

Regulatory Acceptance of (Q)SARs | Mini-Monograph

Use of QSARs in International Decision-Making Frameworks to Predict Ecologic Effects and Environmental Fate of Chemical Substances

Mark T.D. Cronin,¹ John D. Walker,² Joanna S. Jaworska,³ Michael H.I. Comber,⁴ Christopher D. Watts,⁵ and Andrew P. Worth⁶

¹School of Pharmacy and Chemistry, Liverpool John Moores University, Liverpool, England; ²Toxic Substances Control Act Interagency Testing Committee, U.S. Environmental Protection Agency, Washington, DC, USA; ³Procter & Gamble, Strombeek-Bever, Belgium; ⁴Product Stewardship & Regulatory Affairs, ExxonMobil Biomedical Sciences Inc., Machelen, Belgium; ⁵WRc-NSF, National Centre for Environmental Toxicology, Medmenham, Marlow, Bucks, England; ⁶European Chemicals Bureau, Institute for Health & Consumer Protection, Joint Research Centre, European Commission, Ispra, Italy

This article is a review of the use, by regulatory agencies and authorities, of quantitative structure–activity relationships (QSARs) to predict ecologic effects and environmental fate of chemicals. For many years, the U.S. Environmental Protection Agency has been the most prominent regulatory agency using QSARs to predict the ecologic effects and environmental fate of chemicals. However, as increasing numbers of standard QSAR methods are developed and validated to predict ecologic effects and environmental fate of chemicals, it is anticipated that more regulatory agencies and authorities will find them to be acceptable alternatives to chemical testing. *Key words:* bioaccumulation, chemical substances, persistence, QSAR, regulatory agencies, toxicity. *Environ Health Perspect* 111:1376–1390 (2003). doi:10.1289/ehp.5759 available via <http://dx.doi.org/> [Online 6 February 2003]

Introduction

Aims, Goals, and Scope of Review

The manufacture, storage, distribution, and release to the environment of xenobiotic substances worldwide are controlled and regulated at local, national, and international levels. Control and regulation of chemical substances are mandated by legislation and implemented by regulatory agencies and authorities. To regulate the use of chemicals successfully, authorities require reliable data to be produced on the effects and fate of chemicals in the environment. Traditionally, these data have been provided by testing chemical substances by a number of well-defined protocols. However, until the passage of the Toxic Substances Control Act (TSCA 1976) in the United States and similar legislation elsewhere, such data were publicly available for only a fraction of industrial chemicals. Since the enactment of TSCA and creation of the TSCA Interagency Testing Committee (ITC 2002), the ITC has recommended testing for hundreds of chemicals, the U.S. Environmental Protection Agency (U.S. EPA) has implemented these recommendations, and the producers of these chemicals have conducted more than 900 tests for these chemicals (Walker 1993a, 1993b). The development of these data, in response to the ITC's recommendations, proved advantageous for subsequent nonstatutory programs designed to encourage the voluntary development of data. For example, the Organisation for Economic Cooperation and Development's (OECD) Screening Information Data Set (SIDS) and the U.S. EPA High Production Volume (HPV) Challenge programs benefited because most of the chemicals in the early phases of those programs had already been recommended for testing by the ITC (Walker

1993a). However, even with the creation of these nonstatutory programs, there are potentially thousands of non-HPV industrial chemicals that will continue to go untested. Prioritization of these industrial chemicals for screening and testing needs the development and validation of standard methods to predict the ecologic effects and environmental fate of chemicals using quantitative structure–activity relationships (QSARs) (Russom et al. In press). Guidelines have been published for developing and using QSARs [Walker. In press (a); Walker et al. In press (a)] and for predicting ecologic effects (Bradbury et al. In press). Nonetheless, further development and validation of standard QSAR methods is needed to gain widespread acceptance by the regulatory and regulated communities.

The aim of this article is to review the worldwide regulatory use of QSARs for predicting the ecologic effects and environmental fate of chemical substances (the regulatory use of QSARs for predicting human health effects forms the basis of a second review; Cronin et al. 2003). The use of QSARs by a number of regulatory agencies to prioritize chemicals for testing and to fill data needs is described. Although QSARs are applied by agencies worldwide, this review focuses on their use in North America and by the European Union (EU) and their member states. It should be emphasized that the purpose of this article is not to provide an extensive review of the use of QSARs per se but to review their regulatory application; further details on this complex and evolving topic may be obtained from a recent review by Walker et al. (2002) and a web-based database developed by the OECD (2002). QSARs themselves have been the subject of a number of excellent recent reviews

[Boethling and Mackay 2000; ECETOC 1998; Karcher 1995; Lyman et al. 1990; Nendza 1997; Walker 2003, In press (a–e); Walker and Schultz 2002].

Regulatory Use—Europe

European Union. Risk assessment of the ecologic effects of chemicals is well described by van Leeuwen and Hermens (1995). The potential application of QSARs in this process is described in Nendza and Hermens (1995), Comber (In press), and Comber et al. (In press). In the EU, risk assessment of chemical substances is driven by the requirements of Commission Directive 93/67/EEC on Risk Assessment for New Notified Substances and Commission Regulation (EC) No. 1488/94 on Risk Assessment for Existing Substances (EEC 1993a, 1994; Joint Research Centre 1998). To ensure consistency of application of the environmental risk assessment (ERA) process, in 1996 the EU produced a comprehensive technical guidance document (TGD) to support the Directive on New Substances and the Regulation on Existing Substances (EEC 1996). This document includes a substantial chapter providing guidance on the use of QSARs in the ERA process in terms of where they should be used, how they should

This article is part of the mini-monograph "Regulatory Acceptance of (Q)SARs for Human Health and Environmental Endpoints."

Address correspondence to M. Cronin, School of Pharmacy and Chemistry, Liverpool John Moores University, Byrom Street, Liverpool, L3 3AF, England. Telephone: 44-0-151-231-2066. Fax: 44-0-151-231-2170. E-mail: m.t.cronin@livjm.ac.uk

The funding of the European Chemical Industry Council for the partial preparation of this review (to M.T.D.C. and C.D.W.) is gratefully acknowledged (Contract NMLRI-WRc-NSF0112). In addition, we thank a large number of scientists around the world who have contributed information, answered questions, and made useful suggestions to this review.

The authors, the organizations, institutes, and the businesses they represent, and the European Chemical Industry Council are not responsible for the accuracy of the information used in this review, or any omissions from it. Citations of QSARs and software in this review do not constitute recommendations or endorsements for use by any of the authors.

The authors declare they have no conflict of interest.

Received 2 May 2002; accepted 9 December 2002.

be used, and which ones should be used. Four specific uses for QSARs in ERA are identified in the EU TGD, and these are presented in more detail below.

Data Evaluation

Acceptable QSARs may be used as a supporting tool to evaluate the adequacy of the available experimental data, for example, when the validity of the test data is not obvious. This may occur when incomplete data are available on the test and/or the test differs in some way from current methods, for example, OECD test guidelines.

Decision for Further Testing/Testing Strategies

In those cases where a predicted environmental concentration (PEC)/predicted no effect concentration (PNEC) ratio, established using test data, is greater than 1, there will be a requirement to determine whether additional testing is needed to allow a refinement of the PEC/PNEC ratio. To facilitate this decision, all available test data should be reconsidered along with estimates established using acceptable QSARs. If PEC/PNEC ratios derived using the QSARs suggest that further testing is required, then generally a chronic test should be conducted on the species that showed the lowest estimated no-observable-effect concentration (NOEC).

Establishing Specific Parameters

Acceptable QSARs can be used for the estimation of specific (input) parameters used in the risk assessment, particularly in the exposure assessment when no measured data are available to enable derivation of the PEC.

Identifying Data Needs on Effects of Potential Concern

Acceptable QSARs can be used for preliminary assessment of endpoints that are not part of the base set of data and for which information is not available.

Thus QSARs are used to estimate missing physicochemical parameters in environmental exposure assessment. Rules for the use of QSARs in environmental effects assessment are also provided in the TGD. These simply state that a QSAR is considered to be acceptable for a particular use with the risk assessment process if *a*) the QSAR applied has been validated by an appropriate process, and *b*) the estimate possesses the necessary accuracy for the intended use.

At the time of our writing this article, the TGD is nearing the end of a major revision that, among other things, will incorporate a risk assessment of the marine environment into the process. This will include manual comparison of persistent, bioconcentration, and toxicity

properties against specific criteria; the use of QSARs is allowed where experimental data are not available. This is an implementation of the European Commission's interim strategy for management of persistent bioaccumulative and toxic substances (PBTs) and very persistent, very bioaccumulative substances (vPvB) (CEC 2001b), and the criteria used are shown in Table 1.

The Convention for the Protection of the Marine Environment of the Northeast Atlantic DYNAMEC prioritization of hazardous substances. The Commission for the Protection of the Marine Environment of the Northeast Atlantic (OSPAR Commission) has developed a strategy regarding hazardous substances that involved the development and application of a dynamic selection and prioritization mechanism (DYNAMEC) to produce an updated list of chemicals for priority action. The prioritization mechanism was based on persistence, toxicity, and bioaccumulation properties of chemicals and was applied to a data set of more than 180,000 substances (OSPAR Commission 2000). For many of the chemicals, no experimental test data were available, and the data used for the prioritization exercise were generated using QSARs, including the Syracuse Research Corporation (SRC, Syracuse, New York, USA) BIOWIN model for the prediction of persistence, KOWWIN and BCFWIN for bioaccumulation, and ECOSAR for toxicity.

REACH initiative. The European Commission (CEC 2001b) has recently published a white paper setting out a "strategy for a future chemicals policy," which discusses the registration, evaluation, and assessment of chemicals—the so-called REACH initiative—for new and existing chemical substances marketed in quantities of more than 1 metric ton/enterprise/year. The 30,000 existing substances affected will be processed on a phased basis over a period of 11 years (ending 2012), starting with those marketed in the highest volumes, as well as those with very high hazard (once data have been registered). There is a desire to decrease the time taken to assess the risk of thousands of existing chemicals that are in current use. An important part of this chemicals policy is the fostering of research on development and validation of testing methods as

alternatives to animal testing, including QSAR models. This theme has been taken up by the European Parliament (2001a, 2001b), which has requested the use of screening procedures based on simplified risk assessment using data modeling, for example, QSARs and use patterns, to prioritize substances of possible concern "in order to speed up risk assessments" (European Parliament 2001a). In addition, the European Commission (CEC 2001c) has recently published an "Interim Strategy for Management of PBT and vPvB Substances," which states that the identification and verification of PBTs and vPvBs among new and existing substances will use QSAR models "where experimental data do not exist." The Danish Environmental Protection Agency (Danish EPA) have proposed that their QSAR database be used for this purpose.

EU Scientific Committee on Toxicity, Ecotoxicity and the Environment. The EU Scientific Committee on Toxicity, Ecotoxicity and the Environment (CSTEE) recommended, in their general data requirements for regulatory submissions, that QSAR data may also be used. However, they cautioned that for predictions, chemicals should be assessed not only to ensure they are of similar structure, but also to ensure that they are operating by a similar mode of action as the chemicals used to elicit the QSAR. Moreover, in order to take into account the higher level of uncertainty of predicted data compared with experimental data, different rules (i.e., stricter application of safety factors and triggers) should be applied to QSAR information.

Danish self-classification of dangerous substances. The Danish EPA (2001) developed an advisory list for self-classification of dangerous substances using QSAR models. Of approximately 47,000 substances that were examined, 20,624 substances were identified as requiring classification for one or more of the following dangerous properties: acute oral toxicity, sensitization by skin contact, mutagenicity, carcinogenicity, and danger to the aquatic environment. The Danish EPA stated that "the [QSAR] models used here are now so reliable that they are able to predict whether a given substance has one or more of the properties selected with an accuracy of approximately 70–85%" (Danish EPA 2001). With regard to

Table 1. Criteria for identification of PBT and vPvB substances.

Criterion	PBT criteria	vPvB criteria
P	Half-life > 60 days in marine water or > 40 days in fresh water ^a or half-life > 180 days in marine sediment or > 120 days in freshwater sediment ^a	Half-life > 60 days in marine water or fresh water or > 180 days in marine or freshwater sediment
B	BCF > 2,000	BCF > 5,000
T	Chronic NOEC < 0.01 mg/L or CMR or endocrine-disrupting effects	Not applicable

Abbreviations: B, bioaccumulative; CMR, carcinogenic, mutagenic, reprotoxic; P, persistent; T, toxic.

^aFor marine environment risk assessment, half-life data in fresh water and freshwater sediment can be overruled by data obtained under marine conditions.

environmental effects, specific endpoints predicted included bioconcentration, biodegradation, and fish acute toxicity.

Danish QSAR database. The Danish EPA has made extensive use of QSARs and has developed a QSAR database that contains predicted data on more than 166,000 substances (OSPAR Commission 2000). A recent publication from the Danish EPA (Tyle et al. 2001) reports the use of QSARs for identification of potential PBT and vPvB substances from among the HPV and medium-production-volume chemicals used in the EU.

German use of QSARs. Lange and Vormann (1995) of the German Umweltbundesamt (UBA) have reported on the use of well-designed QSARs both to fill missing data needs and to give some assurance of the quality of the available experimental test data. They described how the German Regulatory Authority developed a QSAR-based software system to assess the validity of predictions made for data supplied with newly notified chemicals. The system predicted a wide range of fate and effect endpoints, including fish and *Daphnia* acute toxicity. However the authors explained that for 64% of the substances assessed, some of the QSARs were not applicable because the substances were ionic, contained heavy elements, hydrolyzed rapidly, or were reaction mixtures. Generally, the German UBA uses these QSAR approaches and software within the framework of the assessment of chemicals.

Use of QSARs in the Netherlands. Many of the applications of QSARs in The Netherlands are reported by van Leeuwen and Hermens (1995). An approach was developed to classify environmental pollutants according to mode (or mechanism of action) and subsequently to make predictions of toxicity (Verhaar et al. 1992, 1996, 2000). The approach allows for the calculation of a "baseline" toxicity and then a classification scheme for four classes of toxicity: class 1 (inert chemicals, baseline toxicity, typically nonpolar narcosis), class 2 (relatively inert chemicals, typically polar narcosis), class 3 (reactive chemicals, typically compounds capable of covalent electrophilic and nucleophilic interactions with biological macromolecules), class 4 (specifically acting chemicals such as pesticides). The classification approach was applied to predict the toxicity of HPV chemicals (Bol et al. 1993; Verhaar et al. 1994). For inert toxicants (i.e., nonpolar narcotics), QSARs were published for the no-observable-effect concentration values for 19 different species of bacteria, algae, fungi, protozoa, fish, amphibians, and other invertebrates (van Leeuwen et al. 1992). The information from the QSARs was used to calculate hazard concentrations for 5% of the species, with a low probability of having an effect at the ecosystem level (van Leeuwen 1995).

Regulatory Use—United States

ITC. The ITC is not a regulatory organization; however, many of the 16 U.S. government organizations represented on ITC have regulatory responsibilities. The ITC was created under section 4(e) of TSCA as an independent advisory committee to the U.S. EPA administrator (U.S. EPA 2002e). The ITC was created to identify industrial chemicals in need of testing and to add them to the priority testing list in May and November reports to the U.S. EPA administrator (Walker 1993a). The ITC has a statutory responsibility to use SARs to identify these chemicals (Walker 2003). U.S. governmental organizations represented on the ITC that apply QSARs include the U.S. EPA, Agency for Toxic Substances and Disease Registry (ATSDR), and the U.S. Food and Drug Administration (U.S. FDA) (Walker 2003).

U.S. EPA. The U.S. EPA has received about 38,000 premanufacture notifications (PMNs) for new chemicals and currently receives about 2,000 per annum. Because TSCA does not require testing before submission of a PMN, SARs and QSARs are used to predict the environmental fate and ecologic effects (Walker 2003).

To assess the risk of a new chemical, the U.S. EPA makes predictions concerning chemical identity, physical/chemical properties, environmental transport and partitioning, environmental fate, environmental toxicity, engineering releases to the environment, and environmental concentrations. The agency uses a variety of methods to make predictions that include SARs, nearest analogue analysis, chemical class analogy, mechanisms of toxicity, chemical industry survey data, and professional judgment. The agency uses these methods to fill data gaps in an assessment and to validate submitted data in notifications (Nabholz 2001). Predictions are made by the U.S. EPA Office of Pollution Prevention and Toxics (OPPT) under TSCA (Zeeman et al. 1995). The OPPT has routinely used QSARs to predict ecologic hazards, fate, and risks of new industrial chemicals, as well as to identify new chemical testing needs, for more than two decades.

OPPT (Q)SARs for physical/chemical properties used for new chemical assessments are available (U.S. EPA 2002b). Many of the QSARs applied by the U.S. EPA to make predictions are based on estimates of the logarithm of the octanol-water partition coefficient ($\log K_{ow}$). $\log K_{ow}$ is predicted by the KOWWIN software that is available within the EPIsuite software (which can be downloaded from the OPPT website; U.S. EPA 2002b). Additional sources of information regarding U.S. EPA activities are provided. For example, validation studies of OPPT SARs used in the new chemical program are available

(U.S. EPA 2002f) as well as a database of aquatic toxicity values from the literature (U.S. EPA 2002a).

Regulatory Use—Canada

The Canadian Environmental Protection Act (1999) requires that the 23,000 substances on the Canadian Domestic Substance List (DSL) (MacDonald et al. 2002) should be categorized and screened for persistence or bioconcentration and inherent toxicity. To assist with the categorization process, Environment Canada formed a technical advisory group (TAG) in December 1998. The TAG consists of scientific and technical experts from government, industry, environmental organizations, and consultant groups and was formed to act as a resource to Environment Canada for identifying and resolving scientific and technical issues that emerge from the implementation of the categorization program (Environment Canada 2002). As recommended by the TAG, an international workshop on QSARs hosted by Environment Canada was held in Philadelphia in November 1999 to establish "rules of thumb" for predicting properties and effects of the structurally diverse DSL chemicals (MacDonald et al. 2002).

Regulatory Use—Australia

Use of SARs and QSARs by Australian government organizations has been described previously (Walker et al. 2002). Australian regulatory authorities do not currently use QSARs in relation to the listing of chemicals on the Australian Inventory of Chemical Substances (<http://www.nicnas.gov.au/obligations/aics.htm>). With regard to new chemical assessments, limited use may be permitted in the National Industrial Chemicals Notification and Assessment Scheme (<http://www.nicnas.gov.au>). For example, data from chemical analogues may be used from structurally similar chemicals if toxicologic and ecotoxicologic data are unavailable on the notified chemical. From the analogue data, conclusions may be drawn on the health and environmental effects of the notified chemical. However, Australian authorities do not routinely accept SAR predictions without support from actual test data, for example, physicochemical properties and ecotoxicologic effects.

Regulatory Use—Japan

Japanese regulatory authorities do not currently make predictions of ecologic effects and environmental fate using QSARs. However, this situation is under review by the Japanese Ministry of the Environment (Walker et al. 2002).

Organizations Involved with the Validation of QSARs

Despite not being formal regulatory agencies, two bodies, the European Centre for

Validation of Alternative Methods (ECVAM) in the EU and the Interagency Coordinating Committee on the Validation of Alternative Methods in the United States, have responsibility for the validation of alternative methods to the use of animals in the safety evaluation of chemical substances. Alternative methods include *in vitro* tests as well as QSARs and other computer modeling techniques. ECVAM has evaluated the development and the potential for the validation of expert systems, including those using QSARs for predicting toxicity (Dearden et al. 1997).

The OECD has also been involved in the assessment of QSARs to predict toxicity and physicochemical properties. Workshop reports made recommendations for the use of QSARs in aquatic toxicity prediction (OECD 1992) and for physicochemical properties (OECD 1993a) and biodegradation (OECD 1993b). Many of these recommendations were incorporated into the EU TGD. The recommendations of the 1993 report for aquatic toxicity prediction have been updated recently (OECD 2000).

The OECD was also responsible for collating the results from a tripartite (United States, EU, Japan) assessment of QSARs to predict toxicity (Comber and Feijtel 1997; Feijtel 1995; Karcher et al. 1995; OECD 1994a, 1994b; U.S. EPA 1994a). This study compared ecologic effects, physical properties, and environmental fate predictions made by the U.S. EPA with the EC's minimum premarket data (MPD) for 144 chemicals. The results of the study were quite useful in judging many of the strengths and weaknesses of the U.S. EPA approach, as well as determining the utility of MPD-type data in improving the U.S. EPA predictive methods. There were some limitations to the exercise; namely, only small data sets were available, the endpoints used for comparison were limited to the tests included in the MPD data set, different approaches were used to ascertain certain parameters, and indirect measurement in some MPD data sets was used for one or more physical/chemical properties (i.e., extrapolation), which may or may not give a "true" result [for more details on this comparative study, refer to U.S. EPA (2002f)].

QSARs to Predict Acute Aquatic Toxicity

QSARs Based on Chemical Classes

Toxicologic QSARs were based originally on congeneric series of chemicals (Schultz et al. 2003). Such QSARs are based on the assumption that compounds with the same functional group (e.g., an aliphatic alcohol) have the same mode of action. It can be limited, however by the identification of the correct class for chemicals with more than one functional group and more than one mode of action. The ecotoxicity

structure activity (ECOSAR) database developed by the U.S. EPA (2002b) includes more than 150 QSARs for more than 50 chemical classes. These chemical classes are listed in Table 2.

QSARs Based on Modes of Action

QSARs based on mode of action represent an alternative to QSARs for classes of chemicals. Several modes of action have been predicted based on fish toxicity studies (Russom et al. 1997). These predicted modes of action have been incorporated into the U.S. EPA Assessment Tools for the Evaluation of Risk (ASTER) system, which links two empirical databases: the Aquatic Toxicity Information Retrieval (AQUIRE) database (Pilli et al. 1989) component of the U.S. EPA ECOTOX database (U.S. EPA 2002a), and the physical/chemical properties database (Russom et al. 1991). Further tools exist for classifying compounds according to mechanisms of action, including nonpolar and polar narcosis reactive and specific categories (Boxall et al. 1997; Verhaar et al. 1992, 1996, 2000).

Expert Systems

A number commercially available QSAR-based "expert systems" have been developed to predict different toxicity endpoints. For general reviews of expert systems, the reader is referred to Dearden et al. (1997) and Hulzebos et al. (1999).

Toxicity prediction by Komputer-Assisted Technology (TOPKAT). TOPKAT contains a suite of predictive models for both environmental and human health effects. It is marketed currently by Accelrys Inc. (Cambridge, UK). It has two models suitable for the prediction of acute toxicity. These ecotoxicity models are based on the fathead minnow and *Daphnia magna* toxicity data. The TOPKAT methodology is based on the collation of the toxicity data for large, chemically heterogeneous groups of chemicals. The chemicals may then be subdivided into any number of groups on the basis of chemical structure (e.g., aromatic, aliphatic, heterocyclic compounds), and then QSARs can be developed for each group. The QSARs are typically developed using topological, or atom-based, descriptors. An integral part of the prediction is an assessment of whether the chemical for which the prediction is made falls within the optimum prediction space (OPS) of the model (Gombar 1999). This is the "coverage" of the model as determined by the physicochemical descriptors from which it was derived. Predictions for compounds that fall outside of this range should be discarded.

The Fathead Minnow LC₅₀ module of the TOPKAT package is composed of eight statistically significant and cross-validated quantitative structure-toxicity relationship (QSTR) models, and the data from which the models

are derived. Each QSTR model assesses acute median lethal concentration (LC₅₀) to fathead minnows in weight/volume units for a specific class of chemicals. The *Daphnia magna* module of the TOPKAT package is composed of *a*) four statistically significant and cross-validated, QSTR models, and *b*) a database of 252 uniform experimental acute median effective concentration (EC₅₀) values selected after critical review of the open literature and the AQUIRE database.

Computer-automated structure evaluation (CASE) methodology. CASE methodology, and all its variants, has been developed by Klopman colleagues (Klopman 1992; Klopman and Rosenkranz 1991). There are a multitude of models available for a variety of endpoints and hardware platforms [more details are available from the manufacturer: MultiCASE Inc. (2002)]. The CASE technique is based on the development of regression-type models after the analysis of the toxicity of a large and chemically heterogeneous data set. The technique splits the molecules into all possible fragments and seeks to find those two-dimensional fragments that are associated with (either by promoting or reducing) toxicity. A regression model is then developed from those fragments normally without the use of electronic or three-dimensional structural descriptors. There are many forms of the CASE models; the software is variously called CASE, MultiCASE (MCASE), CASE-TOX, and TOXALERT [MultiCASE Inc. (2002)], depending on the endpoint being modeled and the hardware platform. For the successful prediction of acute toxicity, the baseline activity identification algorithm

Table 2. Classes of chemicals for which toxicity models exist in ECOSAR.

Acid chloride/halide	Neutral organics
Acrylamides	Peroxy acids
Acrylates	Phenols
Aldehydes	Phenols (dinitro)
Aliphatic amines	Propargyl alcohols
Anilines (amino- <i>meta</i>)	Propargyl ethers
Anilines (amino- <i>ortho</i>)	Quinone
Anilines (amino- <i>para</i>)	Salicylates
Aromatic amines	Salicylic acid
Aziridines	Schiff bases
Benzotriazoles	Silamines
Benzyl alcohols	Silanes (alkoxy)
Benzyl amines	Surfactants, anionic
Benzyl halides	Surfactants, cationic
Diazoniums	Surfactants, nonionic
Diepoxides	Thiazolidinones
Diketones	Thiazolinone (iso-)
Dinitro aromatic amine	Thiocyanates
Dinitrobenzenes	Thiols (mercaptans)
Epoxides	Thiophenes
Esters	Triazines
Esters (phosphate)	Ureas (substituted)
Haloacetamides	Vinyl/allyl alcohols
Hydrazines	Vinyl/allyl ethers
Imides	Vinyl/allyl halides
Isocyanates	Vinyl/allyl ketones
Malononitriles	Vinyl/allyl sulfones
Methacrylates	

(Klopman 1998) simulates the “baseline” effect of narcosis. In this manner, CASETOX models have been developed for the fathead minnow on the basis of 479 test values (Klopman et al. 2000) and for the guppy on the basis of 219 test values (Klopman et al. 1999).

Optimized approach based on structural indices set (OASIS). OASIS is a shell system for screening chemical inventories for physicochemical and toxic endpoints accounting for conformational flexibility of chemicals. It was developed and defined by Mekenyan et al. (1990, 1994). The OASIS Forecast software (designed for personal computers with Microsoft Windows; Laboratory of Mathematical Chemistry, University “Prof. As. Zlatarov,” Bourgas, Bulgaria) is an interfacing program providing screening of chemicals by making use of QSAR models. Presently, models are available for bioavailability, $\log K_{ow}$, acute toxicity, phototoxicity, estrogen and androgen binding affinity, and mutagenicity. The system is flexible in terms of its ability to be employed as an expert-system shell for implementing subsequently derived QSARs or QSARs derived by other laboratories. QSAR models are described as logistical decision trees that consist of multiple hierarchically ordered rules based on parameter ranges that comprise reactivity patterns associated with toxic or physicochemical endpoints. These patterns could be expressed by steric, electronic, and two-dimensional (fragment) structural requirements. Boolean logic operators are used to establish “rules” in the decision tree. If the value of a parameter calculated for a conformer of a chemical is found in a range of the molecular descriptors defined by certain confidence limits, it is assumed that a chemical meets the specific requirement for eliciting the endpoint with a certain probability. OASIS Forecast is unique in terms of accounting for conformational flexibility of chemicals during the screening. It provides two approaches for that purpose. When applied to large databases of three-dimensional structures where a chemical is represented by a single conformer, one can use the “tweak” algorithm, which is based on the directed conformational search aiming to render the distance between the fragments into the specified limits. In fact the rotatable bonds of the structures are adjusted to produce a conformation that matches as closely as possible a given three-dimensional requirement. The second alternative, to account for conformational flexibility of the screened molecules, employs the recently developed method for conformational coverage based on a genetic algorithm. The generation of sets of energetically reasonable conformers, providing coverage of conformational space, however, requires subsequent quantum chemical assessment of their electronic structures. Hence, the second alternative is evaluated to be much more time-consuming than the directed

(tweak) conformational analyses. The conformational analysis and subsequent quantum chemical calculations are performed on the fly during the screening process.

The OASIS expert system has been developed for the prediction of acute aquatic toxicity of noncongeneric chemicals using a two-step SAR approach (Karabunarliev et al. In press). It employs *a*) a computerized rule-based discrimination of chemicals and *b*) correlative QSARs using $\log K_{ow}$ and quantum-chemical descriptors. The Duluth database (Russom et al. 1997) of acute toxicities to fathead minnow for about 660 organic chemicals served as the data source to develop an explicit rule-based discrimination scheme and QSARs for the separate chemical categories. The OASIS system has been assessed for its capabilities to predict fish toxicity (Moore et al. In press).

CATABOL. CATABOL is a mechanistic modeling approach for quantitative assessment of biodegradability. The system generates the most plausible biodegradation products and provides quantitative assessment for their solubility and toxic endpoints. The core of CATABOL is the biodegradability simulator including a library of hierarchically ordered individual transformations (catabolic steps). The catabolic steps are derived from a set of the most plausible metabolic pathways predicted by experts for each chemical from the Japanese Ministry of International Trade and Industry (MITI) database training set. The data in the training set agreed well with the calculated biological oxygen demands ($r^2 = 0.90$) in the entire range; that is, a good fit was observed for readily degraded, intermediate degraded, and difficult-to-degrade chemicals. After introducing 60% theoretical oxygen demand as a cutoff value, the model predicted correctly 98% readily biodegradable structures and 96% not readily biodegradable structures. Cross-validation by leaving out 25% of the data and recalculating the model 4 times resulted in $Q^2 = 0.88$ between observed and predicted values.

CATABOL is a hybrid of a knowledge-based expert system for predicting biotransformation pathway, working in tandem with a probabilistic model that calculates probabilities of the individual transformations and overall biological oxygen demand and/or extent of CO_2 production. The novelty of the model is that the biodegradation extent is assessed on the basis of the entire pathway and not, as with all other models, on the parent structure alone. The second novelty is that CATABOL explicitly considers effects of adjacent fragments before executing each transformation step. CATABOL was described in detail by Jaworska et al. (2002).

Substructure-based computerized chemical selection expert system (SuCCSES). After years of recommending chemical classes based on SARs, the ITC determined that a computerized

system was needed to use historical information and expert opinions to facilitate the ITC’s review of large groups of chemicals with similar substructures (and modes of action, if available). Historical information and expert opinions used to identify ecologic and health effects associated with chemicals were captured in SuCCSES. Historical information was obtained from the ITC scoring exercises 1, 2, 3, 4 and 5 that were convened from 1978 to 1983 (Walker 1993a, 1993c, 1995). For each workshop the ITC invited expert toxicologists, pharmacologists, oncologists, geneticists, and ecologists to score (within their discipline) several hundred chemicals using criteria or thresholds for potential to cause acute, subchronic, mutagenic, carcinogenic, developmental, or reproductive effects to humans or acute effects to aquatic organisms, and for bioconcentration potential (Walker 1995).

Expert opinions were based on SARs, empirical data, and professional judgment. Opinions were developed by first convening separate panels of health effects experts and ecologic effects experts. These experts were asked to identify chemical substructures that were associated with chemicals likely to cause health or ecologic effects, respectively. Each panel was instructed to *a*) identify appropriate toxicologic endpoint categories for health effects or ecologic effects that could be used to classify chemicals with identical substructures and similar modes of toxic action, *b*) describe chemical substructures that could be associated with toxicologic endpoint categories for health effects or ecologic effects, *c*) develop a questionnaire that could be sent to health effects or ecologic effects experts to solicit their opinions on the reliability of the toxicologic endpoint categories and associated chemical substructures, and *d*) provide names of health effects or ecologic effects experts to whom the questionnaire then could be sent. Additional information on health effects that are coded in SuCCSES is provided in the companion article (Cronin et al. 2003).

For ecologic effects, numerous international experts were sent a questionnaire listing more than 100 different chemical substructures and asked to predict (based on their field of expertise related to ecologic effects and knowledge of modes of action) the potential for chemicals containing any of the substructures to cause effects to algae, aquatic invertebrates, birds, fish, mammals, microorganisms, plants, or terrestrial invertebrates. Opinions from these ecologic effects experts were converted to codes that identified chemical substructures and indicated potential of chemicals containing one or more substructures to cause effects to the previously listed organisms.

SuCCSES is a relational database with fields indexed on Chemical Abstract Service (CAS) Registry numbers. It was created using

Molecular Design Limited Information Systems' Integrated Scientific Information System. For chemicals in SuCCSES, chemical formulas, molecular weights, two-dimensional chemical structures, and simplified molecular input line entry system (SMILES) notations (Weininger 1988; Weininger et al. 1989) are provided. Features of SuCCSES promote substructure searching to identify structurally related chemical classes of chemicals and their potential ecologic or human health effects. For each record in SuCCSES, a computer screen displays fields for CAS Registry number, molecular weight, molecular formula, SMILES notation, two-dimensional chemical structure, chemical name, predicted mode of action, and potential health or ecologic effects. SuCCSES is not available to the public because it contains confidential business information.

SuCCSES is used to facilitate the ITC's review of large groups of chemicals (Walker 1991, 1995; Walker and Brink 1989). The aldehyde substructures in SuCCSES that were associated with potential to cause acute effects to aquatic organisms were included in recent publications (Walker and Printup In press; Walker et al. In press (b)). A forthcoming book chapter summarizes the development of SuCCSES and its applicability to the ITC's statutory mandate to use SARs before recommending chemicals for testing in May and November reports to the U.S. EPA administrator (Walker and Gray. In press).

Other approaches to toxicity prediction. Other approaches to predict toxicity include those for predicting water quality objectives (Vighi et al. 2001). This was a multivariate QSAR approach that involved the assessment of water quality objectives for a set of 125 chemicals [derived from the European priority list in compliance with Directive 76/464/EEC (EEC 1976)]. Predictions from classification models (based on algae, *Daphnia*, and fish toxicity values) were shown to perform satisfactorily compared with the classifications using literature toxicity data.

EU Regulatory Use

Technical guidance documents. In Europe, data obtained from QSARs can be used according to the guidance on the use of QSARs for specific groups of substances found in Part IV of the TGD (EEC 1996). The TGD provides recommendations for the use of QSARs to predict acute toxicity to fish (96 hr LC₅₀), *Daphnia* (48 hr EC₅₀) and algae (72- to 96-hr EC₅₀). In particular, QSARs are provided for chemicals acting by nonpolar narcosis and polar narcosis mechanisms of action. No QSARs have been recommended for substances that act by more specific modes of action.

Danish EPA. Tyle et al. (2001) reported that toxicity is considered only for substances with a bioconcentration factor (BCF) between

2,000 and 5,000. For such compounds, the QSAR predictions of the Danish EPA are used. The Danish EPA also used the MCASE fathead minnow model to predict acute aquatic toxicity of substances. The toxicity predictions were then used to assign EU risk classifications (e.g., R50, R51, and R53) (CEC 2001a)

U.K. Department of Environment, Food, and Rural Affairs (DEFRA). In the United Kingdom, DEFRA has established a Chemical Stakeholders Forum (CSF) (DEFRA 2001), which gives a voice to those in society with an interest in chemicals and their effects on the environment and, through the environment, on human health. The CSF has established criteria for identifying chemicals of concern based on specific persistence, bioaccumulation, and toxicity values. The approach recommended by the CSF implies that appropriate QSARs can be used for estimation of persistence, bioaccumulation, and toxicity properties where experimental data are lacking. The proposed approach has been endorsed by DEFRA's Advisory Committee on Hazardous Substances (ACHS 2001), which also has agreed that appropriate QSAR models can be used to fill data gaps. The Environment Agency has carried out an initial screen of the International Uniform Chemical Information Database (Heidorn et al. 2003) for substances that fulfill the CSF PBT criteria and reported the initial findings from this study to the ACHS in September 2001 (ACHS 2001).

German Umweltbundesamt (UBA). A QSAR-based software system has been developed by the UBA (Lange and Vormann 1995). Its use is intended both to fill missing data needs and to give some assurance of the quality of the available experimental test data. The system is capable of predicting a wide range of fate and effect endpoints, including fish and *Daphnia* acute toxicity.

North American Regulatory Use

U.S. EPA. The OPPT SARs for aquatic toxicity are used for new chemical assessments and are available from U.S. EPA (2002b). An environmental toxicity profile may consist of the following effects:

- Freshwater organisms (SARs and analogues): acute toxicity to fish, daphnid, and green algal species; fish, daphnid, and green algal chronic values (ChV)
- Saltwater organisms (SARs and analogues): fish, mysid shrimp, and green algal acute toxicity; fish, algal, and green algal ChV
- Benthic or sediment-dwelling organisms (analogues)
- Terrestrial plants (analogues)
- Terrestrial soil, including earthworms (SARs and analogues) and insects (analogues)
- Birds (analogues): mallards, quail, raptors
- Wild mammals (human health effects profile and analogues): marine and terrestrial

- Terrestrial insects (analogues)

The ECOSAR software is used to make predictions for these effects.

Environment Canada. Environment Canada recently funded a study to assess and evaluate six modeling packages to predict acute toxicity, with particular application to prioritizing chemicals within the Canadian DSL (Moore et al. In press). The six packages assessed were ECOSAR, TOPKAT, a probabilistic neural network (PNN) model, a computational neural network model, the QSAR components of the ASTER system, and the OASIS system. Of these, the PNN model provided the best predictions on the basis of an external test set of compounds. The TOPKAT model also provided excellent predictions, but only for compounds classified as falling within the OPS of the model.

QSARs to Predict Chronic Aquatic Toxicity

QSARs

In general few QSARs for chronic toxicity have been published. Those available are generally for unreactive, narcotic chemicals.

EU Regulatory Use

Technical guidance documents. The TGD provides recommendations for the use of QSARs to predict long-term toxicity to fish (NOEC, 28 days) and to *Daphnia* (NOEC, 21 days). In particular QSARs are provided for chemicals acting by nonpolar narcosis and polar narcosis mechanisms of action. No QSARs have been recommended for substances that act by more specific modes of action.

North American Regulatory Use

U.S. EPA. The U.S. EPA uses the chronic toxicity QSARs from the ECOSAR system.

QSARs to Predict Estrogen Receptor Binding

The issue of identifying chemicals that may elicit endocrine disruption has grown immeasurably in importance in the last decade. Interest in the prediction of endocrine disruption by QSAR methods has also been increased by the lack of historical data for this endpoint. This has made the need to screen the databases of existing chemicals a priority. The fact that endocrine disruption is an effect that could be initiated by a receptor-binding event, as well as being an immensely complex response that may be brought about by a number of interactions (e.g., interactions with estrogen, androgen hormones, etc.), has changed the challenges faced by modelers. There are thus two predictions that should be made: the first is whether a compound has the capability to elicit a response; the second is the magnitude of that response. With these challenges in

mind, modelers in the environmental field have turned to many of the techniques traditionally used in drug design and optimization. It should be noted that “endocrine disruption” covers a broad range of physiologic effects *in vivo* and may result in a number of events such as estrogenicity, androgenicity, and so forth (Zacharewski 1997). QSARs for predicting estrogen receptor binding affinity of structurally diverse chemicals have recently been reviewed (Schmieder et al. In press).

QSARs for Classes

Given that a compound has the capability for binding at a receptor site, altering the structure of the compound should alter its affinity for the binding site. This phenomenon has been well demonstrated by a number of researchers; for instance, Itz et al. (2000) demonstrated that the binding of a series of phenols to the human estrogen receptor was related to the size of the phenol. Binding to this receptor is known to require a relatively bulky, hydrophobic group. These types of models provide an accurate prediction of potency but are not applicable outside the training set.

Pattern Recognition Techniques

Common reactivity pattern (COREPA). Pattern recognition methods are statistical techniques that enable the classification and prediction of activity. Schmieder et al. (2000) reported the successful application of COREPA [implemented in Mekenyan’s OASIS software; Mekenyan et al. (1990, 1994)]. COREPA has been used to predict a chemical’s potential to bind to the estrogen receptor (Mekenyan et al. 2002) or androgen receptor (Serafimova et al. 2002).

Comparative Molecular Field Analysis

Much of the early progress in the prediction of estrogenicity was spawned as an offshoot of pharmaceutical studies. The study of Waller et al. (1996) indicated that three-dimensional techniques such as comparative molecular field analysis (CoMFA) were able to predict potency of estrogens. These techniques are limited to series of similar compounds (most commonly congeneric series), such as is often observed in classic drug optimization studies. As such, they are limited in their applicability outside of the training set. CoMFA techniques also still remain controversial regarding the success, or otherwise, of the fitting (molecular overlaying) of the compounds. Nevertheless, they have been used to predict estrogen receptor binding affinity (Hong et al. 2002; Shi et al. 2002).

Tiered Assessment Approach

The Integrated 4-Phase model was developed at the U.S. FDA’s National Center for Toxicological Research (Hong et al. 2002; Shi et al. 2002). The Integrated 4-Phase model is

composed of four sequential phases. Phase I employs two rejection filters; that is, if the molecular weight is less than 94 or greater than 1,000, and if there is no ring structure, it significantly, and with high confidence, eliminates those chemicals extremely unlikely to bind to the estrogen receptor (Shi et al. 2002). Phase II chemicals pass through phase I and are assigned as yes/no for estrogen receptor binding using 11 models using three different methods: three structural alerts, seven pharmacophore searching methods, and one classification model (Shi et al. 2002).

These 11 models are complementary and were designed to distinguish active from inactive chemicals. Only chemicals identified as inactive by all 11 models were eliminated from further evaluation in phase III. In phase III, three-dimensional chemical structure assessments and molecular alignment were used to develop a CoMFA model (Shi et al. 2001). Chemicals with higher predicted binding affinity are given higher priority for further evaluation in phase IV.

Phase IV is a knowledge-based system that can be used to foster definitive decision making and facilitate priority assessments of chemicals with high predicted estrogen relative binding affinities (Perkins et al. In press).

EU Regulatory Use

TGD. The TGD makes no recommendations for the use of QSARs for endocrine disruption.

North American Regulatory Use

U.S. EPA. The U.S. EPA established the Endocrine Disruptors Screening and Testing Advisory Committee (EDSTAC). EDSTAC was chartered in October 1996, under the Federal Advisory Committee Act (Federal Advisory Committee Act 1972). The EDSTAC developed a conceptual framework to screen and test chemicals for endocrine disruption. The Endocrine Disruption Priority Setting Database (EDPSD1) was developed by Walker et al. [In press (c)] for EDSTAC as a hierarchical database. Details of EDPSD1 and QSARs to predict estrogen receptor binding have been described elsewhere [Walker et al. In press (c)].

After the development and implementation of EDPSD1, EDPSD2 was created with substantially more resources [Walker. In press (e)]. EDPSD2 is a decision support tool that will be used to select chemicals for endocrine disruption screening assays.

QSARs for Predicting Degradation (or Persistence)

The persistence of a chemical is a vital area in establishing the ecologic effects and environmental fate of chemicals. A compound that is not persistent is generally considered to provide less risk than a persistent chemical with a

similar toxic profile. However, persistence in the environment is a complex phenomenon to model because it depends on chemical structure, environmental conditions and, for biodegradation, the ability of available microorganisms to degrade a chemical. A number of excellent reviews have been published recently in the general area of biodegradation (Raymond et al. 2001; Jaworska et al. 2002, In press). General models of persistence are also provided by Gramatica et al. (1999, 2001)

Biodegradation

QSARs for classes. The vast majority of published quantitative structure–biodegradation relationships (QSBRs) rely on octanol–water partition coefficients (K_{ow}), van der Waal’s radii, alkaline (abiotic) hydrolysis rate constants, and various molecular connectivity indices. Classes of chemicals covered by such models include chlorophenols and chloroanisoles, *n*-alkyl phthalates, alcohols, (2,4-dichlorophenoxyacetic acid) esters, *para*-substituted phenols, *meta*-substituted anilines, esters, carbamates, ethers, and ketones (Howard 2000; Howard and Banerjee 1984; Howard et al. 1987). Generally, the correlation coefficients between physical/chemical properties or molecular descriptors and biodegradation rates have been good, but overall these models have not seen much use. Their applicability was limited to a very specific class, and it was inappropriate to predict biodegradation rates for chemicals outside of that class. These models were reviewed by Howard (2000) and Raymond et al. (2001).

Generally applicable QSARs. A number of multivariate QSAR approaches, including linear and nonlinear regression, partial least squares, and neural networks, have been applied to predict biodegradability. These were reviewed recently by Raymond et al. (2001) and Jaworska et al. (2002, In press).

Expert systems. The most recently developed class of QSBRs consists of expert systems that represent artificial intelligence approaches. These so-called knowledge-based expert systems tend to act as a repository of expert knowledge about phenomena or a process that, like biodegradation, can be described by a set of rules. The library of rules or transformations is organized in a hierarchy that orders the rules by their likelihood of being executed. Because they predict (or attempt to predict) the biodegradation pathway, such models are clearly mechanistic. Expert systems are qualitative in nature, but they can be linked to other models to provide a quantitative assessment. Most recently, an expert system, CATABOL (described previously) was developed (Jaworska et al. 2002). The available expert systems to predict biodegradation have been reviewed by Jaworska et al. (2002).

Other Persistence Endpoints

Methods to predict the persistence of chemicals are well reviewed, for example, for the oxidation (Canonica and Tratnyek. In press) and reduction of chemicals in water (Tratnyek et al. In press)

Expert Systems That May Be Applied to Predictions of Persistence

SRC Software. SRC's BIODEG (Biodegradation Probability Program for Windows, BIOWIN) program. BIOWIN calculates the probability that a chemical under aerobic conditions with mixed cultures of microorganisms will biodegrade rapidly or slowly. It uses fragment constants developed using multiple linear and nonlinear regressions and data from SRC's database of evaluated biodegradation data (Howard et al. 1992). The previous version (version 3) of the program added new expert survey data. The model uses a slight revision of the previous fragments and molecular weight to *a*) calculate the probability of rapid biodegradation from the experimental data and *b*) estimate the primary and ultimate biodegradation times for complete degradation (days, weeks, months, longer) using the evaluations of 200 chemicals by 17 biodegradation experts. A description is available in Boethling et al. (1994).

SRC's BIOWIN (version 4.00). BIOWIN version 4 adds two new predictive models for assessing a chemical's biodegradability in the MITI biodegradation test. The new models use an approach similar to the linear/nonlinear regression models already included in BIOWIN. A description of the MITI biodegradation models is available (Tunkel et al. 2000). Under its Chemical Substances Control Law (CSCL 1973), the Japanese have tested approximately 900 discrete substances using the MITI test. This protocol, which determines ready biodegradability, is among six guidelines officially approved by the OECD. A total data set of 884 chemicals was compiled to derive the fragment probability values that are applied in this MITI biodegradability method. The data set consists of 385 chemicals critically evaluated as "readily degradable" and 499 chemicals critically evaluated as "not readily biodegradable."

BIOWIN produces two separate MITI probability estimates for each chemical. The first estimate is based upon the fragments derived through linear regression. The second estimate is based upon the fragments derived through nonlinear regression.

Other SRC models. SRC's atmospheric oxidation program. The Atmospheric Oxidation Program for Microsoft Windows (AOPWIN) estimates the rate constant for the atmospheric, gas-phase reaction between photochemically produced hydroxyl radicals and organic chemicals. It also estimates the rate constant for the gas-phase reaction between

ozone and olefinic/acetylenic compounds. The rate constants estimated by the program are then used to calculate atmospheric half-lives for organic compounds on the basis of average atmospheric concentrations of hydroxyl radicals and ozone. The estimation methods used by the Atmospheric Oxidation Program are based upon the SAR methods developed by Atkinson and colleagues (Atkinson 1987; Atkinson and Carter 1984; Kwok and Atkinson 1995; Kwok et al. 1996). In addition, SRC has derived some new fragment and reaction values from new experimental data. The AOPWIN program is described by Meylan and Howard (1993). AOPWIN comes with a database of experimental values for 780 chemicals for reaction rates with hydroxyl radicals, ozone, and nitrate radicals.

SRC's HydroWIN Program. The HydroWIN program calculates the hydrolysis rate constant for esters, carbamates, halomethanes, alkyl halides, and epoxides using a method developed by the U.S. EPA OPPT. It calculates a second-order acid- or base-catalyzed hydrolysis rate constant at 25°C. Acid- and base-catalyzed half-lives are calculated for pH 7 and/or pH 8. The prediction methodology was developed for the U.S. EPA and is outlined by Mill et al. (1987).

TOPKAT. The Aerobic Biodegradability Module of the TOPKAT package is a single, self-contained module, consisting of four structurally based submodels. A single study that reported the biodegradability of 894 compounds, as assessed by the Japanese MITI test protocol, was used to develop these models. The discriminant models compute the probability of a submitted structure being capable of aerobic biodegradation (probability > 0.7) or incapable of being degraded aerobically (probability < 0.3). Probability values between 0.3 and 0.7 refer to an indeterminate region in which decisions should not be made except in special circumstances or under further analytical assessments.

MultiCASE. There are two (aerobic and anaerobic) MCASE models to assist in the prediction of biodegradability. The anaerobic biodegradation model was developed based on only 79 chemicals, and its applicability is limited. The aerobic biodegradation model is combined with an expert system that predicts a biodegradation pathway. It was developed on the training set of 200 chemicals.

EU Regulatory Use

TGD. For persistence, the TGD recommends two of the SRC BIOWIN models, namely, the BIOWIN2 nonlinear model and the BIOWIN3 survey model for ultimate biodegradation. The exact cutoff points for these models have been "calibrated" on the basis of the model score for 1,2,4-trichlorobenzene, a substance that is known to be relatively

persistent under environmental conditions. In model terms the cutoff values for identifying potentially persistent substances are [BIOWIN2] < 0.5 and [BIOWIN3] < 2.2.

The European Commission (2000) reported the use of QSARs to predict persistence for compounds that were to be assessed for endocrine disruption capability.

Selection of "high-production-volume" chemicals was based on the HPV list from Regulation (EEC) No. 793/93 (EEC 1993b), of chemicals with a production volume of more than 1,000 metric tons per year. For selection of persistent chemicals the SRC Biodegradation program is used as a first indication of the persistence of a substance. Two SRC models were used: the linear regression method and the ultimate degradation method. The linear regression method leads to the definition of classes of biodegradation probability. The ultimate degradation model predicts the time for ultimate degradation (complete mineralization) of a substance. This model is based on the results of a survey of 17 biodegradation experts who were asked to evaluate 200 chemicals in terms of the time required to achieve ultimate biodegradation. The substances were rated to time units: 5 = hours; 4 = days; 3 = weeks; 2 = months; 1 = more than months. The results were averaged per substance and formulated to 36 fragments and a molecular weight parameter as with the probability estimation on linear regression. Substances that take more than months (level 1) to biodegrade, combined with a biodegradation probability of < 0.1, are considered highly persistent. Substances not fulfilling both criteria are not considered to be highly persistent.

Danish EPA. A recent publication from the Danish EPA (Tyle et al. 2001) reports the use of QSARs for identification of potential PBTs and vPvBs from among the HPV and medium-production-volume chemicals used in the EU. Proposals are made on which QSAR models should be used for persistence (the nonlinear BIOWIN model). The SRC BIOWIN program was used to assist in the advisory classification of substances according to the EU risk phrase R53 (may cause long-term effects to the aquatic environment).

OSPAR. The briefing document on the OSPAR DYNAMEC exercise (OSPAR Commission 2000) describes how two of the models incorporated in the SRC BIOWIN software were used to provide information on persistence for those substances for which experimental data were not available. The models used were those for estimating that a substance is not rapidly biodegradable in the environment and for estimating the likely time scale of ultimate biodegradation in the environment. The cutoff point was "calibrated" on the basis of experimental data for 1,2,4-trichlorobenzene because it is a substance that

is known to be rather persistent under environmental conditions.

North American Regulatory Use

U.S. EPA. The U.S. EPA has routinely applied the BIOWIN software as well as other SRC software in the assessment of persistence of substances submitted for marketing approval.

Bioaccumulation

The ability of a chemical to bioaccumulate is also an important part of an ERA. Bioaccumulation is usually assessed by measurement of the BCF. On the whole, this has been assumed to be a simple passive accumulation process driven by the thermodynamic capability to leave an aqueous environment and enter the lipid tissue of an aquatic species. Because of this, many, if not most or all, QSARs for estimating bioaccumulation have been based upon estimates of hydrophobicity.

General QSARs

A number of general models for bioaccumulation have been proposed, including linear and nonlinear approaches (Meylan et al. 1999; Dimitrov et al. 2002a, 2002b, 2003). Generally, because bioaccumulation has been considered to be a partitioning process into an organism, there has been considerable emphasis on the use of $\log K_{ow}$.

Expert Systems: SRC's BCFWIN Program

BCFWIN estimates the BCF of an organic compound using the compound's $\log K_{ow}$. The estimation methodology (developed at SRC in conjunction with the U.S. EPA) is described by Meylan et al. (1999). Measured BCF data and other key experimental details were collected for 694 chemicals. Log BCF was then regressed against $\log K_{ow}$ and chemicals with significant deviations from the line of best fit were analyzed by chemical structure. The resulting algorithm classifies a substance as either nonionic or ionic, the latter group including carboxylic acids, sulfonic acids, and their salts, and quaternary nitrogen compounds. Log BCF for nonionics is estimated from $\log K_{ow}$ and a series of correction factors if applicable; different equations apply for $\log K_{ow}$ 1.0–7.0 and > 7.0 . For ionics, chemicals are categorized by $\log K_{ow}$ and a log BCF in the range of 0.5–1.75 is assigned. Organometallics (tin and mercury), nonionics with long alkyl chains, and aromatic azo compounds receive special treatment. The correlation coefficient ($r^2 = 0.73$) and mean error (0.48) for log BCF ($n = 694$) indicate that the new method is a significantly better fit to existing data compared with other methods.

EU Regulatory Use

Hansen et al. (1999) reported the use of a $\log K_{ow}$ -based QSAR to predict bioaccumulation

factors of chemicals for use within the EU with regard to priority setting for environmental exposure of existing methods.

TGD. For bioaccumulation, BCF values may be estimated from K_{ow} using QSAR models where experimental data are not available. For highly hydrophobic substances, for example, with $\log K_{ow} > 6$, substances the available BCF models can lead to very different results. Hence, an assessment must be done on a case-by-case basis taking into account what is known about the BCF QSAR models and the specific properties of the substance, in particular, what is known to affect uptake and the potential for metabolism in aquatic organisms.

Danish EPA. The Danish EPA reported the use of SRC BCFWIN software to assist in the classification of chemical substances. A recent publication from the Danish EPA (Tyle et al. 2001) reports the use of QSARs for identification of potential PBTs and vPvBs from among the HPV and medium-production-volume chemicals used in the EU. Proposals are made on which QSAR models should be used for bioaccumulation (the SYRACUSE BCFWIN model). Because the report shows that toxicity is only relevant for substances with BCFs between 2,000 and 5,000, it is suggested that available experimental and QSAR data be examined and expert evaluation made on a case-by-case basis. Hasse diagrams were used to prioritize 50 PBTs identified by the Danish EPA based on their PBT characteristics (Carlsen and Walker 2003).

North American Regulatory Use

U.S. EPA. The U.S. EPA routinely applies the SRC BCFWIN software for risk assessment purposes of chemicals.

Soil and Sediment Sorption

Soil and sediment sorption is a further fundamental property to assess with regard to risk assessment of xenobiotic chemicals in the environment. The main approaches used predict the organic carbon partition coefficient (K_{oc} ; the ratio of the chemical adsorbed per unit weight of organic carbon in the soil or sediment to the concentration of the chemical in solution at equilibrium), which is widely used, in conjunction with the fractional organic carbon content, to model the partitioning of organics to sediments and soils.

General QSARs

As with bioaccumulation, soil sorption is a physical process that involves the partitioning of chemicals between soil and sediment, water, and the other organic fractions. It is assessed with measures such as the soil or sediment adsorption coefficient (K_{oc}). Models for K_{oc} based on $\log K_{ow}$ have normally been developed for separate chemical classes (Sabljić et al.

1995). Other general models for the prediction of K_{oc} are also available (Gramatica et al. 2000).

Expert Systems: SRC's PCKOC Program

The PCKOC program calculates the soil or sediment adsorption coefficient from a correlation to the first-order molecular connectivity indices and a series of statistically derived fragment contribution factors for polar compounds.

EU Regulatory Use: TGD

The TGD discusses a number of QSARs to predict soil and sediment sorption coefficients. In total 19 QSARs based solely on $\log K_{ow}$ are reported. These QSARs are for distinct chemical classes and were taken from the article by Sabljic et al. (1995).

North American Regulatory Use

U.S. EPA. The U.S. EPA routinely uses the PCKOC program to estimate soil sorption coefficients.

Physicochemical Properties

Physicochemical properties are important in the ERA of chemicals, both in their own right (for registration purposes for specific purposes, e.g., explosivity, etc.) and as descriptors in QSARs. There are numerous software programs to predict physicochemical properties (Lyman et al. 1982). Those described below are provided by the SRC and have been developed in collaboration with the U.S. EPA. Although these programs may not be the most accurate available, they are provided free of charge to interested parties (U.S. EPA 2002b) and are probably the most widely applied in a regulatory framework. Detailed compilations of descriptor software are discussed in individual sections above.

n-Octanol–Water Partition Coefficient (K_{ow})

K_{ow} is a physical property used extensively to describe a chemical's lipophilic or hydrophobic properties. It is the ratio of a chemical's concentration in the octanol phase to its concentration in the aqueous phase of a two-phase system at equilibrium. Because measured values range from $< 10^{-4}$ to $> 10^8$ (> 12 orders of magnitude), the logarithm ($\log K_{ow}$) is commonly used to characterize its value. $\log K_{ow}$ is a valuable parameter in numerous QSARs that have been developed for the pharmaceutical, environmental, biochemical, and toxicologic sciences.

As a consequence of the widespread use of $\log K_{ow}$ values, there are probably more models available for predicting $\log K_{ow}$ than for any other endpoint. This is of fundamental importance because $\log K_{ow}$ is essential for many of the QSARs described in this review. A number

of recent reports have assessed the capabilities of the various methods to predict $\log K_{ow}$ (cf. Buchwald and Bodor 1998; Mannhold and Dross 1996). The comparisons of methods tend to suggest that fragment-based methods [e.g., KOWWIN and ClogP (BioByte Corp., Claremont, CA, USA)] perform better than atom-based and conformation-dependent approaches.

SRC's KOWWIN program. The KOWWIN program estimates $\log K_{ow}$ of organic chemicals using an atom/fragment contribution method (Meylan and Howard 1995). The KOWWIN program has a unique feature that allows the user to adjust the estimation using an experimental K_{ow} from a structurally related chemical compound.

Water Solubility

The prediction of water solubility is closely related to prediction of $\log K_{ow}$. However, there are fewer practical approaches to water solubility prediction than for the prediction of $\log K_{ow}$. Solubility remains an important property to estimate, especially for determining cut-offs for acute toxicity and for risk assessment to aquatic species in general.

SRC's WsKow program. WsKow estimates the water solubility of an organic compound using its $\log K_{ow}$. The estimation methodology is described by Meylan et al. (1996). A total of 1,450 compounds (941 solids/509 liquids) that had measured values for K_{ow} , melting point, and water solubility were used in the development of the linear regressions used by WsKow. The correlation coefficient for the method was 0.93, with a mean error of 0.47. The method was evaluated using a separate validation set of 817 chemicals (482 solids, 335 liquids) with measured water solubilities and estimated $\log K_{ow}$ values (from SRC's LogKow program), with a resulting correlation coefficient of 0.90 and a mean error of 0.48.

Henry's Law Constant

Henry's law constant (HLC) is essentially an air-water partition coefficient. As such, it is an essential parameter for understanding and modeling the distribution of chemicals throughout the environment. Methods for the estimation of HLC have been reviewed recently by Dearden and Schüürmann (In press).

SRC's Henry program. This program calculates HLC using both the group contribution and the bond contribution methods of Hine and Mookerjee (1975). SRC has updated and expanded the bond and group contribution methods by developing new fragment constants from experimental data. The methodology is described by Meylan and Howard (1991). Henry version 3 extends the methodology to allow estimation of HLCs over a temperature range (0–50°C). In

addition, version 3 includes an experimental database of HLC values for 1,350 compounds. The Henry program also allows the user to adjust the estimation using an experimental HLC from a structurally related chemical compound.

Other Physicochemical Properties

Methods to predict melting point, boiling point, and vapor pressure are well reviewed by Dearden (In press).

SRC's MPBPVP program. This program estimates the melting point, boiling point, and vapor pressure of organic compounds. The estimation methodology for boiling point is an adaptation of the Stein and Brown (1994) method. Melting point is estimated by two different methods; the first is an adaptation of the Joback group contribution method described in detail by Dearden (1999), and the second is a method that correlates melting point to boiling point. The two melting points are then weighted according to chemical structure and magnitude of the difference between the two estimates to yield a preferred (suggested) melting point. For a description of methods to predict melting point, the reader is referred to the excellent review by Dearden (1999). Vapor pressure is estimated by three methods, all of which use the normal boiling point: the Antoine method, the modified Grain method, and the Mackay method (Neely and Blau 1985). The MPBPVP program will then select or calculate a preferred (suggested) vapor pressure (e.g., the Grain method is recommended for solids).

Regulatory Use of Estimation Methods for Physicochemical Properties

European Union. Many regulatory agencies will accept predicted values for physicochemical properties, especially $\log K_{ow}$. The use of reliable and well-validated methods such as KOWWIN are preferred over less well-known methods.

United States. The U.S. EPA and other agencies will accept predicted values for $\log K_{ow}$. In addition, the SRC software described above is used routinely by the U.S. EPA to calculate the physicochemical properties of new chemical substances for PMNs.

Prediction for Mixtures

The use of QSARs to make toxicity predictions for mixtures has been reviewed recently by Altenburger et al. (In press). Generally speaking, the making of predictions for mixtures or for different formulations is still in its infancy. QSARs are little applied to mixtures by regulatory agencies and the possible additive and/ or synergistic effects occurring from mixtures are not taken into account in risk assessment in a formal manner.

Use of (Q)SARs to Predict the Ecologic Effects and Environmental Fate of HPV Chemicals

Under the U.S. EPA HPV Chemical Challenge Program (Challenge Program) the chemical industry is being challenged to voluntarily compile a SIDS for chemicals on the U.S. HPV list (U.S. EPA 2002c). The SIDS, which has been internationally agreed to by member countries of the OECD, provides basic screening data needed for an initial assessment of the physicochemical properties, environmental fate, and human and environmental effects of chemicals. The information used to complete the SIDS can come from either existing data or from new tests conducted as part of the Challenge Program. The Challenge Program chemicals list, available online (U.S. EPA 2002d), consists of about 2,800 HPV chemicals reported under the TSCA's 1990 and 1994 (<http://www.epa.gov/opptintr/iur/>) inventory update rules. Because of the large number of chemicals on the list, it is important to reduce the number of tests to be conducted, where this is scientifically justifiable. SARs may be used to reduce testing in at least three different ways: First, by identifying a number of structurally similar chemicals as a group, or category, and allowing selected members of the group to be tested with the results applying to all other category members. Second, by applying SAR principles to a single chemical that is closely related to one or more better characterized chemicals (analogues). The analogue data are used to characterize the specific endpoint value for the HPV candidate chemical. Third, a combination of the analogue and category approaches may be used for individual chemicals, for example, searching for a "nearest chemical class" as opposed to a nearest single chemical analogue to estimate a SIDS endpoint. Such an approach is used in ECOSAR, a SAR-based computer program that generates ecotoxicity values.

The U.S. EPA has developed a guidance document to assist sponsors and others in constructing and supporting SAR arguments for potential application in the Challenge Program. The guidance will draw on experience from the OECD SIDS program, the U.S. EPA PMN program, and other sources available in the literature. OECD guidance on the use of QSARs is also available (OECD 2001).

Scope and Application of (Q)SARs in the U.S. HPV Challenge Program

The environmental fate and aquatic toxicity SARs rely heavily on physicochemical properties as inputs and are similarly structured in terms of models, chemical classes, and regression equations. However, "accepted QSARs" (cases in which ample data are available for a

given chemical class) are not available for certain chemical classes for either ecotoxicity endpoints estimated using ECOSAR or biodegradation endpoints estimated using BIOWIN.

The use of SAR/QSAR in the HPV Challenge Program is expected to decrease the number of new tests that will be required to develop a SIDS for each HPV chemical. Their use, by either the category or individual chemical approach, will necessarily be limited by the nature of the SIDS endpoint, the amount and adequacy of the existing data, and the type of SAR/QSAR analysis performed. Measured data developed using acceptable methods are preferred over estimated values. The development and use of SAR/QSAR in the Challenge Program will be different for each of the major categories of SIDS (Table 3). It has been suggested that data from HPV chemicals be used to develop and validate QSARs to make predictions of base-set data endpoints for non-HPV chemicals (Walker et al. 2003c).

Physicochemical properties. It is anticipated that melting point, boiling point, vapor pressure, K_{ow} , and water solubility data will be available for most HPVs. In some cases, this will be in the form of values taken from standard reference books [e.g., the *Merck Index* (O'Neil et al. 2001) and the *CRC Handbook of Chemistry and Physics* (Lide 2002)]. In the event that neither measured data nor reference book values are available, estimations using an appropriate model will be accepted for all physicochemical endpoints.

Environmental fate. Acceptable estimation techniques are available for photodegradation and hydrolysis, whereas biodegradation models are less available and less well accepted. The fourth SIDS endpoint in this category is a model (fugacity models to estimate transport/distribution), and so there is no measured data requirement to fulfill. Thus, estimations will be acceptable in lieu of photodegradation and hydrolysis tests but not for biodegradation.

Ecotoxicity. ECOSAR is an established QSAR program that estimates toxicity to fish, invertebrates, and algae. Even though this approach represents a screening-level characterization, it is of a higher order than either physicochemical or environmental fate tests. This is not to diminish the importance of physicochemical/environmental fate tests, but

there are layers of complexity not present in these endpoints when toxicity is the entity being measured/estimated. Therefore, some measured data must be available to strengthen the use of ECOSAR to characterize aquatic toxicity for an HPV chemical in the Challenge Program. For example, if an ECOSAR (or other aquatic toxicity SAR) estimation procedure is to be presented for any one endpoint, it must be accompanied by experimental data on that endpoint with a close analogue.

SARs for Individual Chemicals

For individual chemicals, SAR is applied in two ways: *a*) by the use of (usually quantitative) predictive models based on well-validated data sets (QSAR); or *b*) by comparing the chemical with one or more closely related chemicals, or analogues, and using the analogue data in place of testing the chemical. In the case of models, the comparison has essentially been incorporated into the model.

In developing a SAR, proposers need to consider the following steps for each HPV chemical they are interested in sponsoring:

- Step 1: Conduct literature search
- Step 2: Determine data adequacy by SIDS endpoint
- Step 3: Identify data gaps by SIDS endpoint
- Step 4: Use SAR or perform test by SIDS endpoint

SARs by SIDS Endpoints

Physicochemical estimation techniques. Methods exist for estimating most of the physicochemical properties required to develop a basic understanding of the behavior of a chemical released to the environment and its potential environmental exposure pathways. Some of the methods require input as simple as chemical structure, whereas others require much less readily available information such as water solubility values, K_{ow} , and so forth. Estimation methods for key physicochemical properties have been reviewed by Howard and Meylan (1997) and are discussed briefly below.

Boiling point, melting point, and vapor pressure. Most comprehensive estimation methods for boiling point, melting point, and vapor pressure are "group contribution" methods, where values assigned to atoms, bonds, and their placement in a molecule are used to estimate their contribution to the inherent physicochemical properties of that molecule.

The Stein and Brown (1994) method for estimating boiling points was developed and validated on a large database (> 10,000 chemicals) and has been integrated into a computer program (MPBPVP) used by OPPT. In contrast, melting points are not very well estimated by this method, so the group contribution method is combined with an algorithm that relates melting point with boiling points to estimate melting point. This method is used in MPBPVP. Recently, attempts have been made to use molecular symmetry (Simamora and Yalkowsky 1994; Krzyzaniak et al. 1995), but the methods have not been well documented or validated.

A limited number of methods are available for estimating vapor pressure. Most rely on estimating the vapor pressure from the boiling point and use melting points when the chemical is a solid at room temperature, which is the method used by OPPT in MPBPVP.

Octanol-water partition coefficient. The literature contains many methods for estimating $\log K_{ow}$. The most common are classified as "fragment constant" methods in which a structure is divided into fragments (atom or larger functional groups), and values of each group are summed together (sometimes with structural correction factors) to yield the $\log K_{ow}$ estimate (Meylan and Howard 1995; Hansch and Leo 1979, 1995; Hansch et al. 1995). The OPPT KOWWIN model is based on the fragment constant method. General estimation methods based upon molecular connectivity indices (Niemi et al. 1992), UNIFAC-derived activity coefficients (Banerjee and Howard 1988), and properties of the entire solute molecule (charge densities, molecular surface area, volume, weight, shape, and electrostatic potential) (Bodor et al. 1989; Bodor and Huang 1992; Sasaki et al. 1991) have also been developed.

Water solubility. Water solubility is a determining factor in the fate and transport of a chemical in the environment as well as the potential toxicity of a chemical. Yalkowsky and Banerjee (1992) have reviewed most of the recent literature on aqueous solubility estimation and concluded that, at present, the most practical means of estimating water solubility involves regression-derived correlations using $\log K_{ow}$. OPPT uses the $\log K_{ow}$ -based WsKow model to estimate water solubility. Recently, direct fragment constant approaches to estimating water solubility have been developed (Kuhne et al. 1995; Myrdal et al. 1995).

Environmental fate estimation techniques.
Biodegradation. Biodegradation (i.e., complete mineralization, or conversion to carbon dioxide and water) is an important environmental degradation process for organic chemicals. Prediction of biodegradability is severely limited because of the lack of reproducibility of biodegradation data (Howard et al. 1987) as well as the numerous protocols that have been

Table 3. Use of SARs in the U.S. HPV Challenge Program.

Approach	SIDS endpoint	Comment
Category	All	Assemble information on all endpoints for all category members to determine whether trends exist that would allow adequate characterization
Nearest chemical class	Ecotoxicity, degradation	Depends upon the placement of the HPV candidate chemical in an existing chemical class that is part of a QSAR
Other QSAR	Physicochemical	Estimations based on chemical bonds and where located in the candidate chemical

used for biodegradation tests (Howard and Banerjee 1984). As a result, quantitative prediction of biodegradation rates has only been attempted on very limited numbers of structurally related chemicals (Howard et al. 1992). A number of comprehensive approaches using fragment constants have been attempted to qualitatively predict biodegradability. Many of the models have used a weight-of-evidence biodegradation database (BIODEG) that was specifically developed for structure–biodegradability correlations (Howard et al. 1986). Boethling et al. (1994) used the experimental BIODEG database as well as results of an expert survey to develop four models (these models are in the OPPT program BLOWIN).

Hydrolysis rates. Hydrolysis is the reaction of a substance with water in which the water molecule or the hydroxide ion displaces an atom or group of atoms in the substance. Chemical hydrolysis at a pH normally found in the environment (i.e., pH 5–9) can be important for a variety of chemicals that have functional groups that are potentially hydrolyzable, such as alkyl halides, amides, carbamates, carboxylic acid esters and lactones, epoxides, phosphate esters, and sulfonic acid esters (Neely 1985). Only a method to predict hydrolysis rate constants for esters, carbamates, epoxides, and halogenated alkanes has been developed using linear free energy relationship (Taft and Hammett constant) methodology. A computer program (HYDROWIN) that uses this methodology is available and is used by OPPT. Also, Ellenrieder and Reinhard (1988) have developed a spreadsheet program that allows hydrolysis rates to be calculated at different pH values and temperatures if adequate data are available in the companion database.

Atmospheric oxidation rates. For most chemicals in the vapor phase in the atmosphere, reaction with photochemically generated hydroxyl radicals is the most important degradation process (Atkinson 1989). Methods for estimating reactivity with hydroxyl radicals have generally relied on fragment constant approaches or molecular orbital calculations. The method validated on the largest number of chemicals (641) is the Atkinson fragment and functional approach method (the method used in AOPWIN, the model used by OPPT), although molecular orbital methodology gives promising results on a much more limited number of chemicals (Meylan and Howard. In press).

Ecologic endpoint estimation techniques. (Q)SARs for aquatic toxicity to fish, aquatic invertebrates, and algae have been developed and used by OPPT since 1979 (U.S. EPA 1994b, 1994c). These (Q)SARs have been incorporated into a software program (ECOSAR) available free from the U.S. EPA. ECOSAR uses molecular weight and structure and log K_{ow} to predict aquatic toxicity.

The predictions are based on actual data of at least one member of a chemical class. The data (measured toxicity values) are correlated with molecular weight and log K_{ow} to derive a regression equation that may be used to predict aquatic toxicity of another chemical that belongs to the same chemical class. ECOSAR contains equations for many chemical classes (> 50; the full list is shown in Table 2) that can be categorized into the following areas of the chemical universe (Nabholz JV. Personal communication):

- Neutral organics that are nonreactive and nonionizable
- Organics that are reactive and/or ionizable and that exhibit “excess toxicity” (toxicity beyond narcosis associated with neutral organic toxicity)
- Surfactants, which are further divided by charge: nonionic, cationic, anionic, and amphoteric
- Polymers, which are further divided by charge: nonionic, cationic, anionic, and amphoteric
- Dyes, which are further divided by charge: nonionic, cationic, anionic, amphoteric
- Inorganic compounds, which are separated from organometallics; inorganics are divided into classes via the periodic table; organometallics are also divided into classes via the periodic table

ECOSAR is being constantly updated and so is only partially programmed for these chemical classes. For instance, SARs for classes with excess toxicity are added as they become available; surfactants are only partially programmed; polymer SARs are not yet programmed but will be as they become available; dyes are not yet programmed, although

neutral dyes are assessed with SARs for neutral organics or excess toxicity as their structure dictates; inorganics are not yet programmed, but many SARs have been developed and await to be programmed; organometallics are not yet programmed, but many SARs have been developed and await to be programmed: the first step in assessing organometallics will be reactivity, either pyrophoric and/or hydrolysis. Stable organometallics will be assessed on the basis of predicted log K_{ow} . To make predictions, ECOSAR now has a programmed decision tree based on SMILES, which selects the SAR that U.S. EPA experts would select. This decision tree is continually refined to reflect how OPPT/U.S. EPA performs new chemical assessments (Nabholz JV. Personal communication).

Therefore, to use ECOSAR for a particular chemical, an appropriate SAR is selected based on the following: chemical structure, chemical class, predicted log K_{ow} , molecular weight, physical state, water solubility, number of carbons, ethoxylates or both, and percentage of amine nitrogen or number of cationic charges or both, per 1,000 molecular weight. Because the regression equations are chemical specific, and because they may vary by species (fish vs. daphnids vs. algae), the most important factor is the identification of the chemical class (U.S. EPA 1994b).

The following presents some guidance on the approach for evaluating the aquatic toxicity (to fish, plants, and invertebrates) of a candidate HPV chemical using ECOSAR:

- Identify the chemical structure and convert it to SMILES notation
- Identify appropriate physicochemical properties: physical state, melting point, water

Table 4. Summary of SAR models used by the U.S. EPA for each SIDS endpoint.

SIDS category	SIDS endpoint	SAR model ^a	Model availability
Chemical and physical properties	Melting point	MPBPVP	http://esc.syrres.com/~esc/
	Boiling point		
	Vapor pressure		
	Log K_{ow}	KOWWIN	
Environmental fate and pathways	Water solubility	WsKow	
	Photodegradation	AOPWIN	
	Stability in water	HYDROWIN	
	Biodegradation	BLOWIN	
Ecotoxicity tests	Acute toxicity to fish, aquatic invertebrates, and algae	ECOSAR	www.epa.gov/opptintr/newchems

^aRequired input: CAS no. and/or SMILES notation.

Table 5. Framework of QSARs used by regulatory agencies for prediction of ecologic fate and environmental fate.^a

Organization	Aquatic toxicity	Benthic toxicity	K_{oc}	K_{ow}	MP	VP	WS	HLC	Atmospheric oxidation	BCF	Biodegradation	Hydrolysis
ATSDR				Q								
U.S. EPA	Q		Q	Q		Q	Q			Q	S	
ITC	Q	S	Q	Q	Q	Q	Q	Q	Q	Q	S	Q
Danish EPA	Q		Q	Q								
Health Canada	Q		Q	Q			Q					
EU (TGD)	Q		Q	Q								

Abbreviations: Q, QSARs; S, SARs. ^aData from Walker et al. (2002).

solubility, vapor pressure, and K_{ow} are required to predict effect concentrations (i.e., EC_{50}). If a chemical is highly water reactive (e.g., a hydrolysis half-life < 1 hr), consider estimating toxicity for the hydrolysis products

- Decide what ECOSAR chemical class best fits your chemical
- Run the ECOSAR program to develop an aquatic toxicity profile for the candidate chemical

Table 4 provides a summary of the SAR models discussed above.

Conclusions

Main Findings

The use of QSAR for regulatory purposes has been growing steadily and a framework of QSARs has been established by regulatory agencies worldwide (Table 5). By far the greatest use and application of QSARs has resulted from the TSCA and the efforts of the U.S. EPA. The regulatory use of QSARs in Europe and elsewhere in the world is less widespread and formalized. Within the EU, the use of QSARs is underpinned by the TGD and is being developed and encouraged by some of the EU environmental protection agencies. Generally, QSARs have been applied by regulatory agencies in two main ways: priority setting for existing chemicals and classification for new chemicals.

Future Outlook

The future will almost certainly bring an increased use of QSARs by regulators for estimating the ecologic effects and environmental fate of chemical substances. Such activities may include the prioritization of existing chemical databases (the examples of the Danish and Canadian authorities are pertinent in this regard). In such cases, well-designed strategic use of QSARs may be capable of flagging compounds of concern. It should also be noted that there will be an increasing requirement for the use of alternative methods, including QSARs, if the EU REACH initiative is to be successful.

The increase in the use of QSARs for the registration of new chemicals will also continue, particularly spurred on by legislation and public pressure against animal testing, and the cost of environmental assessment. However, in the foreseeable future, for the assessment of new chemicals a great reliance will continue to be placed on the use of experimental data.

Wider application and acceptance of the use of QSARs for regulatory purposes will require further development and more thorough validation and assessment of their use. There must be greater appreciation of QSAR "quality" and the appropriateness of their use, both in terms of the chemical domain described, and in terms of the precision of the estimates that are produced.

REFERENCES

- ACHS. 2001. Minutes of the Second Meeting of the Advisory Committee on Hazardous Substances held on 3 September 2001, Ashdown House, London, UK. London:Department for Environment, Food, and Rural Affairs.
- Altenburger R, Nendza M, Schüürmann G. In press. Mixture toxicity and its modeling by quantitative structure-activity relationships. *Environ Toxicol Chem*.
- Atkinson R. 1987. A structure-activity relationship for the estimation of rate constants for the gas-phase reactions of OH radicals with organic compounds. *Int J Chem Kinet* 19:799–828.
- . 1989. Kinetics and mechanisms of the gas phase reactions of the hydroxyl radical with organic compounds. *J Phys Chem Ref Data (Monogr 1)*:1–246
- Atkinson R, Carter WPL. 1984. Kinetics and mechanisms of the gas-phase reactions of ozone with organic compounds under atmospheric conditions. *Chem Rev* 84:437–470.
- Banerjee S, Howard PH. 1988. Improved estimation of solubility and partitioning through correction of UNIFAC-derived activity coefficients. *Environ Sci Technol* 22:839–841.
- Bodor N, Gabanyi Z, Wong CK. 1989. A new method for the estimation of partition coefficient. *J Am Chem Soc* 111:3783–3786.
- Bodor N, Huang MJ. 1992. An extended version of a novel method for the estimation of partition coefficients. *J Pharm Sci* 81:272–281.
- Boethling RS, Howard PH, Meylan W, Beaumann J, Tirado N. 1994. Group contribution method for predicting probability and rate of aerobic biodegradation. *Environ Sci Technol* 28:459–465.
- Boethling RS, Mackay D, eds. 2000. *Handbook on Property Estimation Methods for Chemicals*. Boca Raton, FL:Lewis Publishers.
- Bol J, Verhaar HJM, van Leeuwen CJ, Hermens JLM. 1993. Predictions of the aquatic toxicity of high production volume chemicals. The Hague:The Netherlands Ministry of Housing, Spatial Planning and Environment.
- Boxall ABA, Watts CD, Dearden JC, Bresnen GM, Scoffin R. 1997. Predicting the toxic mode of action for environmental pollutants based on physico-chemical properties. In: *Quantitative Structure-Activity Relationships in Environmental Sciences, VII* (Chen F, Schüürmann G, eds). Pensacola, FL:SETAC Press, 263–275.
- Bradbury S, Russom C, Ankley G, Schultz TW, Walker JD. In press. QSARs for predicting ecological effects of organic chemicals. *Environ Toxicol Chem*.
- Buchwald P, Bodor N. 1998. Octanol-water partition: searching for predictive models. *Curr Med Chem* 5:353–380.
- Canadian Environmental Protection Act. 1999. Chapter 33. Ottawa, Quebec, CN:Minister of Supply and Services.
- Canonica S, Tratnyek P. In press. Oxidation reactions of organic chemicals in water. *Environ Toxicol Chem*.
- Carlsen L, Walker JD. 2003. QSARs for prioritizing PBT substances to promote pollution prevention. *QSAR Comb Sci* 22:49–57.
- CEC (Commission of the European Communities). 2001a. Commission Directive 2001/59/EC of 6 August 2001 Adapting to Technical Progress for the 28th Time Council Directive 67/548/EEC on the Approximation of the Laws, Regulations and Administrative Provisions Relating to the Classification, Packaging and Labelling of Dangerous Substances. *Off J Eur Communities* L225:1–333.
- . 2001b. White Paper on the Strategy for a Future Chemicals Policy. COM(2001) 88 final. Brussels:Commission of the European Communities.
- . 2001c. Joint Meeting of the Competent Authorities for the Implementation of Council Directive 67/548/EEC (New Substances) and Council Regulation (EEC) 793/93 (Existing Substances). An Interim Strategy for Management of PBT and VPVB Substances. ENV/D/432048/01. NOTIOF/36/2001. Brussels:Commission of the European Communities.
- CSCL. 1973. *Chemical Substances Control Law of Japan*. Tokyo:Ministry of Economy, Trade and Industry, Government of Japan.
- Comber MHI. In press. Quantitative structure activity relationships in risk assessment. In: *QSARs for Pollution Prevention, Toxicity Screening and Risk Assessment* (Walker JD, ed). Pensacola, FL:SETAC Press.
- Comber MHI, Fejtel TCJ. 1997. Assessment and use of quantitative structure-activity relationships: a report from two ECE-TOC task forces. In: *Quantitative Structure-Activity Relationships in Environmental Sciences, VII* (Chen F, Schüürmann G, eds). Pensacola, FL:SETAC Press, 365–379.
- Comber MHI, Watts C, Hermens J, Walker JD. In press. QSARs for predicting potential ecological risks of organic chemicals. *Environ Toxicol Chem*.
- Cronin MTD, Jaworska JS, Walker JD, Comber MHI, Watts CD, Worth AP. 2003. Use of QSARs in international decision-making frameworks to predict health effects of chemical substances. *Environ Health Perspect* 111:000–000.
- Danish EPA. 2001. Draft Advisory List for Self-Classification of Dangerous Substances. Environmental Project No. 636. Copenhagen:Danish Environmental Protection Agency. Available: <http://www.mst.dk/chemi/01050000.htm> [accessed 1 May 2002].
- Dearden JC. 1999. The prediction of melting point. In: *Advances in Quantitative Structure Property Relationships* (Charton M, Charton BI, eds). Stamford, CT:JAI Press, 127–175.
- . In press. QSARs for prediction of boiling point, vapour pressure and melting point. *Environ Toxicol Chem*.
- Dearden JC, Barratt MD, Benigni R, Bristol DW, Combes RD, Cronin MTD, et al. 1997. The development and validation of expert systems for predicting toxicity—the report and recommendations of an ECVAM/ECB workshop (ECVAM workshop 24). *Altern Lab Anim* 25:223–252.
- Dearden JC, Schüürmann G. In press. Quantitative structure-property relationships (QSPRs) for predicting Henry's law constant from molecular structure. *Environ Toxicol Chem*.
- DEFRA. 2001. UK Chemical Stakeholders Forum. First Annual Report, 2000–2001. November, 2001, London:Department for Environment, Food, and Rural Affairs.
- Dimitrov SD, Dimitrova NC, Walker JD, Veith GD, Mekenyan OG. 2002b. Predicting BCFs of highly hydrophobic chemicals: effects of molecular size. *IUPAC J Pure App Chem* 74:1823–1830.
- Dimitrov SD, Dimitrova NC, Walker JD, Veith GD, Mekenyan OG. 2003. Bioconcentration potential predictions based on molecular attributes—an early warning system for chemicals found in humans, birds, fish and wildlife. *QSAR Comb Sci* 22:58–68.
- Dimitrov SD, Mekenyan OG, Walker JD. 2002a. Non-linear models for predicting bioconcentration and acute toxicity of narcotic chemicals. *SAR QSAR Environ Res* 13:177–184.
- ECETOC. 1998. QSARs in the Assessment of the Environmental Fate and Effects of Chemicals. Technical Rep no. 74. Brussels:European Centre for Ecotoxicology and Toxicology of Chemicals.
- Ellenrieder W, Reinhard M. 1988. Athias—an information system for abiotic transformations of halogenated hydrocarbons in aqueous solution. *Chemosphere* 17:331–344.
- Environment Canada. 2002. Existing Substances Evaluation. Ottawa, ON:Environment Canada. Available: <http://www.ec.gc.ca/substances/ese/eng/esehome.cfm#autop> [accessed 9 October 2002].
- European Commission. 2000. Towards the Establishment of a Priority List of Substances for Further Evaluation of Their Role in Endocrine Disruption. Document no. M0355008/1786Q/10/11/00. Brussels:European Commission. Available: http://europa.eu.int/comm/environment/docum/bkh_main.pdf [accessed 1 May 2002].
- EEC (European Economic Community). 1976. Council Directive 76/464/EEC of 4 May 1976 on pollution caused by certain dangerous substances discharged into the aquatic environment of the Community. *Off J Eur Communities* L129.
- . 1993a. Commission Directive 93/67/EEC of 20 July 1993. Laying down the principles for assessment of risks to man and the environment of substances notified in accordance with Council Directive 67/458/EEC. *Off J Eur Communities* L227/9.
- . 1993b. Council Regulation (EEC) No 793/93 of 23 March 1993 on the evaluation and control of the risks of existing substances. *Off J Eur Communities* L 84:1–15.
- . 1994. Commission Regulation (EC) No. 1488/94 laying down the principles for assessment of risks to man and the environment of existing substances in accordance with Council Regulation (EEC) No. 793/93. *Off J Eur Communities* L161/3.
- . 1996. Technical Guidance Document in Support of Commission Directive 93/67/EEC on Risk Assessment for New Notified Substances and Commission Regulation (EC) No 1488/94 on Risk Assessment for Existing Substances. Luxembourg:European Commission.
- European Parliament. 2001a. Compromise Amendments 1–23. Draft Report on the Commission White Paper on Strategy for a Future Chemicals Policy. PE 304.693. Strasbourg:European Parliament.
- . 2001b. Draft report by Inger Schorling on the

- Commission White Paper on Strategy for a Future Chemicals Policy. PE 304. 693/COMP. 1–23. Strasbourg:European Parliament.
- Feijt TCJ. 1995. Evaluation of the use of QSARs for priority setting and risk assessment. SAR QSAR Environ Res 3:237–245.
- Gombar VK. 1999. Reliable assessment of log P of compounds of pharmaceutical relevance. SAR QSAR Environ Res 10:371–380.
- Gramatica P, Consolaro F, Pozzi S. 2001. QSAR approach to POPs screening for atmospheric persistence. Chemosphere 43:655–664.
- Gramatica P, Consonni V, Todeschini R. 1999. QSAR study on the tropospheric degradation of organic compounds. Chemosphere 38:1371–1378.
- Gramatica P, Corradi M, Consonni V. 2000. Modelling and prediction of soil sorption coefficients of non-ionic organic pesticides by molecular descriptors. Chemosphere 41:763–777.
- Hansch C, Leo AJ. 1979. Substituent Constants for Correlation Analysis in Chemistry and Biology. New York:Wiley.
- . 1995. Exploring QSAR: Fundamentals and Applications in Chemistry and Biology. Washington, DC:American Chemical Society.
- Hansch C, Leo A, Hoekman D. 1995. Exploring QSAR: Hydrophobic, Electronic, and Steric Constants. Washington, DC:American Chemical Society.
- Hansen BG, van Haelst AG, van Leeuwen K, van der Zandt P. 1999. Priority setting for existing chemicals: European Union risk ranking method. Environ Toxicol Chem 18:772–779.
- Heidorn CJA, Rasmussen K, Hansen BG, Nrago O, Allanou R, Seynaeve R, Scheer S, Kappes D, Bernasconi R. 2003. IUCLID: an information management tool for existing chemicals and biocides. J Chem Inf Comput Sci 43:779–786.
- Hine J, Mookerjee PK. 1975. The intrinsic hydrophilic character of organic compounds. Correlations in terms of structural contributions. J Org Chem 40:292–298.
- Hong HW, Tong H, Fang L, Shi R, Xie Q, Wu J, et al. 2002. Prediction of estrogen receptor binding for 58,000 chemicals using an integrated system of a tree-based model with structural alerts. Environ Health Perspect 110:29–36.
- Howard P. 2000. Biodegradation. In: Handbook of Property Estimation Methods for Chemicals (Boethling R, Mackay D, eds). Boca Raton, FL:Lewis Publishers, 281–310.
- Howard PH, Banerjee S. 1984. Interpreting results from biodegradability tests of chemicals in water and soil. Environ Toxicol Chem 3:551–562.
- Howard PH, Boethling RS, Stiteler WM, Meylan WM, Hueber AE, Beauman JA, et al. 1992. Predictive model for aerobic biodegradability developed from a file of evaluated biodegradation data. Environ Toxicol Chem 11:593–603.
- Howard PH, Hueber AE, Boethling RS. 1987. Biodegradation data evaluation for structure/biodegradability relations. Environ Toxicol Chem 6:1–10.
- Howard PH, Hueber AE, Mulesky BC, Crisman JC, Meylan WM, Crosbie E, et al. 1986. BIODLOG, BIODEG, and fate/expos: new files on microbial degradation and toxicity as well as environmental fate/exposure of chemicals. Environ Toxicol Chem 5:977–980.
- Howard PH, Meylan WM. 1997. Prediction of physical properties, transport, and degradation for environmental fate and exposure assessments. In: Quantitative Structure-Activity Relationships in Environmental Sciences, Vol 7 (Chen F, Schüürmann G, eds). Pensacola, FL:SETAC Press, 185–205.
- Hulzebos EM, Schielen PCJ, Wijkhuizen-Maslaniewicz L. 1999. (Q)SARs for human toxicological endpoints: a literature search. RIVM Report 601516.001. Bilthoven:RIVM (Research for Man and Environment).
- ITC 2002. Toxic Substances Control Act (TSCA) Interagency Testing Committee. Washington, DC:U.S. Environmental Protection Agency. Available: <http://www.epa.gov/opptintr/itc/> [accessed 9 October 2002].
- Jaworska JS, Boethling RS, Howard PH. In press. Predicting biodegradability of organic chemicals from their structure. Review of SBRs and QSBRs. Environ Toxicol Chem.
- Jaworska J, Dimitrov S, Nikolova N, Mekenyan O. 2002. Probabilistic assessment of biodegradability based on metabolic pathways: CATABOL system. SAR QSAR Environ Res 13:307–323.
- Joint Research Centre. 1998. Technical Guidance Documents in Support of the Commission Directive 93/67/EEC on Risk Assessment for New Notified Substances and the Commission Regulation (EC) 1488/94 on Risk Assessment for Existing Substances. Ispra, Italy:Joint Research Centre, European Chemicals Bureau, European Commission, 505–566.
- Karabunarliev S, Dimitrov SD, Nikolova N, Mekenyan O. In press. Prediction of acute aquatic toxicity of non-congeneric chemicals: rule-based and quantitative structure-activity relationships. In: QSARs for Predicting Ecological Effects of Chemicals (Walker JD, ed). Pensacola, FL:SETAC Press.
- Karcher W. 1995. Structure-activity relationships in the control of chemical risks. SAR QSAR Environ Res 3:167–245.
- Karcher W, Hansen BG, van Leeuwen CJ, Wagner P, Auer C. 1995. Predictions for existing chemicals—a multilateral QSAR project. SAR QSAR Environ Res 3:217–221.
- Klopman G. 1992. MULTICASE: A hierarchical computer automated structure evaluation program. Quant Struct-Act Relat 11:176–184.
- . 1998. The MultiCASE Program II. Baseline activity identification algorithm (BAIA). J Chem Inform Comp Sci 38:78–81.
- Klopman G, Rosenkranz HS. 1991. Structure-activity relations: maximizing the usefulness of mutagenicity and carcinogenicity databases. Environ Health Perspect 96:67–75.
- Klopman G, Saikhov R, Rosenkranz HS. 2000. Multiple computer-automated structure evaluation study of aquatic toxicity II. Fathead minnow. Environ Toxicol Chem 19:441–447.
- Klopman G, Saikhov R, Rosenkranz HS, Hermens JLM. 1999. Multiple computer-automated structure evaluation program study of aquatic toxicity I: guppy. Environ Toxicol Chem 18:2497–2505.
- Krzyzaniak JF, Myrdal PB, Simamora P, Yalkowsky SH. 1995. Boiling and melting point prediction for aliphatic, non-hydrogen-bonding compounds. Ind Eng Chem Res 34:2530–2535.
- Kuhne R, Ebert R-U, Kleinf F, Schmidt G, Schüürmann G. 1995. Group contribution methods to estimate water solubility of organic chemicals. Chemosphere 30:2061–2077.
- Kwok ESC, Aschmann SM, Atkinson R. 1996. Rate constants for the gas-phase reaction of selected carbamates and lactates. Environ Sci Technol 30:329–334.
- Kwok ESC, Atkinson R. 1995. Estimation of hydroxyl radical reaction rate constants for gas phase organic compounds using a structure-reactivity relationship: an update. Atmos Environ 29:1685–1695.
- Lange AW, Vormann K. 1995. Experiences with the application of QSAR in the routine of the notification procedure. SAR QSAR Environ Res 3:171–177.
- Lide DR. 2002. CRC Handbook of Chemistry and Physics. 83rd ed. Boca Raton, FL:CRC Press.
- Lyman WJ, Reehl WF, Rosenblatt DH, eds. 1982. Handbook of Chemical Property Estimation Methods: Environmental Behavior of Organic Chemicals. New York:McGraw Hill.
- , eds. 1990. Handbook of Chemical Property Estimation Methods: Environmental Behavior of Organic Chemicals. Washington, DC:American Chemical Society.
- MacDonald D, Breton R, Sutcliffe R, Walker J. 2002. Uses and limitations of quantitative structure-activity relationships (QSARs) to categorize substances on the Canadian Domestic Substance List as persistent and/or bioaccumulative, and inherently toxic to non-human organisms. SAR QSAR Environ Res 13:43–55.
- Mannhold R, Dross K. 1996. Calculation procedures for molecular lipophilicity: a comparative study. Quant Struct-Act Relat 15:403–409.
- Mekenyan O, Kamenska V, Serafimova R, Poellinger L, Brower A, Walker J. 2002. Development and validation of an average mammalian estrogen receptor-based QSAR model. SAR QSAR Environ Res 13:579–595.
- Mekenyan OG, Karabunarliev S, Bonchev D. 1990. The microcomputer OASIS system for predicting of the biological activity of chemical compounds. Comput Chem 14:193–200.
- Mekenyan OG, Karabunarliev S, Ivanov J, Dimitrov D. 1994. A new development of the OASIS computer system for modelling of molecular properties. Comput Chem 18:173–187.
- Meylan WM, Howard PH. 1991. Bond contribution method for estimating Henry's law constants. Environ Toxicol Chem 10:1283–1293.
- . 1993. Computer estimation of the atmospheric gas-phase reaction rate of organic compounds with hydroxyl radicals and ozone. Chemosphere 26:2293–2299.
- . 1995. Atom/fragment contribution method for estimating octanol-water partition coefficients. J Pharm Sci 84:83–92.
- . In press. A QSAR review for predicting atmospheric oxidation of organic chemicals. Environ Toxicol Chem.
- Meylan WM, Howard PH, Boethling PS. 1996. Improved method for estimating water solubility from octanol/water partition coefficient. Environ Toxicol Chem 15:100–106.
- Meylan WM, Howard PH, Boethling RS, Aronson D, Printup H, Gouchie S. 1999. Improved method for estimating bioconcentration/bioaccumulation factor from octanol/water partition coefficient. Environ Toxicol Chem 18:664–672.
- Mill T, Haag W, Penwell P, Pettit T, Johnson H. 1987. Environmental Fate and Exposure Studies Development of a PC-SAR for Hydrolysis: Esters, Alkyl Halides and Epoxides. EPA Contract no. 68-02-4254. Menlo Park, CA:SRI International.
- Moore DRJ, Breton RL, MacDonald DB. In press. A comparison of model performance for six QSAR packages that predict acute toxicity to fish. Environ Toxicol Chem.
- MultiCASE Inc. 2002. Home page. Beachwood, OH:MultiCASE Inc. Available: <http://www.multicase.com/> [accessed 9 October 2002].
- Myrdal PB, Manka AM, Yalkowsky SH. 1995. AQUAFAC 3: aqueous functional group activity coefficients; application to the estimation of aqueous solubility. Chemosphere 30:1619–1637.
- Nabholz JV. 2001. Ecological hazard/risk assessment and risk management of chemicals under TSCA in USA. Invited presentation at the International Symposium on Chemical Management for the Protection of Ecosystems, 9 November 2001, Tokyo, Japan. Tokyo:Ministry of the Environment, Government of Japan.
- Neely WB. 1985. Hydrolysis. In: Environmental Exposure from Chemicals, Vol I (Neely WB, Blau GE, eds). Boca Raton, FL:CRC Press, 157–173.
- Neely WB, Blau GE, eds. 1985. Environmental Exposure from Chemicals, Vol I. Boca Raton, FL:CRC Press.
- Nendza M. 1997. Structure-Activity Relationships in Environmental Sciences. Boston:Kluwer Academic.
- Nendza M, Hermens JLM. 1995. Properties of chemicals and estimation methodologies. In: Risk Assessment of Chemicals. An Introduction (van Leeuwen CJ, Hermens JLM, eds). Dordrecht:Kluwer Academic Publishers, 239–292.
- Niemi GJ, Basak SC, Veith GD, Grunwald G. 1992. Prediction of octanol-water partition coefficient (K_{ow}) with algorithmically derived variables. Environ Toxicol Chem 11:893–900.
- O'Neil MJ, Smith A, Heckelman PE, Obenchain JR (eds). 2001. The Merck Index: An Encyclopedia of Chemicals, Drugs, and Biologicals, 13th ed. Whitehouse Station, NJ:Merck and Co., Inc.
- OECD. 1992. Report of the OECD Workshop on Application of Quantitative Structure Activity Relationships (QSARs) in Aquatic Effects Assessment. OECD Environment Monographs, no 58. Paris:Organisation for Economic Co-operation and Development.
- . 1993a. Report of the OECD Workshop on Application of SARs to the Estimation of Properties Important in Exposure Assessment. OECD Environment Monographs, no 67. Paris:Organisation for Economic Co-operation and Development.
- . 1993b. Structure-Activity Relationships for Biodegradation. OECD Environment Monographs, no 68. Paris:Organisation for Economic Co-operation and Development.
- . 1994a. U.S. EPA/EC Joint Project on the Evaluation of (Quantitative) Structure Activity Relationships. OECD Report no OECD/GD(94) 28. Paris:Organisation for Economic Co-operation and Development.
- . 1994b. U.S. EPA/EC Joint Opinion of the Scientific Committee on Toxicity, Ecotoxicity and the Environment (CSTEE) on the "Technical guidance document in support of the Directive 98/8/EC concerning the placing of biocidal products on the market: Guidance on data requirements for active substances and biocidal products. Final version—7 December 1999." Opinion expressed at the 13th CSTEE plenary meeting, 4 February 2000, Brussels, Belgium. Paris: Organisation for Economic Co-operation and Development. Available: http://europa.eu.int/comm/food/fs/sc/sct/out53_en.html [accessed 1 May 2002].
- . 2000. Use of QSAR. In: Draft Proposal for a Revised Draft Guidance Document on the Use of the Harmonized System for the Classification of Chemicals Which Are Hazardous to the Aquatic Environment. OECD Environmental Health and Safety Publications, Series on Testing and Assessment, no 24. Paris:Organisation for Economic Co-operation and Development.
- . 2001. Guidance for the Use of Structure-Activity Relationships (SARs) in the HPV Chemicals Programme. Paris: Organisation for Economic Co-operation and Development. Available: <http://www.oecd.org/pdf/M00017000/M00017236.pdf> [accessed 18 September 2001].
- . 2002. OECD's Database on Chemical Risk Assessment Models. Paris:Organisation for Economic Co-operation and Development. Available: <http://webdomino1.oecd.org/comnet/env/models.nsf> [accessed 9 October 2002].

- OSPAR Commission. 2000. Briefing Document on the Work of DYNAMEC and the DYNAMEC Mechanism for the Selection and Prioritisation of Hazardous Substances. London:OSPAR Commission.
- Perkins RG, Anson JF, Blair RM, Branham WS, Dial SL, Fang H, et al. In press. The Endocrine Disruptor Knowledge Base: a prototype toxicological knowledge base and predictive QSAR models for endocrine disrupting compounds. In: QSARs for Predicting Biological Activities Related to Endocrine Disruption Potential and Environmental-Human Health Interactions of Chemicals (Walker JD, ed). Pensacola, FL:SETAC Press.
- Pilli A, Carle DO, Sheedy BR 1989. AQUIRE: Aquatic Toxicity Information Retrieval Data Base. Technical Support Document. Duluth, MN:U.S. EPA Environmental Research Laboratory.
- Raymond JW, Rogers TN, Shonnard DR, Kline AA. 2001. A review of structure-based biodegradation estimation methods. *J Hazard Mater* B84:189–215.
- Russom CL, Anderson EB, Greenwood BE, Pilli A. 1991. ASTER—an integration of the AQUIRE data-base and the QSAR system for use in ecological risk assessments. *Sci Tot Environ* 109/110:667–670.
- Russom CL, Bradbury SP, Broderius SJ, Hammermeister DE, Drummond RA. 1997. Predicting modes of toxic action from chemical structure: acute toxicity in the fathead minnow (*Pimephales promelas*). *Environ Toxicol Chem* 16:948–967.
- Russom C, Breton R, Walker JD, Bradbury S. In press. QSARs for screening and prioritizing chemicals. *Environ Toxicol Chem*.
- Sabljic A, Gusten H, Verhaar H, Hermens J. 1995. QSAR modeling of octanol/water partition coefficients and systematics of log K-oc vs Log K-ow correlations. *Chemosphere* 31:4489–4514.
- Sasaki Y, Kubodera H, Matsuzaki T, Uneyama H. 1991. Prediction of octanol/water partition coefficients using parameters derived from molecular structures. *J Pharmacobio Dyn* 14:207–214.
- Schmieder P, Ankley G, Mekenyan O, Walker JD, Bradbury S. In press. QSARs for predicting estrogen receptor binding affinity of structurally diverse chemicals. *Environ Toxicol Chem*.
- Schmieder PK, Aptula AO, Routledge EJ, Sumpter JP, Mekenyan OG. 2000. Estrogenicity of alkylphenolic compounds: a 3-D structure-activity evaluation of gene activation. *Environ Toxicol Chem* 19:1727–1740.
- Schultz TW, Cronin MTD, Walker JD, Aptula A. 2003. Quantitative structure-activity relationships (QSARs) in toxicology: a historical perspective. *J Mol Struct Thechem* 622:1–22.
- Schultz TW, Sinks GD, Cronin MTD. 2000. Effect of substituent size and dimensionality on potency of phenolic xenoestrogens evaluated with a recombinant yeast assay. *Environ Toxicol Chem* 19:2637–2642.
- Serafimova R, Walker J, Mekenyan O. 2002. Androgen receptor binding affinity of pesticide “active” formulation ingredients. QSAR evaluation by COREPA method. *SAR QSAR Environ Res* 13:127–134.
- Shi LM, Fang H, Tong W, Wu J, Perkins R et al. 2001. QSAR models using a large diverse set of estrogens. *J Chem Inf Comput Sci* 41:186–195.
- Shi L, Tong W, Fang H, Xie Q, Hong H, Perkins R, et al. 2002. An integrated “4-phase” approach for setting endocrine disruption screening priorities—phase I and II predictions of estrogen receptor binding affinities. *SAR QSAR Environ Res* 13:69–88.
- Simamora P, Yalkowsky SH. 1994. Group contribution methods for predicting the melting points and boiling points of aromatic compounds. *Ind Eng Chem Res* 33:1405–1409.
- Stein SE, Brown RL. 1994. Estimation of normal boiling points from group contribution. *J Chem Inf Comput Sci* 34:581–587.
- Tratnyek P, Weber E, Schwarzenbach R. In press. Chemical reductions of organic contaminants. *Environ Toxicol Chem*.
- TSCA. 1976. Toxic Substances Control Act of 1976. Public Law 94–469.
- Tunkel J, Howard PH, Boethling RS, Stiteler W, Loonen H. 2000. Predicting ready biodegradability in the MITI test. *Environ Toxicol Chem* 19:2478–2485.
- Tyle H, Larsen HS, Wedebye EB, Niemela J. 2001. Identification of potential PBTs and vPvBs by use of QSARs. Copenhagen: Danish Environmental Protection Agency.
- U.S. EPA. 1994a. U.S. EPA/EC Joint Project on the Evaluation of (Quantitative) Structure Activity Relationships. EPA 743-R-94-001. Washington, DC:Office of Pollution Prevention and Toxics, U.S. Environmental Protection Agency.
- . 1994b. Estimating Toxicity of Industrial Chemicals to Aquatic Organisms Using SAR, 2nd ed. EPA 748-R-93-001. Washington, DC:Office of Pollution Prevention and Toxics, U.S. Environmental Protection Agency.
- . 1994c. ECOSAR: A Computer Program for Estimating the Ecotoxicity of Industrial Chemicals. EPA-748-R-93-002. Washington, DC:National Center for Environmental Publication and Information.
- . 2002a. ECOTOX Database. Washington, DC:U.S. Environmental Protection Agency. Available: <http://www.epa.gov/ecotox/> [accessed 9 October 2002].
- . 2002b. Exposure Assessment Tools and Models. Washington, DC:U.S. Environmental Protection Agency. Available: <http://www.epa.gov/oppt/exposure/docs/episutedl.htm> [accessed 9 October].
- . 2002c. High Production Volume (HPV) Challenge Program. Washington, DC:U.S. Environmental Protection Agency. Available: <http://www.epa.gov/chemrtk/sarfin11.htm> [accessed 9 October 2002].
- . 2002d. High Production Volume (HPV) Challenge Program. Washington, DC:U.S. Environmental Protection Agency. Available: <http://www.epa.gov/chemrtk/volchall.htm> [accessed 9 October 2002].
- . 2002e. Minimum Pre-market Data/Structure Activity Relationships Study. Washington, DC:U.S. Environmental Protection Agency. Available: <http://www.epa.gov/oppt-intr/MPD-SAR/> [accessed 9 October 2002].
- . 2002f. New Chemicals Program. Washington DC:U.S. Environmental Protection Agency. Available: <http://www.epa.gov/oppt/newchemicals/> [accessed 9 October 2002].
- van Leeuwen CJ. 1995. Ecotoxicological effects. In: Risk Assessment of Chemicals. An Introduction (van Leeuwen CJ, Hermens JLM, eds). Dordrecht:Kluwer Academic Publishers, 175–237.
- van Leeuwen CJ, Hermens JLM, eds. 1995. Risk Assessment of Chemicals. An Introduction. Dordrecht:Kluwer Academic Publishers.
- van Leeuwen CJ, van der Zandt PTJ, Aldenberg T, Verhaar HJM, Hermens JLM. 1992. Application of QSARs, extrapolation and equilibrium partitioning in aquatic effects assessment. I. Narcotic industrial pollutants. *Environ Toxicol Chem* 11:267–282.
- Verhaar HJM, Ramos EU, Hermens JLM. 1996. Classifying environmental pollutants. 2. separation of class 1 (baseline toxicity) and class 2 (“polar narcosis”) type compounds based on chemical descriptors. *J Chemomet* 10:149–162.
- Verhaar HJM, Solbe J, Speksnijder J, Van Leeuwen CJ, Hermens JLM. 2000. Classifying environmental pollutants. 3. External validation of the classification system. *Chemosphere* 40:875–883.
- Verhaar HJM, van Leeuwen CJ, Bol J, Hermens JLM. 1994. Application of QSARs in risk management of existing chemicals. *SAR QSAR Environ Res* 2:39–58.
- Verhaar HJM, van Leeuwen CJ, Hermens JLM. 1992. Classifying environmental pollutants. 1. Structure-activity relationships for prediction of aquatic toxicity. *Chemosphere* 25:471–491.
- Vighi M, Gramatica P, Consolero F, Todeschini R. 2001. QSAR and chemometric approaches for setting water quality objectives for dangerous chemicals. *Ecotox Environ Saf* 49:206–220.
- Walker JD. 1991. Chemical selection by the TSCA Interagency Testing Committee: use of computerized substructure searching to identify chemical groups for health effects, chemical fate and ecological effects testing. *Sci Total Environ* 109/110:691–700.
- . 1993a. The TSCA Interagency Testing Committee, 1977 to 1992: creation, structure, functions and contributions. In: Environmental Toxicology and Risk Assessment, Vol 2 (Gorsuch JW, Dwyer FJ, Ingersoll CG, La Pointe TW, eds). ASTM STP 1216. Philadelphia:American Society for Testing and Materials, 451–509.
- . 1993b. The TSCA Interagency Testing Committee’s role in facilitating development of test methods: toxicity and bio-concentration testing of chemicals added to sediment. In: Environmental Toxicology and Risk Assessment, Vol 2 (Gorsuch JW, Dwyer FJ, Ingersoll CG, La Pointe TW, eds). ASTM STP 1216. Philadelphia:American Society for Testing and Materials, 688–722.
- . 1993c. The TSCA Interagency Testing Committee’s approaches to screening and scoring chemicals and chemical groups: 1977–1983. In: Access and Use of Information Resources in Assessing Health Risks from Chemical Exposure (Lu PY, ed). Oak Ridge, TN:Oak Ridge National Laboratories, 77–93.
- . 1995. Estimation methods used by the TSCA Interagency Testing Committee to prioritize chemicals for testing: exposure and biological effects scoring and structure activity relationships. *Toxicol Modeling* 1:123–141.
- Walker JD. 2003. Applications of QSARs in toxicology: a US Government perspective. *J Mol Struct Thechem* 622: 167–184.
- . ed. In press (a). QSARs for Pollution Prevention, Toxicity Screening and Risk Assessment. Pensacola, FL:SETAC Press.
- . ed. In press (b). QSARs for Predicting Biological Activities Related to Endocrine Disruption and Environmental-Human Health Interactions. Pensacola, FL:SETAC Press.
- . ed. In press (c). QSARs for Predicting Ecological Effects. Pensacola, FL:SETAC Press.
- . ed. In press (d). QSARs for Predicting Physical Properties, Bioconcentration Potential and Environmental Fate. Pensacola, FL:SETAC Press.
- . In press (e). QSARs for risk assessment—current practices. In: QSARs for Pollution Prevention, Toxicity Screening and Risk Assessment (Walker JD, ed). Pensacola, FL:SETAC Press.
- Walker JD, Brink RH. 1989. New cost-effective, computerized approaches to selecting chemicals for priority testing consideration. In: Aquatic Toxicology and Environmental Fate, Vol 11 (Suter GW II, Lewis MA, eds). ASTM STP 1007. Philadelphia:American Society for Testing and Materials, 507–536.
- Walker JD, Gray DA. In press. The Substructure-Based Computerized Chemical Selection Expert System (SuCCSES). In: QSARs for Predicting Biological Activities Related to Endocrine Disruption Potential and Environmental-Human Health Interactions of Chemicals (Walker JD, ed), Pensacola, FL:SETAC Press.
- Walker JD, Printup H. In press. Structure activity relationships for categories of aldehydes. In: QSARs for Predicting Ecological Effects of Chemicals (Walker JD, ed). Pensacola, FL:SETAC Press.
- Walker JD, Schultz TW. 2002. Structure activity relationships for predicting ecological effects of chemicals. In: Handbook of Ecotoxicology, 2nd ed (Hoffman DJ, Rattner BA, Burton GA, Cairns J, eds). Boca Raton, FL:CRC Press, 893–910.
- Walker JD, Carlsen L, Hulzebos E, Simon-Hettich B. 2002. Global government applications of analogues, SARs and QSARs to predict aquatic toxicity, chemical or physical properties, environmental fate parameters and health effects of organic chemicals. *SAR QSAR Environ Res* 13:607–616.
- Walker JD, Dimitrov S, Mekenyan O. 2003. Using HPV chemical data to develop QSARs for non-HPV chemicals: opportunities to promote more efficient use of chemical testing resources. *QSAR Comb Sci* 22: 386–395.
- Walker JD, Jaworska J, Comber MHI, Schultz TW, Dearden JC. In press (a). Guidelines for developing and using quantitative structure activity relationships. *Environ Toxicol Chem*.
- Walker JD, Koleva V, Dimitrov SD, Mekenyan OG. In press (b). Quantitative structure activity relationships for predicting acute toxicity of aldehydes to aquatic organisms. In: QSARs for Predicting Ecological Effects of Chemicals (Walker JD, ed). Pensacola, FL:SETAC Press.
- Walker JD, Waller C, Kane S. In press (c). The endocrine disruption priority setting database (EDPSD): a tool using QSARs to rapidly sort and prioritize chemicals for endocrine disruption screening and testing. In: QSARs for Predicting Biological Activities Related to Endocrine Disruption Potential and Environmental-Human Health Interactions of Chemicals (Walker JD, ed). Pensacola, FL:SETAC Press.
- Waller CL, Oprea TI, Chae K, Park HK, Korach KS, Laws SC, et al. 1996. Ligand-based identification of environmental estrogens. *Chem Res Toxicol* 9:1240–1248.
- Weininger D. 1988. A chemical language and information system. 1. Introduction to methodology and encoding rules. *J Chem Inf Comp Sci* 28:31–36.
- Weininger D, Weininger A, Weininger JL. 1989. SMILES. 2. Algorithm for generation of unique SMILES notation. *J Chem Inf Comput Sci* 29:97–101.
- Yalkowsky SH, Banerjee S. 1992. Aqueous Solubility Methods of Estimation for Organic Compounds. New York:Marcel Dekker.
- Zacharewski T. 1997. In vitro bioassays for assessing estrogenic substances. *Environ Sci Technol* 31:613–623.
- Zeeman M, Auer CM, Clements RG, Nabholz JV, Boethling RS. 1995. U.S. EPA regulatory perspectives on the use of QSAR for new and existing chemical evaluations. *SAR QSAR Environ Res* 3:179–201.