemicals: Experiences and Impact of AI Tools, Stanford, CA, March 22–24 1999. I thank the organizers of that symposium, Drs. Giuseppina Gini and Emilio Benfenati, for the invitation to participate. I also thank Drs Carl Blackman and Christopher Lau for helpful comments in review of this manuscript. This manuscript has also been reviewed by the National Health and Environmental Effects Research Laboratory, U.S. Environmental Protection Agency and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

References

Arcos, JC, Argus MF and Woo Y-T (eds), 1995, *Chemical Induction of Cancer, Modulation and Combination Effects: An inventory of the many factors which influence carcinogenesis*. Boston, MA: Birkhauser.

Benfenati E and Gini G, 1997, ''Computational predictive programs (expert systems) in toxicology'' *Toxicology* **119** 213–225.

Benigni R, 1997, ''The first US National Toxicology Program exercise on the prediction of rodent carcinogenicity: definitive results'' *Mutation Research* **387**(1) 35–45.

Benigni R, Passerini L, Gallo G, Giorgi F and Cotta-Ramusino M, 1998, ''QSAR models for discriminating between mutagenic and nonmutagenic aromatic and heteroaromatic amines'' *Environ Mol Mutagen* **32**(1) 75–83.

Benigni R and Richard AM, 1998. ''Quantitative structure-based modeling applied to the characterization and prediction of chemical toxicity'' *Methods: A Companion to Methods in Enzymology* **14** 264–276.

Bristol DW, Wachsman JT and Greenwell A, 1996, ''The NIEHS Predictive-Toxicology Evaluation Project: Chemcarcinogenicity bioassays'' *Environ Health Perspect* **104S**(5) 1001–1010 (`http://www.comlab.ox.ac.uk/ou cl/groups/machlearn/PTE/`)

Combes RD and Judson P, 1995, ''The use of artificial intelligence systems for predicting toxicity'' *Pestic Science* **45** 179–194.

Darvas F, Papp A, Allerdyce A, Benfenati E, Gini G, Tichy M, Sobb N and Citti A, 1999, ''Overview of different artificial intelligence approaches combined with a deductive logic-based expert system for predicting chemical toxicity'' *1999 AAAI Spring Symposium on Predictive Toxicology of Chemicals: Experiences and Impact of AI Tools* AAAI Press, Technical Report SS-99–01, 1999.

Dearden JC, Barratt MD, Benigni R, Bristol DW, Combes RD, Cronin MTD, Judson PN, Payne MP, Richard AM, Tichy M, Worth AP and Yourick JJ, 1997, ''The development and validation of expert systems for predicting toxicity – The Report and Recommendations of ECVAM/ECB Workshop 24'' *ATLA* **25** 223–252.

Enslein K, Gombar VK and Blake BW, 1994, ''Use of SAR in computer-assisted prediction of carcinogenicity and mutagenicity of chemicals by the TOPKAT program'' *Mutation Research* **305** 47–61.

Enslein K (ed), 1999, ''Still another rodent carcinogenesis model'' *HDI Computational Toxicology NEWS* No. 22.

Gini G and Katritzky AR (eds), 1999, *Predictive toxicology of chemicals: Experiences and impact of AI tools,*

1999 AAAI Spring Symposium, AAAI Press, Technical Report SS-99–01 (`http://www.aaai.org/Press/Reports/Symposia/Spring/SS-99-01.html`)

Gold LS, Slone TH, Manley NB, Garfinkel GB, Hudes ES, Rohrbach L and Ames BN, 1991, "The Carcinogenic Potency Database: Analyses of 4000 chronic animal cancer experiments published in the general literature and by the U.S. National Cancer Institute/National Toxicology Program" *Environ Health Perspect* **96** 11–5 (`http://potency.berkeley.edu/hybrid.html`)

Greene N, 1996, "Computer software for risk assessment" *J Chem Inf Comput Sci* **37** 148–150.

Johannsen FR, 1990, "Risk assessment of carcinogenic and noncarcinogenic chemicals" *Crit Rev Toxicol* **20** (5) 341–367.

King RD and Srinivasan A, 1996, "Prediction of rodent carcinogenicity bioassays from molecular structure using inductive logic programming" *Environ Health Perspect* **104S**(5) 1031–1040.

Klopman G and Rosenkranz HS, 1994, "Approaches to SAR in carcinogenesis and mutagenesis. Prediction of carcinogenicity/mutagenicity using MULTI-CASE" *Mutation Research* **305** 33–46.

Lee Y, Buchanan, BG and Rosenkranz HS, 1996, "Carcinogenicity predictions for a group of 30 chemicals undergoing rodent cancer bioassays based on rules derived from subchronic organ toxicities" *Environ Health Perspect* **104S**(5) 1059–1063.

Lewis DF, Ioannides C and Parke DV, 1998, "A combined COMPACT and HazardExpert study of 40 chemicals for which information on mutagenicity and carcinogenicity is known, including the results of human epidemiological studies" *Hum Exp Toxicol* **17**(10) 577–586.

Liu M, Sussman N, Klopman G and Rosenkranz HS, 1996, "Structure-activity and mechanistic relationships: the effect of chemical overlap on structural overlap in data bases of varying size and composition" *Mutation Research* **372**(1) 79–85.

Richard AM, 1998a, "Structure-based methods for predicting mutagenicity and carcinogenicity: are we there yet?" *Mutation Research* **400**(1–2) 493–507.

Richard AM, 1998b, "Commercial toxicity prediction systems: A regulatory perspective" *Toxicology Letters* **102–103** 611–616.

Sanderson DM and Earnshaw CG, 1991, "Computer prediction of possible toxic action from chemical structure; the DEREK system" *Human & Exper Toxicol* **10** 261–273.

Service RF, 1998, "Microchip arrays put DNA on the spot" *Science* **282** 396–399.

Singer B and Grunberger D, 1983, *Molecular Biology of Mutagens and Carcinogens*. Plenum Press.

Talafous J, Sayre LM, Mieyal JJ and Klopman G, 1994, "META. 2. A dictionary model of mammalian xenobiotic metabolism" *J Chem Inf Comput Sci* **34**(6) 1326–1333.

Taningher M, Malacarne D, Mancuso T, Peluso M, Pescarolo MP and Parodi S, 1997, "Methods for predicting carcinogenic hazards: new opportunities coming from recent developments in molecular oncology and SAR studies" *Mutation Research* **391**(1–2) 3–32.

Trosko JE, 1997, "Challenge to the simple paradigm that 'carcinogens' are 'mutagens' and to the in vitro and in vivo assays used to test the paradigm" *Mutation Research* **373**(2) 245–249.

Wilson EO, 1998, *Consilience: The Unity of Knowledge*. New York, NY: Knopf.