Assessment of the Oral Rat Chronic Lowest Observed Adverse Effect Level Model in TOPKAT, a QSAR Software Package for Toxicity Prediction

R. Venkatapathy,*[,†,§] C. J. Moudgal,[‡,§] and R. M. Bruce[‡,§]

Oak Ridge Institute for Science and Education, National Center for Environmental Assessment, Office of Research and Development, U.S. Environmental Protection Agency (NCEA-USEPA), 26 West Martin Luther King Drive, Cincinnati, Ohio 45268

Received March 22, 2004

The performance of the rat chronic lowest observed adverse effect level (LOAEL, the lowest exposure level at which there are biologically significant increases in the severity of adverse effects) model in Toxicity Prediction by Komputer Assisted Technology (TOPKAT), a commercial quantitative structure−activity relationship software package, was tested on a database of chemicals that are of interest to the U.S. EPA's Office of Pesticide Programs. The testing was repeated on a database of chemicals from three U.S. EPA sources that report peer-reviewed LOAELs. The results of this study were also contrasted with the results of the testing performed during TOPKAT's model-building process.

INTRODUCTION

Decision criteria for remediating chemicals at hazardous waste sites are based on an assessment of their potential toxicity to human health and the surrounding ecosystem. In general, a given chemical is considered to be toxic and a decision is made to remediate the waste site if exposure to the chemical is greater than health benchmarks such as oral reference doses (RfDs), inhalation reference concentrations (RfCs), carcinogenicity oral slope factors (OSFs), or carcinogenicity inhalation risk units (URIs).[1] Often, there is a lack of experimental toxicity data to derive the benchmarks for a majority of the chemicals at waste sites. Reliance on the traditional short- and long-term animal bioassays for a new chemical would provide the necessary toxicity data but has many practical limitations. For example, generating new data would require experts in toxicology to evaluate the results from the bioassays; there are ethical and social issues involved with animal testing, and the assays themselves are expensive and time-consuming.[2,3] In addition, as hundreds of new chemicals are being synthesized every year, innovative approaches need to be developed to determine their health effects in addition to those chemicals that currently lack toxicity data.

Predictive toxicological approaches such as quantitative structure−activity relationships (QSARs) provide a means to estimate the carcinogenic and noncarcinogenic toxicity of chemicals in the absence of experimental toxicity data. QSARs describe correlations between various descriptors of molecular structures of chemicals and their observed or predicted biological activities. QSARs assume a common mechanism behind the biological activity of a set of chemicals. Hence, differences in the chemical structures that induce the same biological effect and their associated descriptors can then be mapped to changes in activity through mathematical equations. The resulting equation can then be used to calculate the biological activities of new chemicals.

Commercial QSAR software packages such as DEREK (LHASA Ltd., Leeds, U.K.), CASE and MULTICASE (MultiCase Inc., Cleveland, OH), COMPACT,[5,6] HazardExpert (Compudrug Ltd., Budapest, Hungary), ONCOLOGIC (Logichem Inc., Boyertown, PA), and Toxicity Prediction by Komputer Assisted Technology (TOPKAT; Accelrys Inc., Birmingham, MA) are able to predict the toxicity potentials of a wide variety of chemicals.[7−10] However, currently, only TOPKAT provides models for both qualitative and quantitative toxicity predictions. In addition, only DEREK, MULTICASE, and HazardExpert provide the ability for the users to add new toxicity information to their database of chemicals that was used to generate the QSAR models. The other software packages mentioned above do not offer this option and hence may lose the ability to predict the toxicity potentials of new chemicals with time, especially those that contain new and exotic functional groups or combinations of functional groups that were not included in the original model data set.

The National Center for Environmental Assessment (NCEA), U.S. Environmental Protection Agency (EPA), Cincinnati, OH, uses TOPKAT to screen/rank chemicals on the basis of their hazard potential and select surrogates with peer-reviewed toxicity values for chemicals that lack experimental toxicity data.[11] Among TOPKAT's 29 toxicity modules, the oral rat chronic lowest observed adverse effect level (LOAEL), oral rat lethal dose (LD50), rodent carcinogenicity, Ames mutagenicity, and developmental toxicity models are used with great frequency at NCEA. TOPKAT was chosen over other commercial models because (a) only TOPKAT provided quantitative LOAEL models, (b) TOPKAT had a built-in quality control module to verify if a chemical's descriptors were within the model's prediction domain, and (c) TOPKAT had a similarity search feature that was capable of determining if a query chemical was

* Corresponding author phone: (513) 569-7077; fax: (513) 569 7475; e-mail: venkatapathy.raghuraman@epa.gov.

† Oak Ridge Institute for Science and Education.

‡ National Center for Environmental Assessment, Office of Research and Development.

§ Present address: U.S. Environmental Protection Agency (NCEA-USEPA), 26 West Martin Luther King Drive, Cincinnati, OH 45268.

10.1021/ci049903s CCC: $27.50
Published on Web 07/14/2004

Table 1. Accuracy of the Five Submodels for the Prediction of LOAEL by TOPKAT[21−23]

chemical subclass	no. of chemicals[a]	no. of variables	adj r^2	SD	SE	% chemicals predicted within a factor of 2	3	4	5	95% chemicals predicted within a factor of
acyclics	73	17	0.87	0.85	0.31	73	92	97	100	4
alicyclics	39	12	0.98	1.49	0.22	94	100			3
heteroaromatics	68	17	0.85	0.80	0.30	78	92	98	100	4
multiple benzenes	83	14	0.78	0.71	0.34	70	92	96	97	4
single benzenes	130	23	0.79	0.75	0.34	66	88	94	98	5

[a] Total = 393 chemicals.

represented in TOPKAT's model database. The latter two TOPKAT-specific features increase confidence in toxicity predictions for chemicals that lack experimental data.

The main objective of this research was to determine if the rat chronic LOAEL model in TOPKAT could successfully predict the noncancer toxicities of a wide variety of chemicals. A second, but minor, objective of this research was to determine if the rat chronic LOAEL model in TOPKAT needs to be updated to include the toxicity data of chemicals generated since the model's initial development in 1991. To achieve the objectives, rat chronic LOAELs of 343 chemicals that were of interest to the U.S. EPA's Office of Pesticide Programs (OPP) were predicted using TOPKAT's LOAEL model and compared to their respective experimental values (hereafter referred to as the OPP database). The analysis was repeated for 313 chemicals from IRIS,[12] HEAST,[13] and provisional toxicity value[14] (PTV) databases (hereafter referred to as the IHP database), which reported peer-reviewed LOAEL values in oral rat chronic and subchronic studies.

BACKGROUND ON TOPKAT

TOPKAT, originally developed by Health Designs, Inc. (Rochester, NY) and currently marketed by Accelrys Inc. (Burlington, MA), is a PC-based modular software for the prediction of a wide variety of health end points. The qualitative models in TOPKAT provide dichotomous output (yes/no) for rodent carcinogenicity, Ames mutagenicity, developmental toxicity, skin sensitization, skin irritancy, ocular irritation, and aerobic biodegradability.[15] The quantitative models provide point estimates for LOAEL, LD50, lethal concentration (LC50), maximum tolerated dose (MTD), and octanol−water partition coefficient ($V \log P$) along with 95% confidence limits for each.[15]

Each TOPKAT module consists of a specific database of carefully screened chemicals, and several chemical subclass-specific cross-validated QSAR models for predicting a specific toxicity end point. To assess the end-point specific toxicity for any given chemical structure, the software employs appropriate bulk, electronic, and transport attributes that are responsible for the biological activity of the molecule.[3,4]

Electronic attributes were explained using *E*-state values,[16] which encode information on the topology, interaction(s) between atoms or groups of atoms, and the presence of valence, π, σ, and lone-pair electrons in a molecule.[3,17] Bulk attributes were explained using molecular weight and size-corrected *E*-state values for one- and two-atom fragments in the molecule.[16] Transport attributes were taken into account by using the counts of one- and two-atom constituents of the octanol−water partition coefficient,[18] topological shape indices[19] and molecular symmetry indices.[20]

TOPKAT LOAEL Model Database. The LOAEL model in TOPKAT was initially developed in 1991 using oral rat chronic bioassay data of 234 chemicals from the EPA and National Cancer Institute/National Toxicology Program (NCI/NTP) databases, with more weight placed on data from the EPA database when experimental toxicity data were available in both databases. The EPA data consisted of peer-reviewed LOAEL (in units of milligrams of chemical per kilograms of body weight per day (mg/(kg·day))) values, while the NCI/NTP values were extracted from text tables using the lowest dose at which an adverse effect was noted. In all, oral rat chronic LOAEL values for 247 chemicals were available from the two databases.[20] However, during the model construction process, 13 chemicals were rejected because they were outliers, they were influential, or they did not have at least three nonzero descriptor values.

The final version of the LOAEL model was redeveloped in 1995 with 393 chemicals (hereafter referred to as TOPKAT's database) after augmenting the 1991 dataset containing 247 chemicals with additional data from the National Toxicology Program Technical Reports, Food and Drug Administration New Drug Applications, and citations from the open literature. In this version, the 393 chemicals were divided into 5 chemical subclasses: acyclics, alicyclics, heteroaromatics, single benzenes, and multiple benzenes, with each subclass containing 73, 39, 68, 83, and 130 chemicals, respectively.

LOAEL Model Accuracy. Mumtaz et al.[20] assessed the performance of the initial oral rat chronic LOAEL model in TOPKAT by using all 234 chemicals from TOPKAT's model database. Mumtaz et al. indicated that 55% of the 234 chemicals had estimated LOAELs within a factor of 2 of the observed LOAELs, while over 93% of the chemicals were within a factor of 5. In addition, all chemicals were within a factor of 10 of the observed LOAELs. Table 1 lists the summary statistics for the LOAEL model developed with the augmented database of 393 chemicals.[21−23] According to the study results, the alicyclic submodel performed the best in terms of the percentage predicted (100% of the chemicals predicted within a factor of 3), while the single benzene submodel performed the worst (88% of chemicals predicted within a factor of 3).

METHODS

The LOAELs of all chemicals considered in this study were evaluated using the oral rat chronic LOAEL model in TOPKAT 6.1. All statistical analyses in this study were performed using Microsoft Excel 2002 (Microsoft Corp., Redmond, WA).

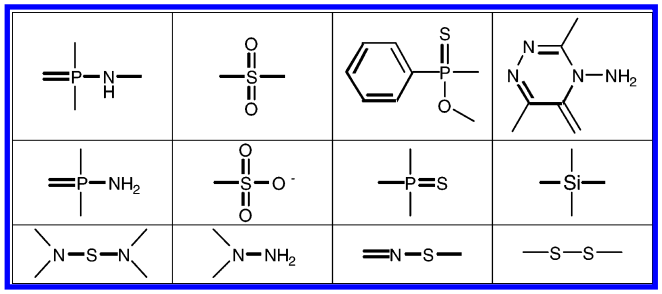

Figure 1. Fragments not represented in TOPKAT's LAOEL model database for the chemicals considered in this study.

TOPKAT Predictions. LOAEL values of the chemicals were estimated by first entering the chemical structure as a simplified molecular input line entry system (SMILES[24]) code in TOPKAT and the rat oral chronic LOAEL model selected as the prediction module. As recommended by TOPKAT, the SMILES for nitro groups was entered as N(=O)(=O) instead of [N+](=O)([O-])(=O), while salts were entered as the appropriate acid or base. The software first screens the chemical structure against the LOAEL model substructural library, determines whether the molecular structure is adequately covered by the model, and automatically chooses the appropriate chemical class-specific QSAR submodel to generate the toxicity prediction.

A successful quantitative LOAEL prediction is made when the results of the substructure analysis satisfy the validation criteria for both the univariate and multivariate procedures that are built into the model. The report generated by TOPKAT was checked to see if there were any warnings (listed under the subsection "TOPKAT Warnings" below). Only those chemicals that had no warnings were used in assessing the performance of the LOAEL model in this study.

Last, the QSTR similarity distance[3,11] of the query chemical from chemicals in the LOAEL model database was determined manually using the "Similarity Search" feature in TOPKAT. A similarity search on a query chemical provides a listing of the experimental LOAEL of chemicals in the model's database, the corresponding similarity distance and LOAEL predicted by TOPKAT. Only chemicals that had a similarity distance ≤ 0.2 from the chemicals in the LOAEL model database were used in this study.[26] A detailed description of the similarity search feature is provided by Moudgal et al.[11]

TOPKAT Warnings. If certain one- or two-atom fragments are missing in the LOAEL model substructural library for any given chemical, or if the fragments are not adequately covered by the chemicals in the model database, TOPKAT issues a warning stating that a prediction cannot be made or may be unreliable. Figure 1 illustrates the fragments that were not covered by TOPKAT's LOAEL model for the chemicals considered in this study. In addition, TOPKAT was also unable to predict LOAELs for charged sulfur species (S^+, S^-) and metal-containing organics, though the metals Sn, Pb, Na, and K were recognized by LOAEL models in TOPKAT versions prior to version 5.0.

TOPKAT automatically compares the predicted LOAEL values with either an experimental LD50 value or an estimate of the LD50 value from the rat oral LD50 model in TOPKAT. TOPKAT issues a warning if the LD50 value is lower than the LOAEL value for a chemical.[3,15] TOPKAT also automatically determines whether the submitted chemical structure is within the model's OPS, the region in descriptor space where LOAELs can be calculated with the greatest confidence.[3,11,25]

Data Sets. Two chemical databases were used to determine the percentage of accurate predictions by the oral rat chronic LOAEL model in TOPKAT. The OPP database contained oral rat chronic LOAELs of 343 chemicals that were of interest to the U.S. EPA's OPP.[27] The second dataset pooled together the LOAELs of 313 chemicals (hereafter referred to as the IHP database) from the U.S. EPA's IRIS,[12] HEAST,[13] and PTV[14] databases. The PTV contains a database of chemical toxicity assessments that can be obtained by contacting regional U.S. EPA offices. Of the 313 chemicals in the IHP database, 262 chemicals were from IRIS, while HEAST and PTV contributed 10 and 41 chemicals, respectively, to the database.

RESULTS AND DISCUSSION

The LOAEL predictions for the chemicals in the OPP and IHP databases were compared separately against the results from the original Mumtaz et al. study[20] and the study by Oxford Molecular.[23] The comparison analysis was performed separately for the two databases because they covered different classes of compounds. While the chemicals in the OPP database were pesticides, the chemicals in the IHP database were mostly industrial chemicals or chemicals that are commonly found at hazardous waste sites. The results of the analyses are presented in the succeeding sections.

OPP Database LOAEL Comparisons. Of the 343 chemicals in the OPP database that had an experimental LOAEL, TOPKAT could not recognize 28 chemicals that were either metal−ligand coordination complexes or polymers. Of the remaining 315 chemicals (343 − 28), TOPKAT was unable to estimate LOAELs for 106 chemicals because of the warnings discussed in Methods. Of the 106 chemicals, 17 had unknown fragments, 27 chemicals had a predicted LOAEL that was greater than the experimental and/or predicted LD50, 53 chemicals were outside the OPS, and 9 chemicals failed the similarity search criteria. Overall, TOPKAT was unable to estimate LOAELs for 134 of the 343 (~39%) of the chemicals in the OPP database.

Of the remaining 209 chemicals (343 − 134) that had a valid LOAEL prediction, 58 chemicals were in TOPKAT's database ("DB" in Table 2). The percentage of chemicals that had a LOAEL predicted to within a factor of 2, 5, and 10 of experimental LOAELs were 43, 66, and 80%, respectively (Table 2; valid %, last row). These results do not compare well with those reported by Mumtaz et al.[20] (namely, 55, 93, and 100% of the chemicals had their LOAELs predicted to within a factor of 2, 5, and 10 of their experimental LOAELs, respectively). These results are not surprising since Mumtaz et al.[20] used all chemicals in TOPKAT's LOAEL model database for their model evaluation analysis, while this study used a significant number (160; 151 of the 209 chemicals + 9 chemicals that failed the similarity search criteria; Table 2) of nondatabase chemicals in the analysis. Hence, the comparison analysis conducted in this study (similar to cross-validation using a test set of chemicals) is a more accurate representation of the true predictive ability of TOPKAT's LOAEL model than the results reported by Mumtaz et al.[20]

Table 2. Comparison Analysis for Chemicals in the OPP Database with the Results from Mumtaz et al.[20] and Oxford Molecular[23,a]

chemical subclass	type	total	no. within a factor of >10	6-10	3-5	1-2	SS	no. misclassified	% misclassified
acyclic	error	23	12		4	2	5	15	68
	DB	6	1		1	4		2	29
	non-DB	9	1	1	3	4		2	22
	sum	38	14	1	8	9	5	19	50
	valid	15	2	1	4	8			
	valid %		13	7	27	53			
alicyclic	error	11	3	2	1		5	7	88
	DB	7	1			6		4	40
	non-DB	13	2	1	2	2	6	10	77
	sum	31	5	1	3	8	11	21	68
	valid	14	3	1	2	8			
	valid %		21	8	14	57			
heteroaromatic	error	29	13	1	3	6	6	18	67
	DB	14	1	1	2	10		1	6
	non-DB	65	16	7	18	23	1	27	42
	sum	108	30	9	23	37	7	46	43
	valid	78	17	8	20	33			
	valid %		22	10	26	42			
multiple benzene	error	6	1	1	1	1	2	4	67
	DB	7		1	3	3		1	14
	non-DB	32	8	7	5	12		15	47
	sum	45	9	9	9	16	2	20	44
	valid	39	8	8	8	15			
	valid %		21	21	21	38			
single benzene	error	28	10	4	3	5	6	16	59
	DB	24	2	5	5	12		6	24
	non-DB	41	9	7	9	14	2	17	41
	sum	93	21	15	17	31	8	39	42
	valid	63	11	12	14	26			
	valid %		17	19	22	41			
overall	error	97	39	8	12	14	24	62	64
	DB	58	5	7	11	35		12	21
	non-DB	160	36	23	37	55	9	72	45
	sum	315	80	38	60	104	33	146	46
	valid	209	41	30	48	90			
	valid %		20	14	23	43			

[a] The numbers of missing fragments were not included in this table. The factors in the table were rounded to the nearest integer. In the table, DB indicates chemicals that are in TOPKAT's database, non-DB indicates chemicals that are not in TOPKAT's database, sum is the number of chemicals, valid is the number of valid predictions (chemicals that passed the validation criteria and did not have any of the warnings mentioned under Methods), valid % is the percentage of valid predictions, SS is the number of chemicals that failed the similarity search criteria, and error is the number of chemicals that had missing fragments, OPS, and/or LOAEL > LD50 type of errors. The last two columns represent the number of chemicals that are misclassified (chemicals whose experimental LOAELs fall outside the 95% confidence intervals predicted by TOPKAT) and the percentage of chemicals that are misclassified.

In a second analysis, the comparison exercise was repeated on the 209 OPP chemicals (with valid LOAEL predictions) after being broken into 5 groups depending on their molecular structure: acyclic, alicyclic, heteroaromatic, and single and multiple benzenes. The percentage of chemicals that were predicted accurately to within a factor of 2 and 5 were 53 and 80% (Table 2), respectively, vs the reported values of 73 and 100, respectively[21−23] (Table 1) for acyclic chemicals. The corresponding percentages were 57 and 71% vs 94 and 100%, respectively, for alicyclics; 42 and 68% vs 78 and 100%, respectively, for heteroaromatics; 38 and 59% vs 70 and 97%, respectively, for multiple benzenes; and 41 and 63% vs 66 and 98%, respectively, for single benzenes. For all chemical subclasses, the percentage of LOAEL values predicted accurately decreased considerably when compared to the original study.[21−23] In addition, the acyclic submodel (80%) seemed to predict the greatest percentage of chemicals accurately to within a factor of 5 of experimental LOAEL, while the multiple benzene submodel (59%) seemed to be the poorest. However, considering the number of erroneously predicted LOAELs (last three columns in Table 2), the alicyclic submodel seemed to predict the poorest followed by acyclic, single benzene, heteroaromatic, and multiple benzene submodels.

IHP Database Comparisons. Since a majority of the chemicals that were used in the development of the initial LOAEL model construction process were from U.S. EPA documents, a lot of commonality was expected between the chemicals in TOPKAT's database containing 393 chemicals and the 313 chemicals in the IHP database. Of the 313 chemicals that had a reported chronic or subchronic LOAEL in the IHP database (including 9 chemicals with missing fragments not included in Table 3), 84 chemicals (∼27%) had missing fragments, were outside the OPS, or had an experimental/predicted LD50 value lower than the calculated LOAEL. Of the remaining 229 chemicals, 100 chemicals were in TOPKAT's database (Table 3). The percentage of chemicals that were predicted to within a factor of 2, 5, and 10 of the reported LOAEL values were 48, 69, and 79%, respectively (Table 3). These results do not compare well with those reported by Mumtaz et al.[20] (55, 93, and 100%, respectively), and were similar to those obtained for the chemicals in the OPP database.

In a second analysis, the comparison exercise was repeated on the IHP database of chemicals after being broken into the 5 groups mentioned in the previous section. Results of the analysis indicate that the percentage of chemicals that were predicted to have a LOAEL within a factor of 2 and 5 of their experimental LOAELs were 54 and 70%, respectively, for acyclics; 65 and 83%, respectively, for alicyclics; 47 and 69%, respectively, for heteroaromatics; 50 and 79%, respectively, for multiple benzenes; and 40 and 61%, respectively, for single benzenes (Table 3). Except for an ∼10% difference in the percentage of predictions for alicyclics and multiple benzenes, the results were similar to those obtained for the chemicals in the OPP database.

Experimental Data Quality Analysis. The role of the quality of experimental LOAELs on TOPKAT's predictive ability was analyzed by choosing only those chemicals in the OPP database that met the OPP core grade classification guidelines[28] (which explicitly states the minimum experimental conditions of a bioassay and what adverse effects to look for, and hence presumably reflects studies of high quality) during their bioassay process. Hence, any differences in the predictive ability of TOPKAT for the whole database vs core-grade-certified chemicals could have been due to using poor-quality experimental data for comparison purposes as opposed to poor predictive ability. Results of the analysis done using the subset of 113 chemicals that met the quality criteria are summarized in Figure 2. In the figure, thick gray, dotted black, and thick black lines enclose chemicals whose LOAEL values are within a factor of 2, 5, and 10 of their respective observed values. Numerically, 49% of the chemicals had estimated LOAELs within a factor of 2 of their experimental values (vs 55% for the study by Mumtaz et al.[20]), while 71% were within a factor of 5 (vs 93% for the

Table 3. Comparison Analysis for Chemicals in the IHP Database with the Results from Mumtaz et al.[20] and Oxford Molecular[23,a]

chemical subclass	type	total	no. within a factor of >10	6-10	3-5	1-2	SS	no. misclassified	% misclassified
acyclic	error	31	18	3	1	4	5	23	74
	DB	25	2		5	18		3	12
	non-DB	32	11	4	4	12	1	16	50
	sum	88	31	7	10	34	6	42	48
	valid	56	13	4	9	30			
	valid %		23	7	16	54			
alicyclic	error	6	3	1	1		1	5	83
	DB	13	1			12		1	8
	non-DB	11	3		4	3	1	8	73
	sum	30	7	1	5	15	2	14	47
	valid	23	4	0	4	15			
	valid %		17	0	17	65			
hetero-aromatic	error	10	2	1		4	3	3	30
	DB	11			3	8		0	0
	non-DB	25	10	1	5	9		14	56
	sum	46	12	2	8	21	3	17	37
	valid	36	10	1	8	17			
	valid %		28	3	22	47			
multiple benzene	error	10	4		2	2	2	6	60
	DB	16		1	5	10		1	6
	non-DB	18	3	3	5	7		6	33
	sum	44	7	4	12	19	2	13	30
	valid	34	3	4	10	17			
	valid %		9	12	29	50			
single benzene	error	16	10	2	2	1	1	10	63
	DB	35	6	7	5	17		13	37
	non-DB	45	11	7	12	15		17	38
	sum	96	27	16	19	33	1	40	42
	valid	80	17	14	17	32			
	valid %		21	18	21	40			
overall	error	73	37	7	6	11	12	47	64
	DB	100	9	8	18	65		18	18
	non-DB	131	38	15	30	46	2	61	47
	sum	304	84	30	54	122	14	126	41
	valid	229	47	23	48	111			
	valid %		21	10	21	48			
subchronic	error	20	14	2	1	2	1	17	85
	DB	13	1		3	9		1	8
	non-DB	50	18	4	12	16		22	44
	sum	83	33	6	16	27	1	40	48
	valid	63	19	4	15	25			
	valid %		30	6	24	40			
chronic	error	53	23	5	5	9	11	30	57
	DB	87	8	8	15	56		17	20
	non-DB	81	20	11	18	30	2	39	48
	sum	221	51	24	38	95	13	86	39
	valid	166	28	19	33	86			
	valid %		17	11	20	52			

[a] The numbers of missing fragments were not included in this table. The factors in the table were rounded to the nearest integer. Terms used in this table are explained in Table 2.

study by Mumtaz et al.). The corresponding values obtained for the study involving 209 chemicals (see OPP database LOAEL comparisons section) were 43 and 66%, respectively. These results indicate that while the percentage of accurate predictions increased slightly for chemicals that followed OPP core-grade guidelines, the overall predictions were still below the correlations reported by Mumtaz *et al.*[20] Hence, the predictive ability of TOPKAT's LOAEL model was not dependent on the quality of experimental LOAELs considered in this exercise.

Misclassification Analysis. In the literature, there is often a 5−10-fold gap between the dose at which no adverse effects are noted in test animals (no observed adverse effect level, NOAEL) and the dose at which the lowest adverse effect is observed (LOAEL). Hence, the "true" LOAEL may be any dose between the NOAEL and the reported LOAEL. To account for the difference between the true and reported LOAELs while accounting for statistical errors during the model building process, the experimental LOAELs for the chemicals in the OPP database were compared against the 95% confidence intervals of the computed LOAEL. The difference between the upper and lower bounds of the 95% confidence intervals generally was an order of magnitude (10-fold) for the LOAEL model.[15] Thus, experimental LOAELs falling within the 95% confidence intervals were classified as "accurate" for the purposes of this analysis.

Table 2 lists the number and percentage of chemicals whose LOAELs fell outside the 95% confidence limits (referred to as misclassification in the table). The results of the comparison indicate that 21% of chemicals in TOPKAT's database were misclassified, while 45% of the nondatabase chemicals were misclassified (last column of Table 2). Overall, the misclassification rate was 46% for the OPP database, which indicates that the experimental LOAEL is well outside the range predicted by the LOAEL model, and, therefore, TOPKAT is a poor predictor of LOAELs for the OPP database chemicals. These results indicate that TOPKAT is expected to misclassify roughly 20% of the query chemicals that are in TOPKAT's database and 45% of the query chemicals that are not present in TOPKAT's database.

The misclassification rates for the chemicals in the IHP database were almost similar to those obtained for the chemicals in the OPP database. For the IHP database chemicals, the misclassification rate for the chemicals in TOPKAT's database was 18% (DB; Table 3) while the misclassification rate for chemicals not in the database was 47% (non-DB; Table 3). The rates compared well with those obtained for the chemicals in the OPP database (21 and 45%, respectively).

TOPKAT vs IHP Database LOAEL Comparison. Since the performance of a QSAR model depends on the numerical value of the LOAELs used in generating the model (i.e., the training set), LOAELs of the chemicals in TOPKAT's database were compared with those in the IHP database. There were 113 chemicals that were common to the two databases. Only LOAELs reported on IHP databases were used in the comparison process because a majority of the chemicals in TOPKAT's database were from IRIS.[12]

The results of the comparison indicate a wide difference between the experimental LOAELs reported in the IHP database vs those used in TOPKAT's database (Figure 3). In Figure 3, subchronic LOAELs in the IHP database are illustrated as open circles and chronic LOAELs are illustrated as filled diamonds, while the thick line represents the 1:1 line. As seen in Figure 3, most of the data points fall away from the 1:1 line, indicating a difference in LOAELs from the two databases. The LOAEL values used in TOPKAT's model database for a particular chemical may differ from those reported in the IHP database either because of differing experimental procedures, because of lower LOAELs reported after the completion of the LOAEL model in TOPKAT, or because newer experiments were conducted using tighter dose groups (e.g., 2-fold vs 5-fold variation between doses). Since TOPKAT's database and, hence, the QSARs for LOAEL prediction cannot be updated with new toxicity data nor can new chemicals be added to the existing database,

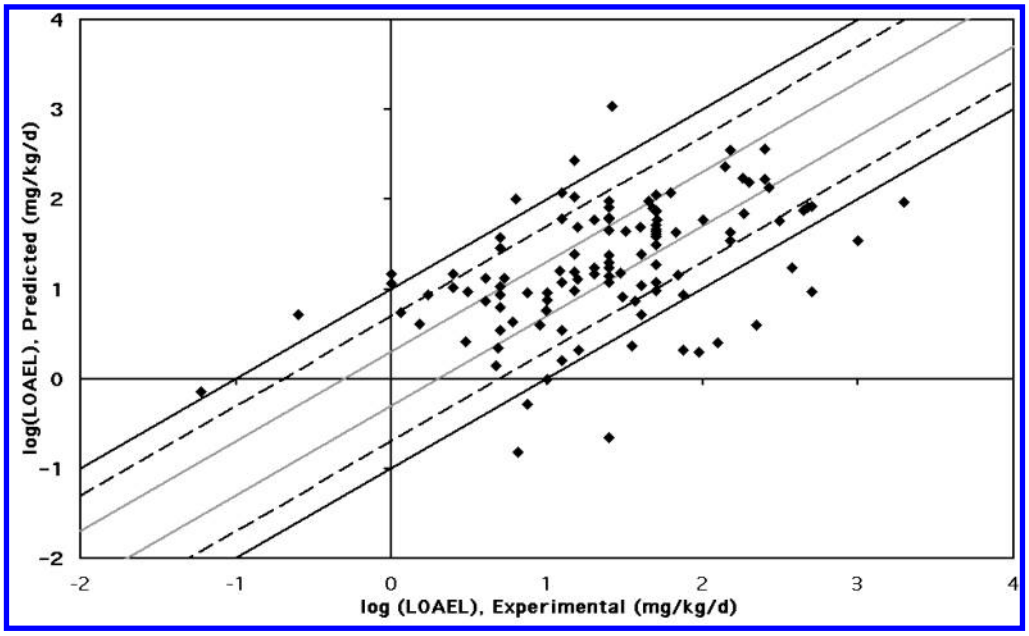

Figure 2. Predicted vs observed LOAEL for the 113 OPP database chemicals that met the EPA guidelines for animal testing. Both axes are in log units. LOAELs are in mg/(kg•day) units. Data points within the thick gray lines are within a factor of 2 of the experimental values, while those predicted to within a factor of 5 and 10 are within the dotted black line and thick black line, respectively.

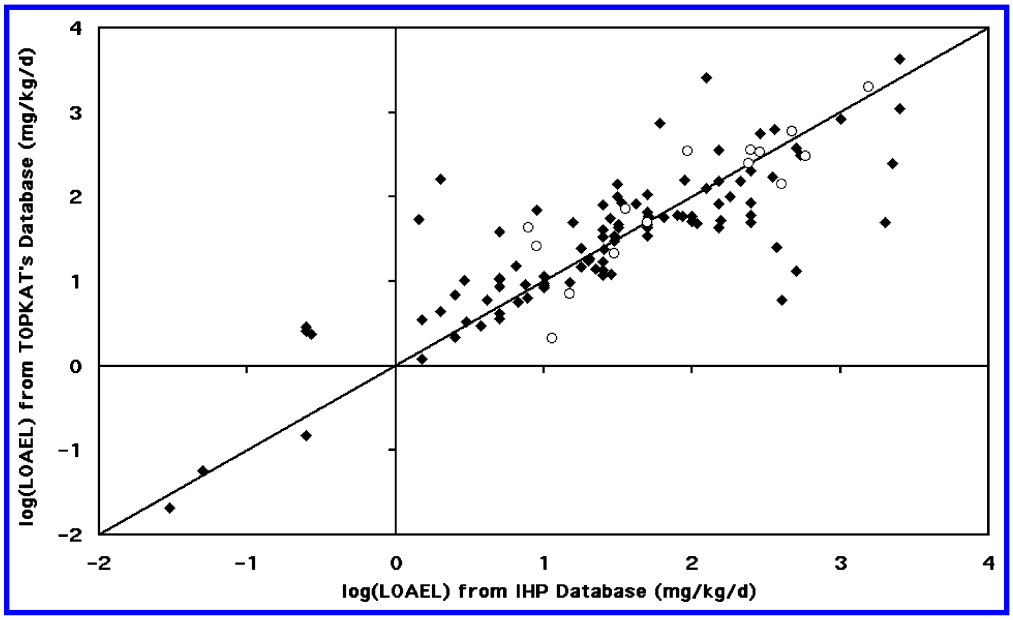

Figure 3. Comparison of experimental LOAELs reported in TOPKAT's database vs those reported in the IHP database for 113 chemicals. The thick line that passes through the origin is the 1:1 line. Filled diamonds represent LOAELs from chronic studies while open circles represent LOAELs from subchronic studies.

the performance of the LOAEL model would be expected to decrease over time.

Chronic vs Subchronic LOAEL Comparison. The duration of the study (subchronic vs chronic) seemed to have an affect on the predictive ability of TOPKAT's LOAEL model. On separating the IHP database chemicals into two groups based on the duration of the study, the percentage of chemicals within a factor of 2, 5, and 10 of the experimental LOAELs were 40, 64, and 70%, respectively, for subchronic vs 52, 72, and 83%, respectively, for chemicals with chronic LOAELs (Table 3). The results indicate that while the oral rat chronic LOAEL model can be used to predict oral subchronic LOAELs, there is a slight loss of accuracy over chronic LOAEL predictions. However, in general, the percentages of accurate predictions for the two durations were similar to that of the chemicals in the OPP database.

Model Robustness. The robustness of the TOPKAT rat chronic LOAEL model was also investigated in this study for the chemicals in the OPP and IHP databases. A robust QSAR is generally able to accurately predict end points for a wide variety of chemicals, while a nonrobust QSAR generally fails to accurately predict end points for chemicals that have not been used in the initial QSAR model development process. The percentage of LOAELs predicted accurately to within a factor of 2, 5, and 10 of experimental values were 60, 79, and 91%, respectively, for OPP chemicals in TOPKAT's database (DB, Table 2) and 65, 83, and 91%, respectively, for IHP chemicals in TOPKAT's

database (DB, Table 3). The corresponding percentages for nondatabase chemicals were 34, 57, and 72%, respectively, for chemicals in the OPP database (non-DB, Table 2) and 35, 58, and 69%, respectively, for chemicals in the IHP database (non-DB, Table 3). The percentages of accurate predictions for nondatabase chemicals did not compare well with predictions for chemicals in TOPKAT's database. These results thus indicate that the accuracy of LOAEL prediction is highly dependent on the chemicals used in the initial model development process.

Implications on Future Toxicity Assessment. One of TOPKAT's greatest strengths is that its models are statistically based and are hence able to include a database of chemically diverse structures in its various models as well as provide toxicity estimates for several complex toxicity end points. However, the LOAELs for a given chemical may be due to various toxicity end points or biological activities such as hepatotoxicity, reproductive/developmental toxicity, neurotoxicity, and immunotoxicity. Each specific toxicity end point is in general due to a relatively simple mechanism of action, which in turn contributes unique bulk, electronic, and transport attributes to QSARs that quantify the biological activity. The attributes for end-point-specific QSARs may differ from those for a combined toxicity end-point QSAR such as TOPKAT's LOAEL model. Hence, the predictions made by the TOPKAT model may be different from those made by end-point-specific QSARs, which may be one of the reasons behind the misclassifications reported in this manuscript. Thus, TOPKAT's major strength may also be one of its major weaknesses.

TOPKAT provides several validation procedures that help increase confidence in predictions, provides access to its database of chemicals, and provides a means to automatically identify chemical analogues based on structural features that has helped NCEA develop the surrogate approach.[11] However, the models in TOPKAT lack the ability to "learn" as users cannot update the database of chemically diverse structures with new data on chemicals as and when they become available in the literature. Moreover, most of the QSAR models in TOPKAT have not been updated since their initial release. As the number of chemically diverse structures with exotic functional groups that are produced every year increases, the percentage of model predictions with missing fragments, unrecognized functional groups, and estimations outside the model prediction domain will also increase, thus reducing the overall predictive ability of the models. Hence, since TOPKAT does not provide the ability for users to input their own toxicity data and/or take steps to keep the models updated on a regular basis, the use of TOPKAT is expected to be limited. However, since TOPKAT is the only software currently on the market that provides quantitative LOAEL values, the feasibility of using the LOAEL model in TOPKAT for various applications is left to the sole discretion of the researcher/user. In the meantime, the authors suggest that only those chemicals that meet the validation criteria described in Methods be used in user applications. The authors also recommend that future models allow users the capability of updating their model database on a regular basis.

ACKNOWLEDGMENT

The authors would like to thank Jeff Swartout of NCEA, U.S. EPA, Cincinnati, for the use of his database of OPP chemicals. This research was supported in part by the appointment of R.V. to the Postgraduate Research Program at the National Center for Environmental Assessment administered by the Oak Ridge Institute for Science and Education, through an interagency agreement between the U.S. Department of Energy and the U.S. Environmental Protection Agency. The views expressed in this manuscript are those of the authors alone and do not necessarily reflect the views and policies of the U. S. Environmental Protection Agency. Mention of trade names or commercial products in the manuscript does not constitute endorsement or recommendation for use.

REFERENCES AND NOTES

(1) *Guidelines for Exposure Assessment*: 1992. U.S. Environmental Protection Agency, Risk Assessment Forum, No. EPA/600Z-92/001. USEPA: Washington, DC, 1992.
(2) Lewis, D. F. V. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH Publishers: New York, 1992; Vol. III.
(3) Gombar, V. K. In *Advances in Molecular Toxicology*; Reiss, C., Labbe, G., Parvez, H., Eds.; VSP: Utrecht, Netherlands, 1998; pp 125−199.
(4) Purcell, W. P.; Bass, G. E.; Clayton, J. M. *Strategy of Drug Design: A Guide to Biological Activity*; Wiley: New York, 1973.
(5) Lewis, D. F. V.; Ioannides, C.; Parke, D. V. In *Proceedings of the Conference on Predicting Chemical Carcinogenesis in Rodents* (Research Triangle Park, NC); National Institute of Environmental Health Sciences: Washington, DC, 1992; pp 45−49.
(6) Lewis, D. F. V.; Ioannides, C.; Parke, D. V. *Mutat. Res.* **1993**, *291*, 61−77.
(7) Dearden, J. C.; Barrett, M. D.; Benigni, R.; Bristol, D. W.; Combes, R. D.; Cronin, M. T. D.; Judson, P. N.; Payne, M. P.; Richard, A. M.; Tichy, M.; Worth, A. P.; Yourick, J. J. *ATLA, Altern. Lab. Anim.* **1997**, *25*, 223−252.
(8) Wang, S.; Milne, G. W. *Chem. Res. Toxicol*. **1993**, *6*, 748−753.
(9) Richard, A. M. *Mutat. Res*. **1994**, *305*, 73−97.
(10) Richard, A. M. *Mutat. Res*. **1998**, *400*, 493−507.
(11) Moudgal, C. J.; Venkatapathy, R.; Choudhury, H.; Bruce, R. M.; Lipscomb, J. C. *Environ. Sci. Technol*. **2003**, *37*, 5228−5235.
(12) *Integrated Risk Information System*, 2002. National Center for Environmental Assessment, U.S. Environmental Protection Agency: Washington, DC, 2002.
(13) *Health Effects Assessment Summary Tables*, 1997. Office of Emergency and Remedial Response, U.S. Environmental Protection Agency: Washington, DC, 1997; NTIS No. PB97-921199.
(14) *Provisional Toxicity Database*, 2003. National Center for Environmental Assessment, U.S. Environmental Protection Agency: Cincinnati, OH, 2003.
(15) *TOPKAT User Guide 6.1*; Accelrys, Inc.: Burlington, MA, 2001.
(16) Hall, L. H.; Mohney, B.; Kier, L. B. *J. Chem. Inf. Comput. Sci*. **1991**, *31*, 76−82.
(17) *TOPKAT 5.0 Reference Manual*; HDi Inc.: Rochester, NY, 1997.
(18) Fujita, T.; Iwasa, J.; Hansch, C. *J. Am. Chem. Soc.* **1964**, *86*, 5175−5180.
(19) Kier, L. B. *Quant. Struct.−Act. Relat*. **1986**, *5*, 1−7.
(20) Mumtaz, M. M.; Knauf, L. A.; Reisman, D. J.; Peirano, W. B.; DeRosa, C. T.; Gombar, V. K.; Enslein, K.; Carter, J. R.; Blake, B. W.; Huque, K. I.; Ramanujam, V. M. S. *Toxicol. Lett*. **1995**, *79*, 131−143.
(21) *Final Report to the U. S. Environmental Protection Agency* (Cincinnati, OH, in fulfillment of Contract No. 68-C4-0006); HDi Inc.: Rochester, NY, 1995.
(22) Enslein, K. In *Advances in Molecular Toxicology*; Reiss, C., Labbe, G., Parvez, H., Eds.; VSP: Utrecht, Netherlands, 1998; pp 141−164.
(23) *TOPKAT User Guide 6.0*; Oxford Molecular Ltd.: Oxford, U.K., 2001.
(24) Weininger, D. *J. Chem. Inf. Comput. Sci*. **1988**, *28*, 31−36.
(25) Gombar, V. K.; Enslein, K. *J. Chem. Inform. Comput. Sci.* **1996**, *36*, 1127−1134.
(26) Gombar, V. K. GlaxoSmithKline, Research Triangle Park, NC, personal communication, 2001.
(27) National Pesticide Information Retrieval System. Online. http://ceris.purdue.edu/npirs/index.html (accessed Nov 2003).
(28) *Pesticide Assessment Guidelines, Subdivision F. Hazard Evaluation: Human and Domestic Animals*; Office of Prevention, Pesticides and Toxic Substances, Hazard Evaluation Division, U.S. Environmental Protection Agency, U.S. Government Printing Office: Washington, DC, 1984.

CI049903S