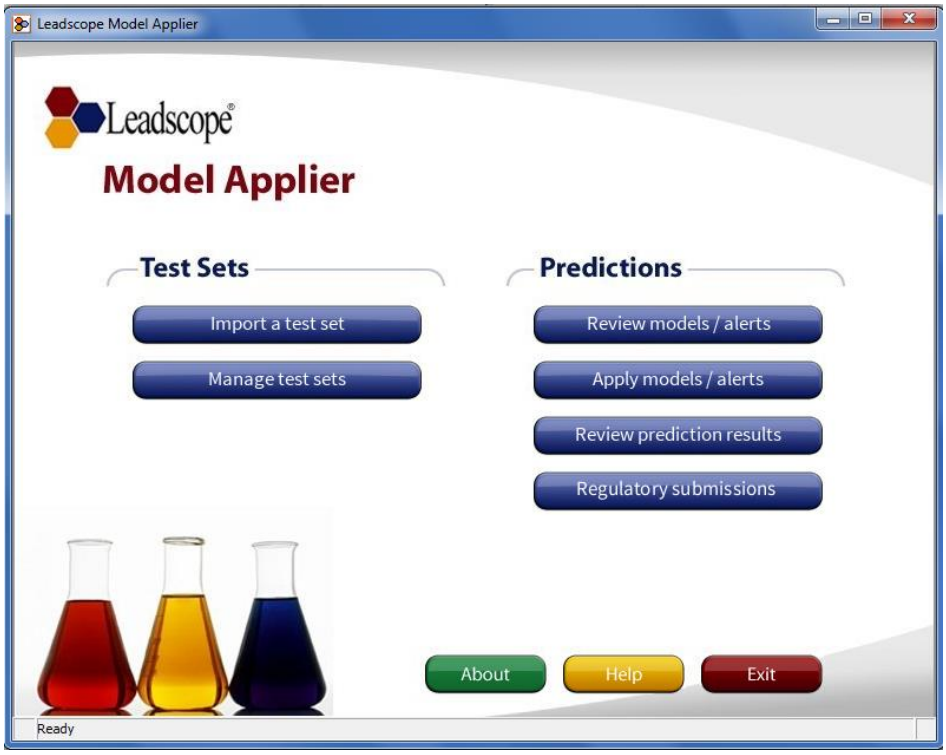
Leadscope Model Applier

Documentation



Version 2.2
May 2017

Leadscope Inc.
1393 Dublin Road
Columbus, OH 43215

Disclaimer

US FDA Disclaimer: This Research Collaboration Agreement has been developed between the US FDA and Leadscope, Inc. and does not explicitly state or imply that the US FDA endorses, recommends, or requires the use of this product or any other Leadscope, Inc. product.

Leadscope® Model Applier containing (Q)SAR models developed under the US FDA/Leadscope, Inc. Research Collaboration Agreement do not necessarily reflect the US FDA's current regulatory position on any given compound. Furthermore, no training set data that are supplied by stakeholders and interpreted by US FDA scientists and contractors in the review of submissions to the US FDA have been included in this product.

Leadscope hereby specifically disclaims liability for any and all adverse consequences or effects of any nature, which arise, may arise, or may be considered to have arisen, as a result of using this software application, Leadscope® Model Applier, by any person in the creation of any type of new drug or pharmacological product or substance.

Leadscope® Model Applier and the related prediction data may not be used for any purpose outside Licensee's internal use. Licensee agrees not to redistribute or export the results from the Leadscope Model Applier to any country to which such export or transmission is restricted by applicable U.S. regulation or statute. Licensee agrees not to redistribute or export training set structures and/or data from the Leadscope Model Applier. Licensee agrees not to publish any prediction results from the Leadscope Model Applier over the internet or journal articles without the prior expressed written consent of Leadscope. Leadscope and U.S. Food & Drug Administration disclaim any liability for any damages, whether direct, indirect, special, incidental or consequential arising from Licensee's use of the Leadscope Model Applier or any results generated there from. Portions of Leadscope Model Applier and data are copyrighted © Copyright 2017, U.S. Government. All rights reserved.

Table of Contents

A. General Description

The Leadscope® Model Applier provides easy-to-use (Q)SAR models and expert alerts to obtain decision support information on the potential toxicities of chemicals. All the (Q)SARs were constructed at the Food and Drug Administration (FDA) by the Division of Applied Regulatory Science (DARS), previously known as the Division of Drug Safety Research Staff. The models were built under a Research Collaboration Agreement (RCA) using the Leadscope Prediction Data Miner software. The training data sets were compiled by DARS; complete documentation of the weight of evidence methodology used for the preparation of the model training sets and the sources of the data have been published by the DARS group [1-22,26-29]. In this document, these (Q)SAR models will be referred as "RCA models". Prediction modeling methodology employed in Leadscope Prediction Data Miner for both binary and continuous data is described in the literature [22]. The Leadscope Genetox Expert Alerts were constructed by Leadscope [23].

The RCA (Q)SAR models were built from public information and include the training set structures and calls as part of the model description. The RCA models are implemented with molecular descriptors that include structural features and up to 8 calculated properties. The structural features include selected Leadscope® default hierarchy features plus the predictive scaffolds generated with default settings. The eight calculated properties include parent molecular weight, parent atom count, LogP, polar surface area, hydrogen bond acceptors, hydrogen bond donors, number of rotational bonds and, Lipinski score (rule violation). More details are presented in the section C-3.

In version 2.1, there are eight suites containing RCA (Q)SAR models with over 80 biological endpoints. The suites can be divided into two major groups of endpoint models: 1) human clinical endpoints; and 2) non-human toxicity endpoints. The first group includes three suites of models which predict the effects of pharmaceuticals based upon human clinical data, including: adverse cardiac effects, adverse hepatobiliary effects, and adverse urinary tract effects. The second group includes five different suites, predicting toxicities of organic chemicals based upon results of *in vivo* animal toxicity and *in vitro* studies. They include: carcinogenicity in rodents, genetic toxicity (i.e., mutagenicity and clastogenicity), reproductive toxicity in male and female rodents, developmental toxicity (i.e., dysmorphogenesis, fetal development, and survival of the rodent fetus), and neurotoxicity in newborn rodents. The majority of the (Q)SAR models are specifically designed to provide decision support information for toxicological and clinical endpoints that are considered important in the safety analyses of FDA-regulated substances. However, the Model Applier also includes a subset of models that have proven to be useful for research on toxicological properties of chemicals, but are not used internally within the FDA to support regulatory decisions.

Each suite has many different RCA (Q)SAR models for individual endpoints and they are designed to be used in sets to make predictions of specific chemical toxicities. For some toxicity

endpoints, sub-models are constructed to improve the predictive performance of the models. The predictive performance of global models depends highly on the ratio of actives (toxic) to inactives (non-toxic) chemicals in a training set. A training set was divided to subsets to maintain optimal active-to-inactive (A/I) ratio to balance high sensitivity with high specificity. The rationale behind these RCA models is that predicting true negatives must be maximized while false negatives must be minimized in product safety analysis within the regulatory agencies. Sometimes several sub-models are built for each model at the optimal A/I ratio as part of an "Average Model". For best results, Leadscope runs each of the sub-models behind the scene and displays an overall result. The overall prediction results are based on averaging the probabilities (likelihood of being positive) from appropriate sub-models. The statistics of overall models are listed in the next section B. The overall models as well as sub-models are available for review in the "Review a predictive model" wizard. More details are available in Section C-3.

The prediction results for each model are presented as the "prediction" and the "positive prediction probability". The prediction can be "Positive", "Negative", and "Not-In-Domain". The positive prediction probability is given as the likelihood value between 0 (non-toxic) and 1 (toxic). If the prediction was a result of using existing experimental data, the word is a hyperlink which links to the found data (e.g. Positive). The higher the probability is, the greater chance of the test chemical being toxic in a particular endpoint. For most models, a test chemical is evaluated as active for a set of models (e.g., mouse composite) if the average probability is ≥ 0.5 and inactive if the average probability is <0.5. The exceptions are the Genetox Salmonella and E. coli AT models where the accepted cutoffs are inactive <0.4 and active ≥ 0.6. As or version 2.2, the models themselves contain the correct cutoff settings and the user no longer has to set them. Test chemicals outside the domain space of the models are so noted. The model domain is defined within the Leadscope application by two criteria: 1) the test compound containing structural model features in addition to property descriptors; 2) the test compound being similar to at least one training set compound (with at least 30 % similarity). More details can be found in the references cited in Section D.

As of version 1.6, the Leadscope SAR Genetox and SAR Carcinogenicity Databases were provided with the purchase of those model suites respectively in Leadscope Model Applier and Leadscope Enterprise. The Model Applier allows analog browsing of these databases after a prediction has been made on a test set of compounds.

B. Suites and Models

B-1. Rodent Carcinogenicity Statistical Suite

Most of the models in the Rodent Carcinogenicity Suite are intended to support regulatory decision-making processes. The RCA (Q)SAR models for prediction of carcinogenicity in rodents are based upon the DARS publications [2-5, 12, 14]. Rodent carcinogenicity was modeled for study calls, where the positive calls are trained as binary 1 and negative calls as binary 0. The outcome of a (Q)SAR prediction is given as the probability of being carcinogenic on a scale of 0 to 1 as described in the Section A. The lower the probability value, the lower the carcinogenic potential of a test chemical. The level of concern for test chemical carcinogenic potential is proportional to the number of sets of models with positive predictions.

In this suite, there are two sets of RCA (Q)SAR models for predicting carcinogenicity endpoints. There are rodent models based on the 2-year rodent bioassays as well as cell transformation *in vitro* assays. The rodent models are used by the FDA as supporting information for decision-making processes involved in the regulatory workflow. However, the cell transformation models are not used by the FDA for decision-making processes but rather only for R&D purposes.

The rodent carcinogenicity models were rebuilt for version 2.0 of the Leadscope Model Applier software [26]. This represents the third version of the rodent carcinogenicity models. A large, high-quality rodent carcinogenesis database (1682 compounds), covering a large number of structural alerts and characteristics of both genotoxic and non-genotoxic carcinogens, was reconstructed and used as a basis for constructing (Q)SAR models for predicting four rodent carcinogenicity study groups: male rat, female rat, male mouse, and female mouse carcinogenicity. All original training set source data was re-examined and tumor findings were reported per major organ for each study group. Newly declassified data and additional public data were added to the training set. A tumor severity scoring system was developed to encode tumor severity and significance for each chemical in each data set. A score of zero corresponds to the absence of tumors. Increasing score values reflect the presence of multi-site tumors, tumors present across genders and/or species, and the conservation of tumor sites across genders and/or species. The maximum score is 50, representing the largest risk of human carcinogenic potential of the test article. For the purpose of constructing these (Q)SAR classification models, an overall binary score was applied, where a significant tumor finding over background in any organ system is defined as a positive.

When compared to performance of earlier generation models for this endpoint, the new models demonstrate comparable performance statistics; however, the models have the benefit of being based on more examples of newly marketed drugs and the use of contemporary scoring criteria to define a positive tumor response. The newly constructed models maintain good overall performance compared to the previous versions, including good sensitivity and negative predictivity, which are critical parameters for the safety assessment of drug products.

The number of molecules in the training set and their cross-validated predictive performance are presented in Table 1. External validation has been reported by the FDA in 2015 [26]. In that report they recommend using a positive probability cutoff of 0.4 for the maximum negative prediction and 0.6 for the minimum positive prediction, with probabilities of 0.4-0.6 representing equivocal or indeterminate predictions.

Table 1. Summary of Statistics of Models in the Rodent Carcinogenicity Suite

Endpoint model names	# training compounds	Sensitivity (%)	Specificity (%)
Carcinogenicity Male Mouse	1267	62.2	84.0
Carcinogenicity Female Mouse	1209	62.8	83.9
Carcinogenicity Male Rat	1406	56.9	76.8
Carcinogenicity Female Rat	1395	62.4	84.5
In Vitro Cell transformation	640**	87.8	50.8
SHE	425***	88.8	55.8
BALB/c-3T3	316†	87.8	54.7
C3H10T1/2	138††	93.9	22.5

** The training set has 62% positive chemicals.
*** The training set has 65% positive chemicals.
† The training set has 59 % positive chemicals.
†† The training set has 71 % positive chemicals.

B-2. Genetic Toxicity Statistical Suite

Most of the models in the Genetic Toxicity Suite are intended to support regulatory decision-making processes. The RCA (Q)SAR models for prediction of genetic toxicity in terms of mutagenicity and clastogenicity, are based upon the DARS publications [10,16,17]. Genetic toxicity was modeled for study calls, where the positive calls are trained as binary 1 and negative calls as binary 0. The outcome of a (Q)SAR prediction is given as the probability of being genotoxic on a scale of 0 to 1. The lower the probability value, the lower the genotoxic potential of a test chemical. The level of concern for test chemical genotoxic potential is evaluated in proportion to the number of sets of models with positive predictions.

This suite consists of modules of RCA (Q)SAR models for predicting gene mutation, and clastogenicity.

Some models in this suite are not used as supporting information for decision-making processes involved in the regulatory workflow. Rather, they support R&D for assessing the genotoxicant potential of organic chemicals. However, the same modeling process along with the training set preparation based on weight of evidence approach were applied as described in the DARS publication [10,14,16,17].

There are several RCA (Q)SAR models for predicting additional genotoxicity endpoints. They include *in vitro* chromosome aberrations, *in vitro* sister chromatid exchange, and Mouse Lymphoma mutagenicity.

As of version 1.6 of the Model Applier, the Salmonella Mutagenicity model was updated to version 3.0[27]. This update included additional mutagenicity data and data added for 445 new compounds harvested from FDA approval packages and the published literature, to give a total of 3974 compounds. 247 drug molecules marketed between 1970 and 2011 were added. Data gaps within the training set were identified using structural features derived from known toxicophores [18] and new compounds were added to the training set to fill those gaps.

As of version 1.7 of the Model Applier, the Salmonella Mutagenicity model, the previous E. coli (Q)SAR models were removed and replaced with a new, larger model of 1199 compounds containing either *Escherichia coli* WP2 uvrA, *Escherichia coli*, WP2 uvrA (pKM101) or *S. typhimurium* TA102 strains. This model, called the *in vitro* E Coli - Sal 102 A-T Mutagenicity model is designed to detect a variety of oxidants and other mutagenic carcinogens which modify A-T (adenine-thymine) base pairs [19].

As of version 2.0, the *in vitro* chromosome aberration models for cell lines other than CHO and CHL have been discontinued. The CHO and CHL training sets have been updated and the models improved. The criteria set by OECD 473 were used to govern the data selected for model construction. Datasets were well-balanced, with 53% and 45% positives, respectively, and contained re-evaluated legacy data as well as new data to expand the chemical space of previous models.

As of version 2.1, the *in vivo* micronucleus mouse training set has been updated and the model improved. Legacy data was re-evaluated and new data added to expand the chemical space of previous models [33]. An average model was built to overcome the deficiency of having few positives in the dataset resulting in better sensitivity and more balanced performance statistics. Citations for the source of training set data has now been included as hyperlink references within the study results. The model for overall rodent *in vivo* micronucleus has been removed.

As of version 2.1, the training set for the mouse lymphoma model was updated, re-evaluated and regraded according to the most current criteria as described by the International Workshop on Genotoxicity Testing in 2005 [34]. Separate training sets and models have now been created for both activated and inactivated endpoints. The updated models have better sensitivity and more balanced performance statistics. Citations for the source of training set data has now been included as hyperlink references within the study results.

The number of molecules in the individual models and their predictive performance are presented in the tables below.

Table 2-a. Summary of Statistics for Gene Mutation models

Endpoint Models	# training compounds	Sensitivity (%)	Specificity (%)
in vitro Salmonella	3974	77.0	87.8
in vitro E Coli - Sal 102 A-T Mut	1199	72.9	87.7
in vivo mammalian	213	62.7	88.5
in vivo mammalian dominant lethal	182	61.5	90.6
in vitro CHO V79 hgprt	643	46..5	92.7
Mouse Lymphoma Activated 5178Y-tk†	674	75.2	76.3
Mouse Lymphoma Unactivated 5178Y-tk†	750	79.0	73.8

† These models are used only in R&D purposes.

Table 2-b. Summary of Statistics for *in vivo* Clastogenicity Models

Endpoint Models	# training compounds	Sensitivity (%)	Specificity (%)
Micronucleus *in vivo* mouse	924	74.8	76.3
Chromosome Aberrations *in vivo*	285	48.0	91.4
Chromosome Aberrations *in vivo* rat	110*	6.67	96.8
Chromosome Aberrations: *in vivo* Other Rodent	153	48.1	86.9

* Only 13 % of the training set was positive.

Table 2-c. Summary of Statistics for *in vitro* Clastogenicity Models

Models	# training compounds	Sensitivity (%)	Specificity (%)
in vitro chrom. ab. CHL†	874	80.0	73.9
in vitro chrom. ab. CHO†	819	66.9	77.4
SCE *in vitro*†	758	71.0	72.2
SCE *in vitro* CHO†	624	87.7	42.4
SCE *in vitro* other cells†	204*	96.0	38.7

* Training dataset has 85% positive chemicals.

† These models are used only for R&D purposes.

B-3. Reproductive Toxicity Statistical Suite

The models in the Reproductive Toxicity Suite are intended to support regulatory decision-making processes. The RCA (Q)SAR models for prediction of reproductive toxicity in male and female rodents are based upon the DARS publications [16, 17].

Reproductive toxicity was modeled for study calls (e.g., male mouse), where the positive calls are trained as binary 1 and negative calls as binary 0. The outcome of a (Q)SAR prediction is given as the probability of being reproductive toxicant on a scale of 0 to 1. The lower the probability value, the lower the potential toxicity of a test chemical. The level of concern for test chemical reproductive toxicant potential is evaluated in proportion to the number of sets of models with positive predictions.

This suite consists of models of RCA (Q)SAR models for predicting reproductive toxicity in male and female rats and mice. All of these models were rebuilt in version 2.1 to provide for better sensitivity and balanced performance statistics. The number of molecules in the individual models and their cross-validated predictive performance are presented in Table 3.

Table 3. Summary of Statistics for Reproductive Toxicity Models

Models	# training compounds	Sensitivity (%)	Specificity (%)
Repro Rat Male	714	84.7	72.5
Repro Mouse Male	146	84.5	77.3
Repro Rat Female	894	61.1	95.3
Repro Mouse Female	151	85.4	69.9
Sperm Rat	723	71.8	80.1
Sperm Mouse	261	73.2	77.7

B-4. Developmental Toxicity Statistical Suite

The models in the Developmental Toxicity Suite are intended to support regulatory decision-making processes. The RCA (Q)SAR models of developmental toxicity of the rodent fetus include dysmorphogenesis (structural and visceral birth defects), developmental toxicity (fetal growth retardation and weight decrease), and fetal survival (fetal death, post-implantation loss, and pre-implantation loss). The methods and data sources are described in DARS publications [16,17].

Developmental toxicity was modeled for study calls (e.g., dysmorphogenesis of rat), where the positive calls are trained as binary 1 and negative calls as binary 0. The outcome of a (Q)SAR

prediction is given as the probability of being developmental toxicant on a scale of 0 to 1. The lower the probability value, the lower the potential toxicity of a test chemical. The level of concern for test chemical developmental toxicant potential is evaluated in proportion to the number of sets of models with positive predictions.

This suite contains RCA (Q)SAR models for structural dysmorphogenesis, visceral dysmorphogenesis, fetal survival, and fetal growth.. The number of molecules in the individual models and their predictive performance are presented in tables below.

Table 4-a. Summary of Statistics of Structural Dysmorphogenesis Models

Models	# training compounds	Sensitivity (%)	Specificity (%)
Structural Dysmorphogenesis Rodent	2019	28.6	94.4
Structural Dysmorphogenesis Rat	1759	43.4	89.4
Structural Dysmorphogenesis Mouse	979	34.6	90.5
Structural Dysmorphogenesis Rabbit	1014	50.4	88.8

Table 4-b. Summary of Statistics of Visceral Dysmorphogenesis Models

Models	# training compounds	Sensitivity (%)	Specificity (%)
Visceral Dysmorphogenesis Rodent	2019	36.2	91.1
Visceral Dysmorphogenesis Rat	1654	42.3	90.9
Visceral Dysmorphogenesis Mouse	978	47.1	88.6

Table 4-c. Summary of Statistics of Fetal Growth Models

Models	# training compounds	Sensitivity (%)	Specificity (%)
Fetal Growth Retardation Rodent	2019	22.1	92.6
Fetal Growth Retardation Rat	1759	33.3	89.8
Fetal Growth Retardation Mouse	978	40.4	90.2
Fetal Growth Retardation Rabbit	1013	38.8	89.0
Fetal Weight Decrease Rodent	2019	30.8	91.8
Fetal Weight Decrease Rat	1759	36.7	89.8

Fetal Weight Decrease Mouse	978	43.9	91.0
Fetal Weight Decrease Rabbit	1013	41.1	90.6

Table 4-d. Summary of Statistics of Fetal Survival Models

Models	# training compounds	Sensitivity (%)	Specificity (%)
Fetal Death Rodent	2019	28.6	90.6
Fetal Death Rat	1759	28.9	91.8
Fetal Death Mouse	978	36.9	90.4
Fetal Death Rabbit	1013	42.9	89.5
Post Implantation Loss Rodent	2019	30.9	92.5
Post Implantation Loss Rat	1759	31.4	90.2
Post Implantation Loss Mouse	978	28.3	92.6
Post Implantation Loss Rabbit	1013	49.0	86.2
Pre Implantation Loss Rodent	2019	32.3	90.6
Pre Implantation Loss Rat	1759	38.7	89.0
Pre Implantation Loss Mouse	978	51.2	90.0
Pre Implantation Loss Rabbit	1013	48.9	88.4

B-5. Neurotoxicity Statistical Suite

The models in the Neurotoxicity Suite intend to support regulatory decision-making processes. The RCA (Q)SAR models for neurotoxicity include prediction of new born behaviors. The methods and data sources are described in the DARS publications [16,17].

Neurotoxicity was modeled for study calls, where the positive calls are trained as binary 1 and negative calls as binary 0. The outcome of a (Q)SAR prediction is given as the probability of being neuro toxicant on a scale of 0 to 1. The lower the probability value, the lower the potential toxicity of a test chemical. The level of concern for test chemical neurotoxicant potential is evaluated in proportion to the number of sets of models with positive predictions.

This suite contains RCA (Q)SAR behavioral toxicity models of newborn rodent, rat, and mouse. The number of molecules in the individual models and their predictive performance are presented in Table 5.

Table 5. Summary of Statistics of Behavioral Toxicity Models

Models	# training compounds	Sensitivity (%)	Specificity (%)
Behavioral Toxicity Newborn Rodent	671	60.1	88.8
Behavioral Toxicity Newborn Rat	628	57.4	91.4
Behavioral Toxicity Newborn Mouse	172	75.7	88.1

B-6. Human Adverse Cardiological Effects Statistical Suite

The models in the Human Adverse Cardiological Suite intend to support regulatory decision-making processes. The RCA (Q)SAR models for prediction of human adverse effects of pharmaceutical chemicals are based on the DARS publications [19-22. Data were obtained from the FDA's post market surveillance AERS (Adverse Event Reporting System) and SRS (Spontaneous Reporting System) databases and the literature. Methodology of developing the human clinical endpoints based on data from these reports is published by DARS group [20].

Human adverse cardiological effects were modeled for summarized effects, where the positive effects are trained as binary 1 and negative as binary 0. The outcome of a (Q)SAR prediction is given as the probability of being cardiac toxicant on a scale of 0 to 1. The lower the probability value, the lower the potential toxicity of a test chemical. The level of concern for test chemical cardiac toxicant is evaluated in proportion to the number of sets of models with positive predictions.

In this suite, there are RCA (Q)SAR models for predicting several cardiac endpoints, including: conduction disorders, coronary artery disorders, electrocardiogram disorders, heart failure disorders, arrhythmia disorders, bradycardia disorders, QT prolongation, tachycardia disorders, torsades, myocardial infarct disorders, myocardial disorders, palpitations, and rate rhythm disorders. The number of molecules in the individual models and their predictive performance are presented in Table 6.

Table 6. Summary of Statistics of Human Cardiological Effects Models

Models	# training compounds	Sensitivity (%)	Specificity (%)
Conduction Disorders	1628	61.7	90.4
Coronary Artery Disorders	1628	52.9	89.1
Electrocardiogram Disorders	1628	51.2	87.8
Heart Failure Disorders	1628	46.3	91.2
Arrhythmia Disorders	1509	50.0	91.5
Bradycardia Disorders	1628	59.3	89.3

QT prolongation	1628	61.3	88.5
Tachycardia Disorders	1628	60.3	87.7
Torsades	1628	64.0	87.5
Myocardial Infarct Disorders	1628	61.7	88.3
Myocardial Disorders	1629	47.4	87.9
Palpitations	1628	58.2	88.0
Rate Rhythm Disorders	1628	39.1	90.1

B-7. Human Adverse Hepatobiliary Effects Statistical Suite

The models in the Human Adverse Hepatobiliary Suite intend to support regulatory decision-making processes. The RCA (Q)SAR models for prediction of human adverse effects of pharmaceutical chemicals are based on the DARS publications [20]. Data were obtained from the FDA's post market surveillance AERS (Adverse Event Reporting System) and SRS (Spontaneous Reporting System) databases and the literature.

Drug-induced liver injury (DILI) is one of the most common drug-induced adverse events (AEs) leading to life-threatening conditions such as acute liver failure. It is also the second most common cause for post-market withdrawals or warnings. Efforts to develop new predictive methods to assess the likelihood of a drug being a hepatotoxicant have been challenging due to the complexity and idiosyncrasy of clinical manifestations of DILI. The FDA adverse event reporting system (FAERS) contains post-market data that depict the morbidity of AEs.

The hepatotoxicity suite of models was updated in version 2.0 of the Leadscope Model Applier[28,29]. A new training set of 2029 unique and modelable drug entities with 13,689 drug-AE combinations was extracted from the AERS database using 38 hepatotoxicity-related query preferred terms. Additional filtering was performed to remove low confidence negative data and the resulting dataset of 1314 compounds, where the percentage of actives in the set was 50%, was used to build (Q)SAR for several endpoints. Part of this training set including descriptions for endpoint development is available from the DARS website [24]. Methodology of developing the human clinical endpoints based on data from these reports is published by DARS group [21].

Human adverse hepatobiliary effects are modeled for summarized effects, where the positive effects are trained as binary 1 and negative calls as binary 0. The outcome of a (Q)SAR prediction is given as the probability of being hepatobiliary toxicant on a scale of 0 to 1. The lower the probability value, the lower the potential toxicity of a test chemical. The level of concern for test chemical hepatobiliary toxicant potential is evaluated in proportion to the number of sets of models with positive predictions.

In this suite, there are RCA (Q)SAR models for predicting: bile duct disorders, cholestasis, liver acute damage, and liver enzyme release disorders. The number of molecules in the individual models and their predictive performance are presented in Table 7.

Table 7. Summary of Statistics of Human Hepatobiliary Models

Models	# training compounds	Sensitivity (%)	Specificity (%)
Bile Duct Disorders	1017	75.9	86.6
Cholestasis	1124	74.8	76.9
Liver Acute Damage	1314	73.2	66.1
Liver Enzyme Release Disorders	1134	72.6	76.1

B-8. Human Adverse Urinary Tract Effects Statistical Suite

The models in the Human Adverse Urinary Tract Suite intend to support regulatory decision-making processes. The (Q)SAR models for prediction of human adverse effects of pharmaceutical chemicals are based on the DARS publications [20,22]. Data were obtained from the FDA's post market surveillance AERS (Adverse Event Reporting System) and SRS (Spontaneous Reporting System) databases and the literature. Methodology of developing the human clinical endpoints based on data from these reports is published by DARS group [20-22].

Human adverse urinary tract effects are modeled summarized effects, where the positive effects are trained as binary 1 and negative effects as binary 0. The outcome of a (Q)SAR prediction is given as the probability of being urinary tract toxicant on a scale of 0 to 1. The lower the probability value, the lower the potential toxicity of a test chemical. The level of concern for test chemical urinary tract toxicant potential is evaluated in proportion to the number of sets of models with positive predictions.

In this suite, there are models for predicting: bladder disorders, blood in urine, kidney disorders, kidney function tests, nephropathy disorders and urolithiasis disorders. The number of molecules in the individual models and their predictive performance are presented in Table 8.

Table 8. Summary of Statistics of Human Urinary Tract Models

Models	# training compounds	Sensitivity (%)	Specificity (%)
Bladder Disorders	1591	51.5	89.7
Blood in Urine Disorders	1591	49.7	94.3
Kidney Disorders	1590	36.8	95.8
Kidney Function tests	1589	48.9	89.9
Nephropathy Disorders	1590	52.9	90.8
Urolithiasis Disorders	1591	42.1	94.9

B-9. Genetox Expert Alerts Suite

The Leadscope Genetox Expert Alerts have been implemented as part of the Leadscope Model Applier as of version 1.8 (alongside the existing statistical-based (Q)SAR models). To develop of this system, an initial library of mutagenicity structural alerts was identified from the literature. This process included consolidating the same or similar alerts cited in multiple publications. Information on plausible mechanisms was collected alongside the structural definitions. Factors that deactivate the alerts were also identified from the literature and through an analysis of the corresponding data using the Leadscope data mining software. Over 200 distinct alerts were identified and these alerts were further validated against a reference database of about 10,000 chemicals with known bacterial mutagenesis results. Only validated alerts with a sufficiently strong association with positive expert-reviewed calls from Salmonella and E. coli strains were included. A prediction of the bacterial mutagenesis assay can be made using these validated alerts; however, this is only possible when the compound is within the applicability domain of the alert system. In addition, a confidence score based upon information collected for each alert is provided alongside the positive or negative call.

As of version 2.0 of the Leadscope Model, the Leadscope Genetox Expert Alerts have been updated (to version 2) to include knowledge shared from corporate sponsors. The alerts have been further refined to include additional corporate knowledge in versions 3.0 and 4.0

Leadscope has established a knowledge-sharing program with interested corporate sponsors to address specific (Q)SAR regulatory issues identified through discussions with sponsors and regulatory agencies. The initiative allows the use of proprietary corporate information to be investigated under confidentiality restrictions and identify potential solutions to specific predictivity issues or increase the number of compounds which can be predicted.

Knowledge from proprietary data has also been used by Leadscope to increase both the sensitivity and specificity for selected chemical classes. In version 4 of the alerts, 37 new alerts have been added and 27 alerts modified to support updates when predicting classes of: alkyl halides, methyl halides, aromatic nitros, aromatic amines, halo-amines, aromatic amides, hydrazines, polycyclic aromatics, fluorenes, and boronic acids. Mitigating factors were also identified from fingerprint data and corporate data contributed to the public sector and were used to reduce false positives of common chemical starting materials and reagents.

The process for knowledge sharing was through the use of structural fingerprints for several compound classes. Several thousand chemical fingerprints based on the Leadscope fragment hierarchy, data analysis of a reference set and external knowledge containing a variety of primary aromatic amines were derived. The list of substructures includes meta-, para-, ortho-, hetero-substituted, polycyclic, as well as more complex substitution patterns. When these fingerprints were applied to a proprietary data set by the data owner(s) the result is a listing of named substructures present along with the number of positive and negative bacterial mutagenicity examples which are present. Only the results for the pre-defined substructures in the fingerprint are summarized and it is therefore possible to apply these fingerprints to a

proprietary database without revealing information for individual compounds or data. This project now includes over 13 pharmaceutical companies and regulatory agencies and has resulted in continued improvement in the performance around primary aromatic amines. This fingerprint methodology is also being applied to other chemical classes including boronic acids and alkyl halides.

The responsiveness statistics of bacterial mutation tester strains (and strain combinations) have been updated for the alerts reflecting updates to the alert definitions and Alert Reference Set data.

To assess the performance of this alert system, two data sets were used: (1) the reference set (as described in Reference set section), and (2) the Hansen data set. The Hansen set includes data described in the Hansen et al. publication [25] where the full set includes 6,512 chemicals. Those in the RCA-(Q)SAR training set were removed [18]. A number of other chemicals were also removed based on stereo-chemical considerations or their inability to be modelled leaving 3,903 compounds. Table 9 shows the performance results for the reference set. Table 10 shows the corresponding performance statistics for the Hansen data set. For further information see the white paper [23].

Table 9: Expert Alert Performance (version 4) for the over 10,000 Compound Reference Set

Concordance	86%
Sensitivity	85%
Specificity	87%
Positive Predictivity	89%
Negative Predictivity	84%
Coverage	92%

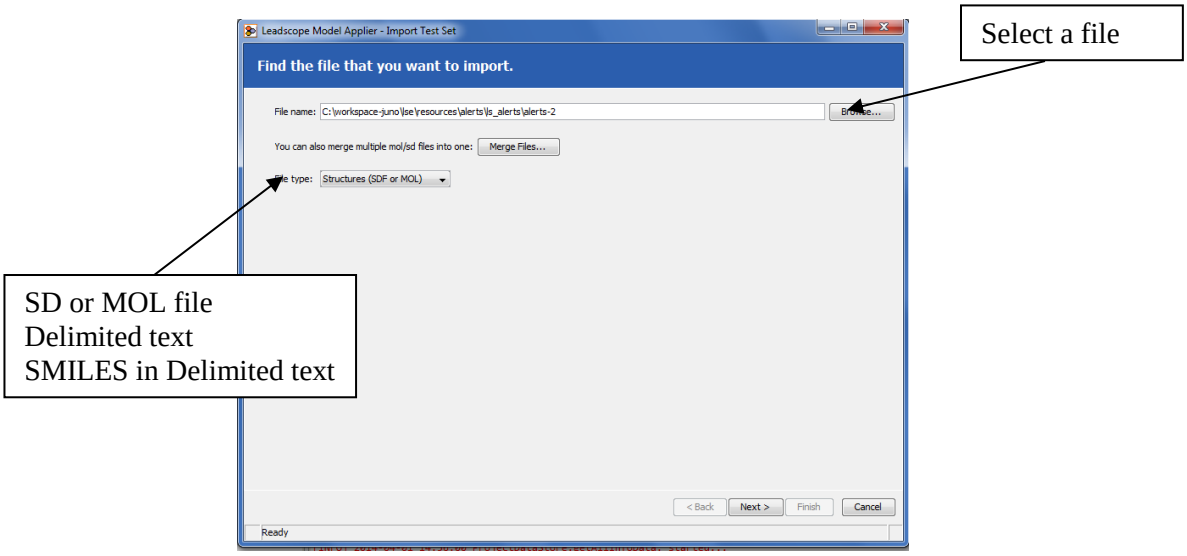
Table 10: Expert Alert Performance (version 4) for the Hansen Set

Concordance	83%
Sensitivity	92%
Specificity	70%
Positive Predictivity	81%
Negative Predictivity	85%
Coverage	94%

C. Model Applier

C-1. Importing a Test Set

Tests can be imported into the application as files or an entry for a single compound. The application handles both MDL MOL or SD files and SMILES.



Importing from a file can be accessed by clicking the "Import a test set" wizard button. The task wizard will go through "file chooser/review structure file (structure id can be specified)/name the test set/Finish". If the test set contains other data, the wizard will guide accordingly. After naming the test set, follow "choose other the data set /define the data type (numeric vs. text)/Finish". If the SD or MOL files do not contain the name or data field for a name, the application will automatically assign "Test-Structure –" prefix in a numeric order. As of version 1.8, multiple MOL files may be imported at the same time.

Importing a single structure from a file can be achieved from this wizard as well. The wizard also allows u*pdating an existing test set.*

Importing a single structure by entry is also allowed from the "Apply models" wizard (see C-4).

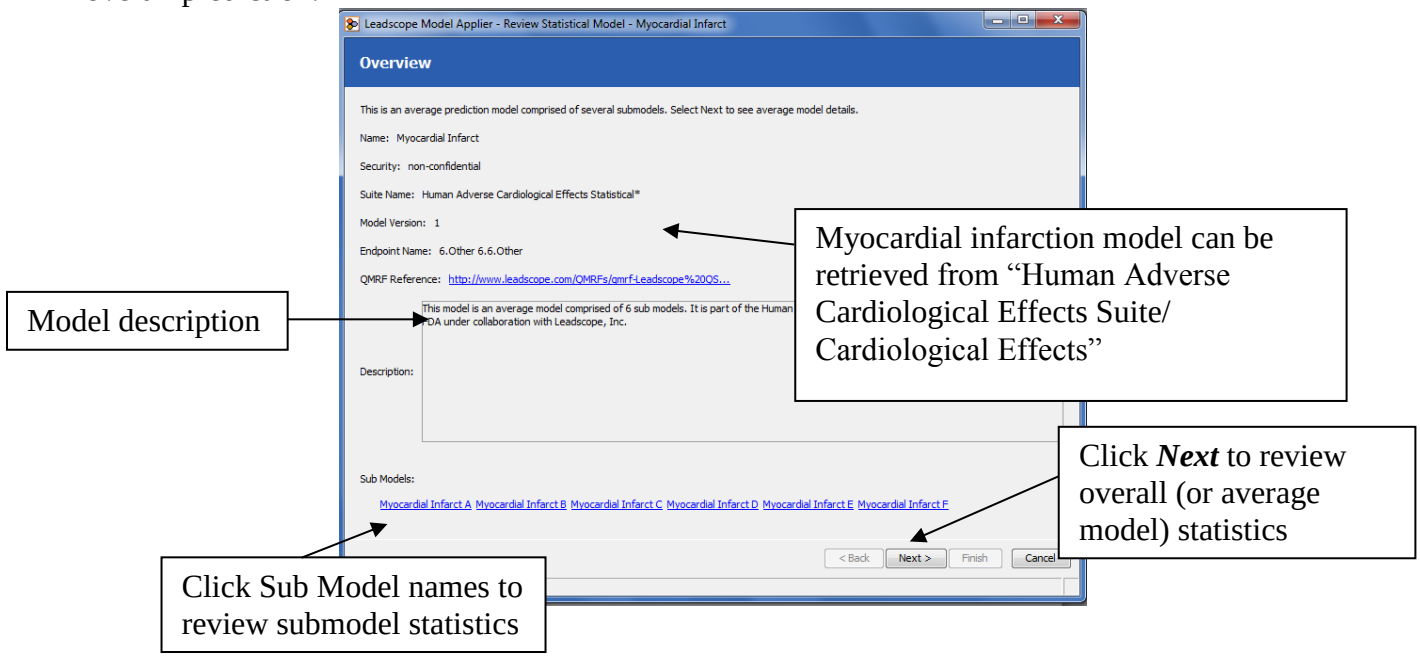
C-2. Managing Test Sets

This wizard provides ways to *browse* imported structures and data, *rename* or *delete* the test sets, and *delete prediction* results.

C-3. Reviewing Prediction Models and Expert Alerts

The RCA (Q)SAR models in this application can be accessed from this wizard. The models are organized hierarchically in a tree in the same order shown in the tables in Section B1-9. When clicking the model name, a short description will appear in the "description" box in the right pane.

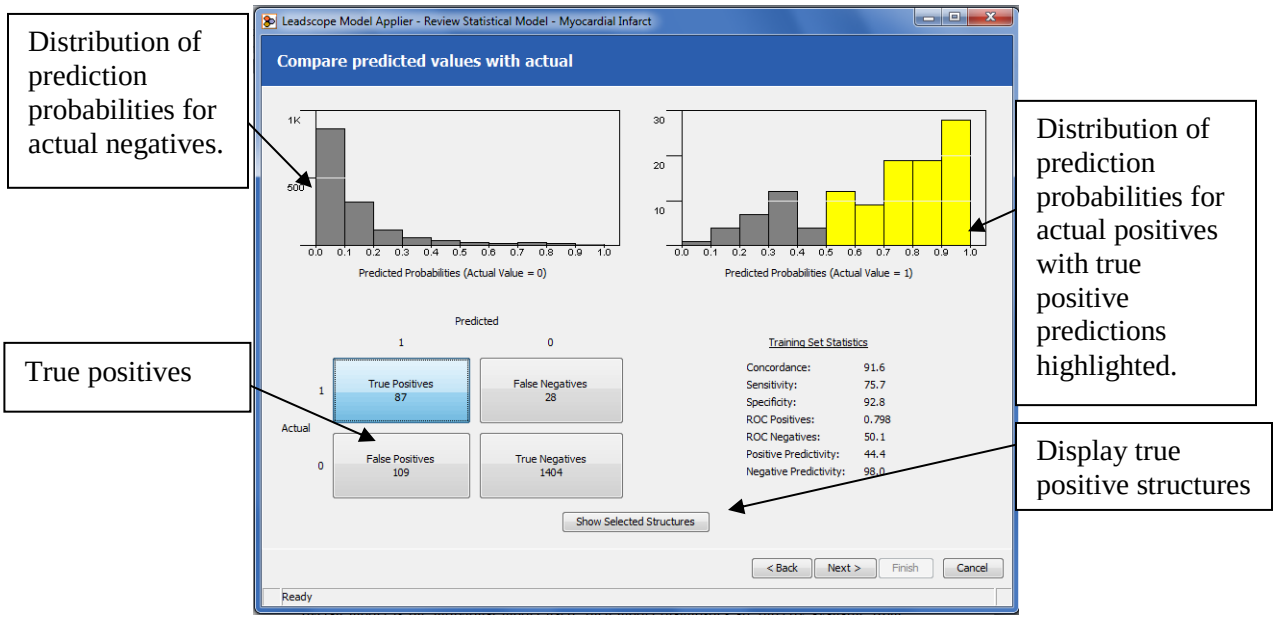
The wizard allows the review of both the overall average model as well as the sub-models. Again, the probabilities of a compound to be positive in each sub-model were averaged to calculate an overall prediction.



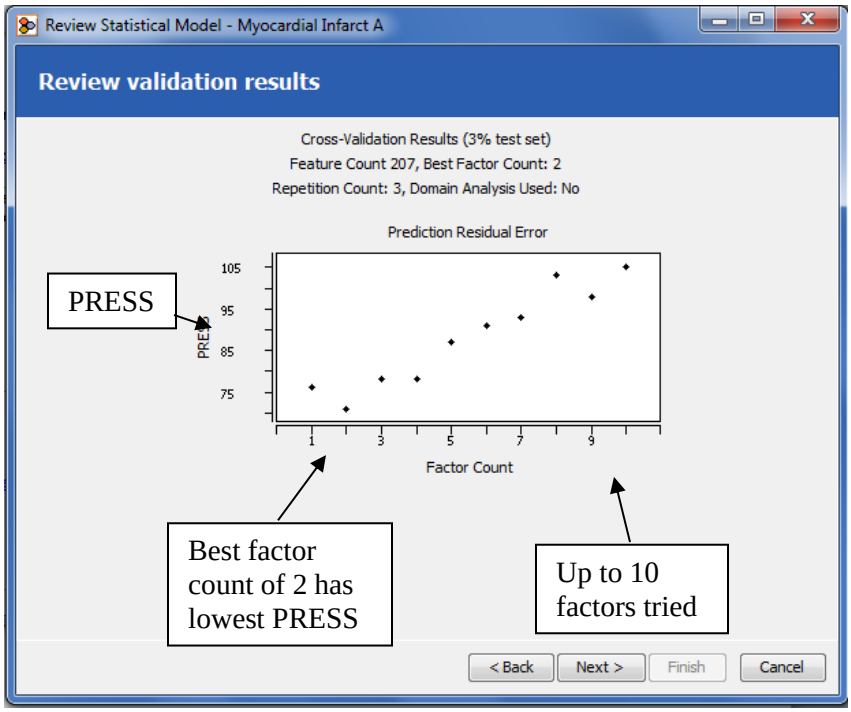
Clicking "Next" from this wizard will display the cross-validated model results for chemicals in the training set. For example, myocardial infarction model results are shown in the next figure below. Leadscope® calculates concordance (overall accuracy), sensitivity, specificity, receiver-operator-constant (ROC) for positives and negatives.

The graphs below indicate how well the structures with experimental values of zero (on the left) and one (on the right) were predicted. In this example, the true positives were selected, which highlights the number of structures with a predicted positive probability of >= 0.5 in the graph on

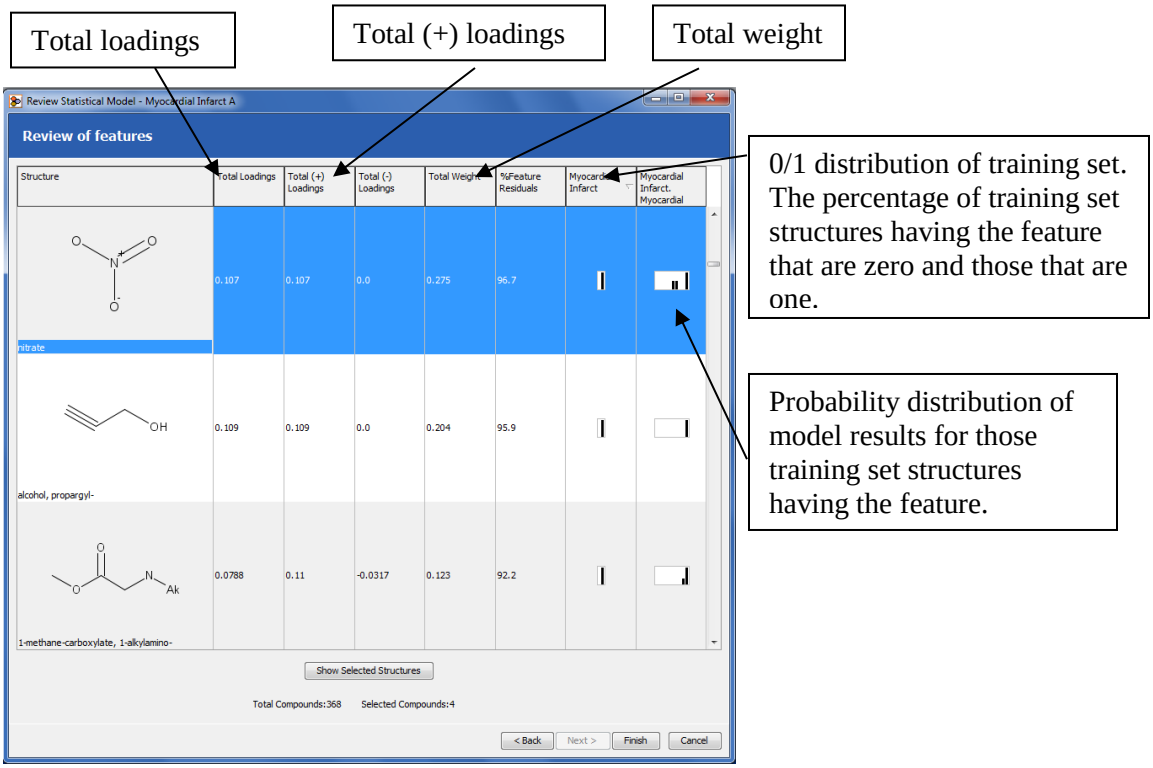
the right. The "Show Selected Structures" button will then display the training set structures and experimental and cross-validated positive prediction probabilities (e.g. toxic) in a table.



Model features and significance can be reviewed for each individual model. If a model is an average of several individual sub-models, the action has to be invoked from the sub-models option in the first panel of the "Review Model Results" wizard as shown on page 21. If the overall model is the individual model itself, then model diagnostics are directly available from the overall model. For example, following the Sub-model A link in the myocardial infarction model leads to a table of prediction results table (actual values, predicted values, and number of compounds in the local neighbor for the test chemical within 60% similarity), followed by the statistics and feature significance wizard panels. The model of this sub-model A for the myocardial infarction model used 207 structural features and 2 PLS factors. This cross-validation was performed at 3%. The graph of PRESS (prediction residual sum of squares) vs. Feature Count is presented, indicating the reason for the selection of 2 PLS factors during the optimization process.



The 207 structural features are then presented in the next screen. The features can be sorted by loadings, which is analogous to a regression coefficient of linear regressions.



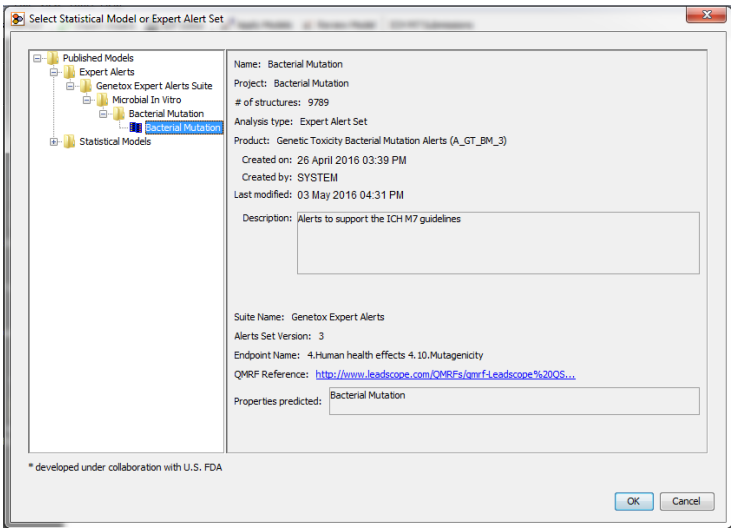
In this model, three highly-weighted features, with high total loadings, and total (+) loadings include: 1) nitrate, 2) propargyl alcohol and 3) 1-methane-carboxylate, 1-alkylamino-. The training set structures containing these features are all positive for this effect, as indicated by the 0/1 frequency distribution of the training set.

The scaffolds (represented as numbers in the feature table) are structural motives generated in Leadscope as "predictive scaffolds" based on FDA DARS training sets at the time of model building. Scaffolds are assembled from features in the Leadscope hierarchy and are optimized against the particular endpoint activity.

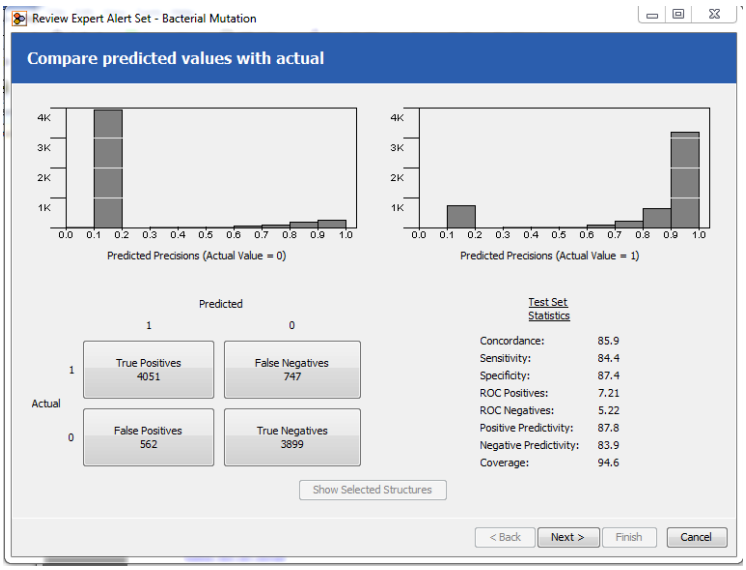
One or more row (i.e. features) may be selected and the structures in the training set matching them may be displayed using the "Show Selected Structures" button.

C-3b. Reviewing Expert Alerts

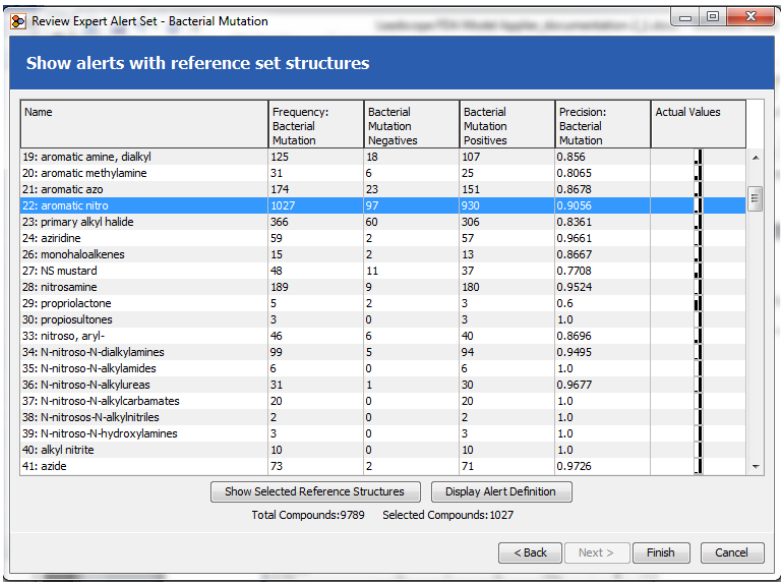
To review the Leadscope expert alerts, click on Endpoint name (e.g. Bacterial Mutation) and click OK.



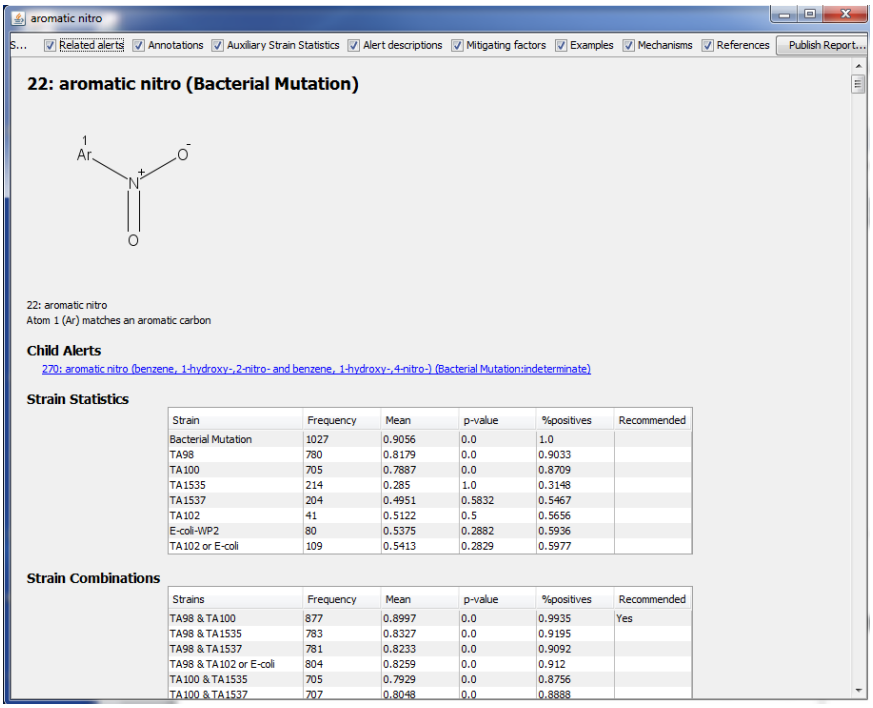
This next screen outlines the performance of the Alerts against the references set.



Next is a listing of all alerts along with information on the performance of each alert, including the number of positive and negative chemicals in the reference set.



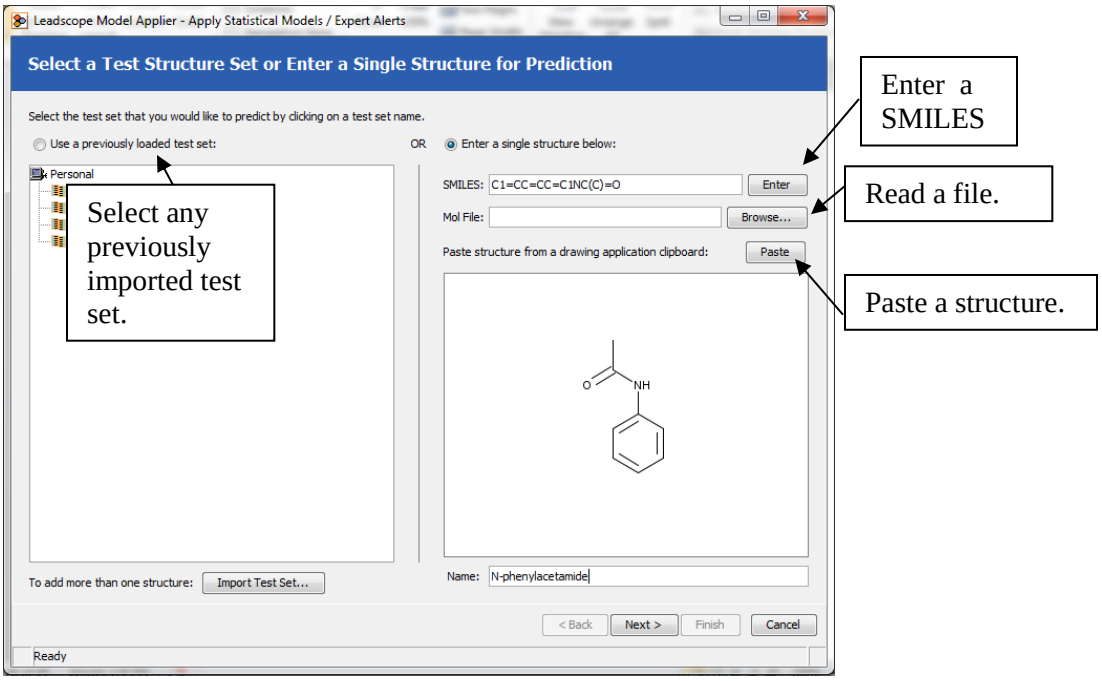
Double clicking on any of the rows will show the chemical structures matching the alert. Right clicking on a row and selecting the option "Display alert definition" will show a definition of the alert, including source, deactivating fragments, examples, mechanisms, and strain information.



As of version 2.1 of the Leadscope Model Applier, the alert definitions (version 3) included strain statistics indicating the repsonsiveness of individual and combination of strains for each alert. This includes a recommendation of strains to target for follow-up Ames testing (per alert) in cases where sufficient amount of test material is not available or is difficult to synthesize (per Note 2 of the M7 Guidelines) [35].

C-4. Applying Models and Expert Alerts

"Apply models for existing imported test set" can be accessed from this panel. Within this "Apply Statistical Models/Alerts" panel, it is possible to enter a single chemical by typing a SMILES string, copying/pasting, or browsing a mol file. Below, N-phenylacetamide was entered as a SMILES string.

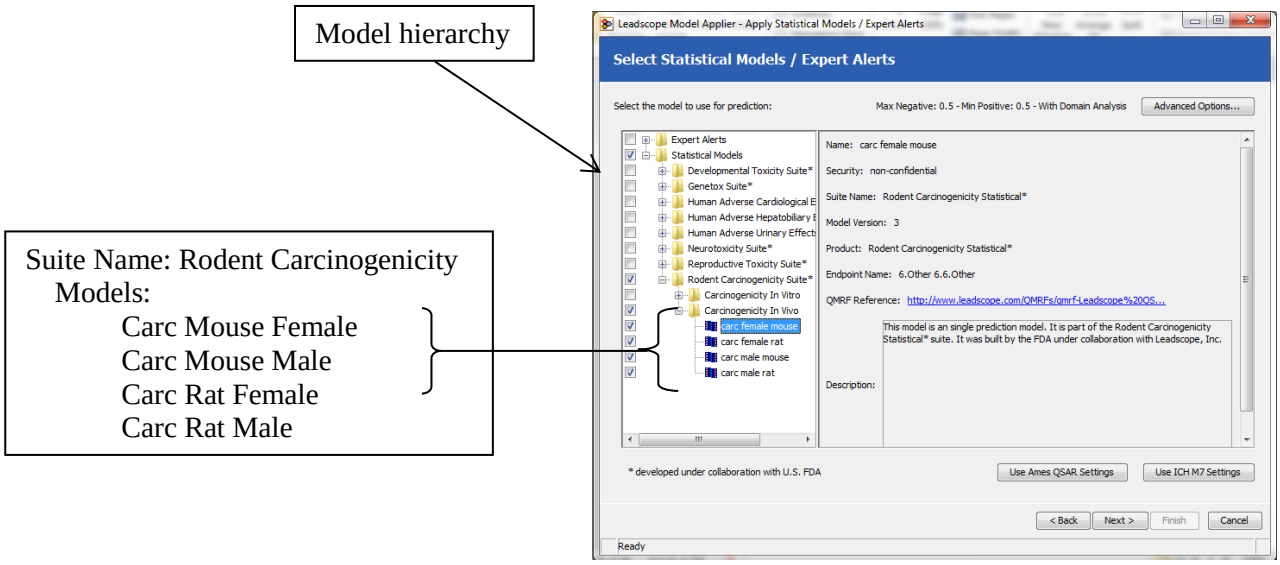


Note that SMILES strings with stereo-chemistry have their stereo-chemistry removed upon entry. The maximum length of a SMILES string is 4000 characters. Stereo-chemistry in MOL or SD files is preserved.

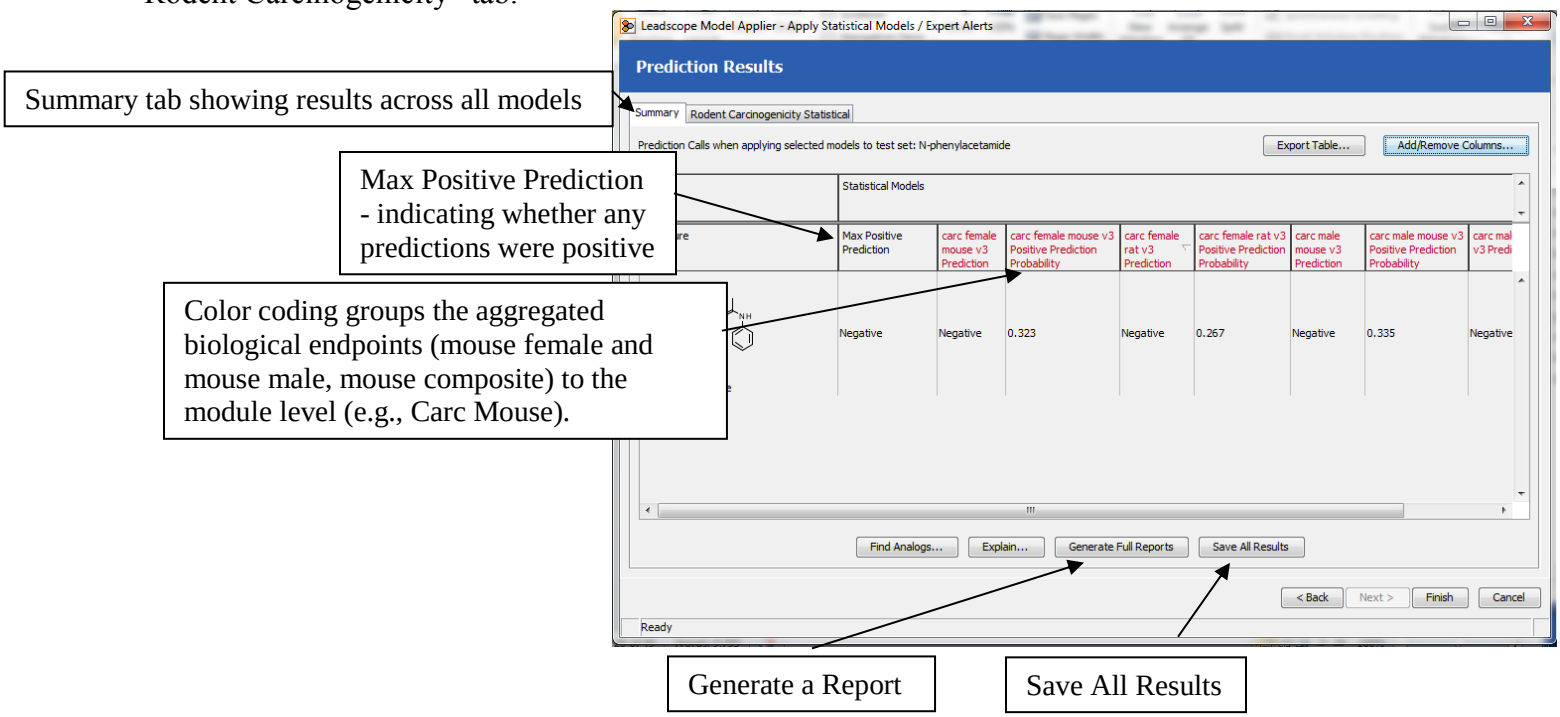
At the time of importing structures, Leadscope will save them in the application as a "test set". This process perceives of all the features and eight calculated properties associated with the chemical and stores the fingerprint and property values. When importing many structures this may by time-consuming. However, fingerprints are calculated only once and prior to applying any of the models. When models are applied only the parent form of the test compound is used for prediction; organic, inorganic salts and water fragments and fragment copies are not considered and the compound is neutralized as appropriate. See Appendix A for the list of organic salts removed.

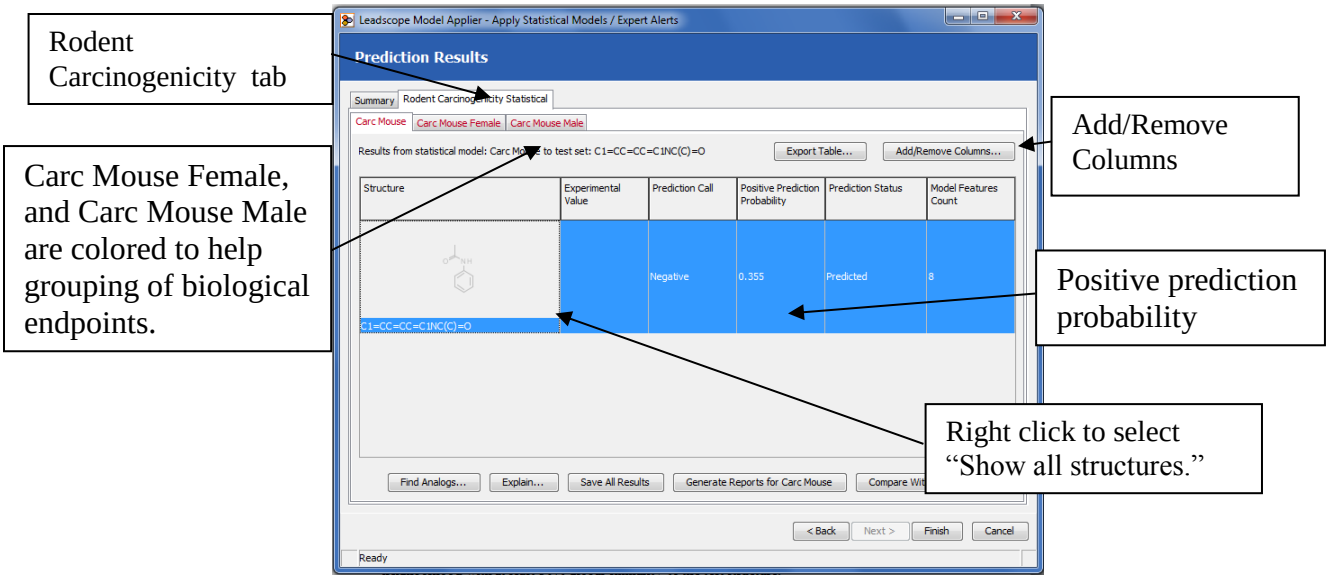
Within the application, running multiple models and expert alerts is possible and the results of the multiple models are displayed in the "Summary" tab of the prediction results table. As an example, mouse models for the RCA (Q)SAR rodent carcinogenicity suite are being illustrated here. In the mouse module, there are three models, i.e., mouse (composite), mouse female, and mouse male. All the models can be selected and run simultaneously, but the results will be presented in a separate tab as well as a summary. Each of the biological endpoint models can be

an overall model based on sub-models of each biological endpoint. Carc Mouse and Carc Rat are separate composite models of mouse female/male and rat female/male data, respectively. They are not averaged results of Carc Rat Female and Rat Male, or Carc Mouse Male/Female models.



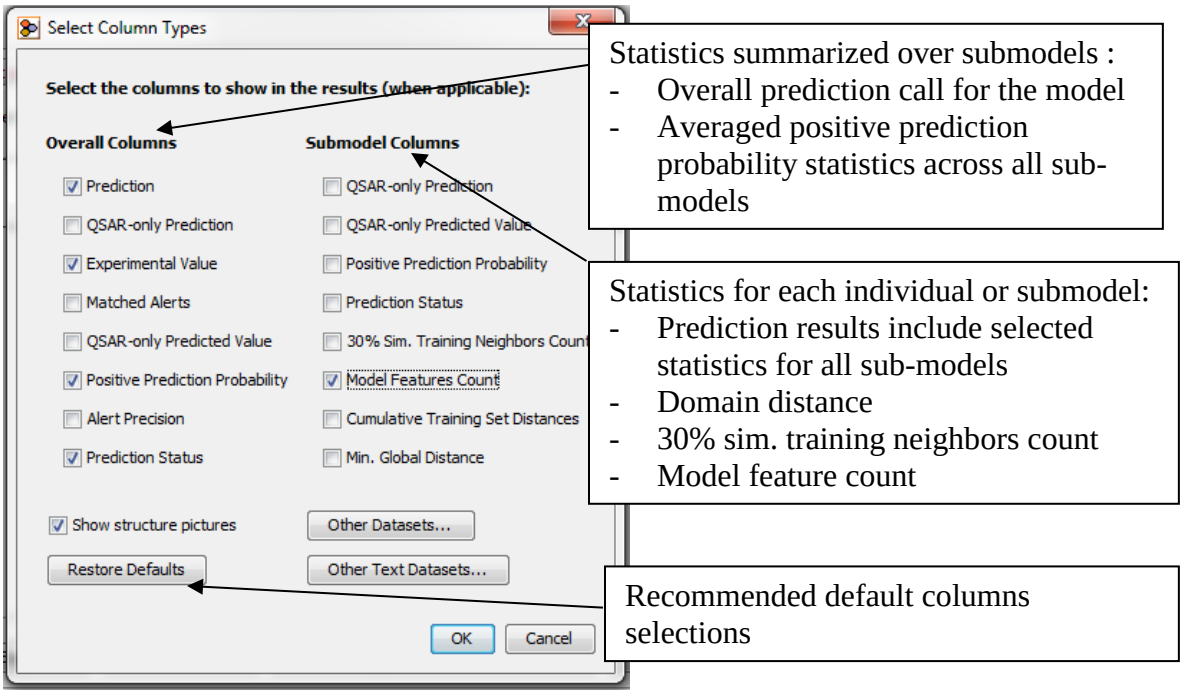
After applying Carc Mouse (composite), Carc Mouse Male, Carc Mouse Female models on N-phenylacetamide, the results are tabulated in a "Summary" table. This summary displays all the requested prediction results for each model as well as indicating whether any prediction calls for from any of the models were positive. Individual prediction results are available from the "Rodent Carcinogenicity" tab.





Prediction calls include positive, negative, not-in-domain or missing descriptors. In RCA models, any probability equal to or greater than 0.5 is considered positive. No predictions will be made if the test chemical is considered not-in-domain or when not all of the descriptors (such as ALogP) can be generated for the chemical structure. Leadscope® uses two parameters to guide the applicability of model domain: 1) having at least one structural feature defined in the model in addition to all the property descriptors; 2) having at least one chemical in a training neighborhood with at least 30% global similarity to the test structure.

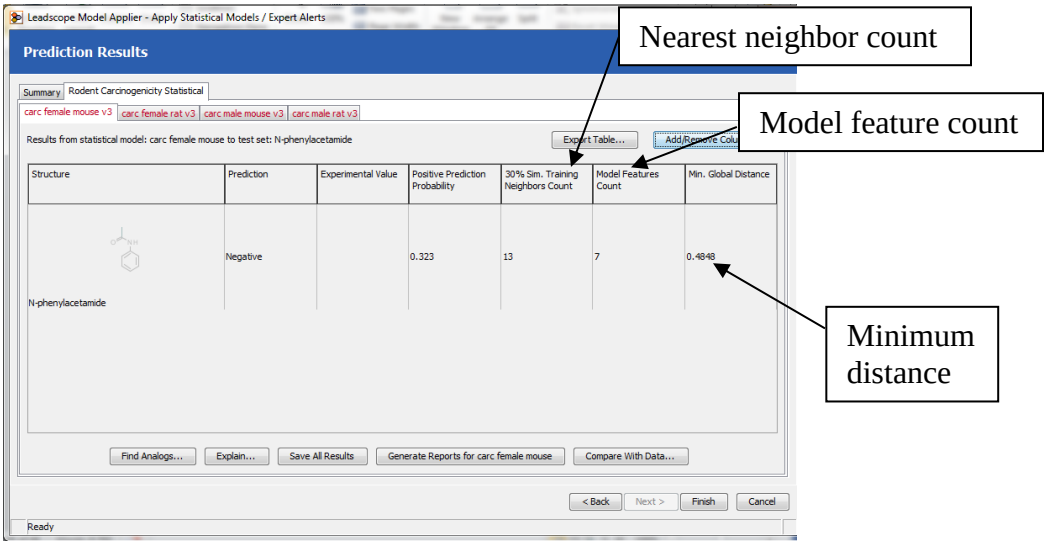
Additional model parameters can be retrieved by clicking the "Add/Remove Columns" button.



The overall Prediction column includes consideration of any experimental data found for the test compound. In this case, the experimental over-rides any predicted value and the experimental call is displayed in the table as a hyperlink. To view the predictions excluding experimental data select the QSAR-only columns from the column chooser.

When submodel columns are selected, the statistics in the spreadsheet for each submodel are grouped together. Different submodel columns have different color column titles to distinguish them apart. However, the specific color has no other meaning.

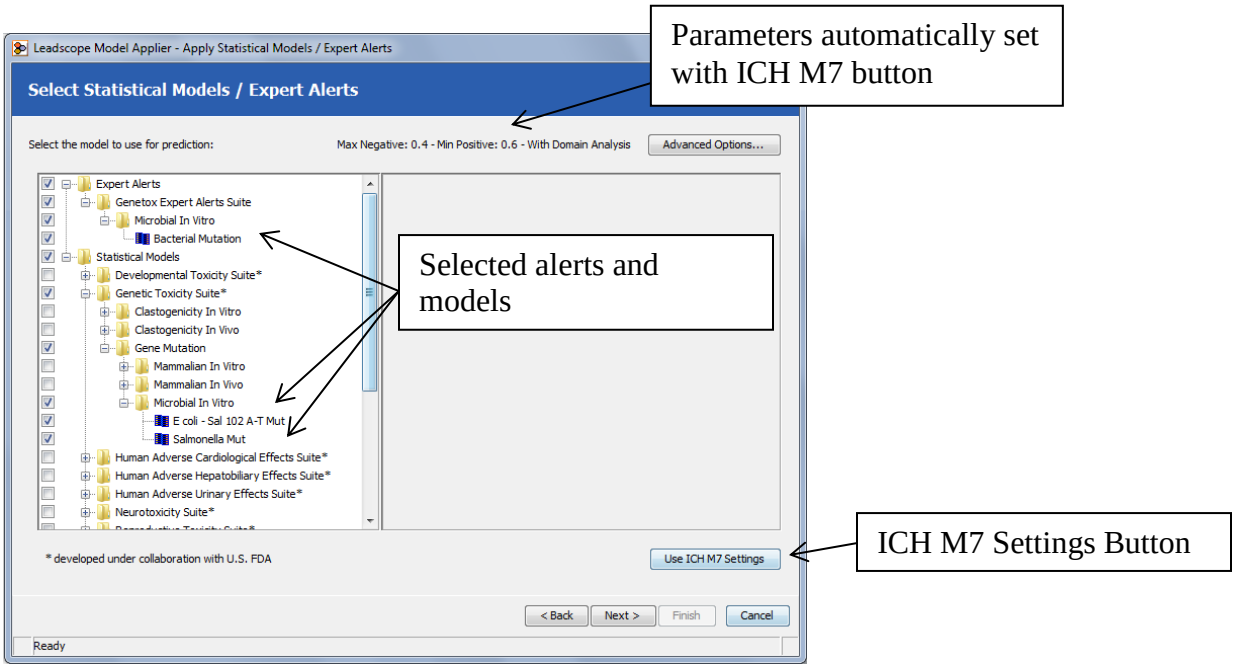
Within the carc mouse female model for N-phenyacetamide, 13 chemicals were found to be within the neighborhood defined by 30% similarity by Leadscope's definition. There were 7 model features used to calculate this prediction. The minimum distance of the test chemical to all the training set structures (based on all chemical features) is 0.4848.



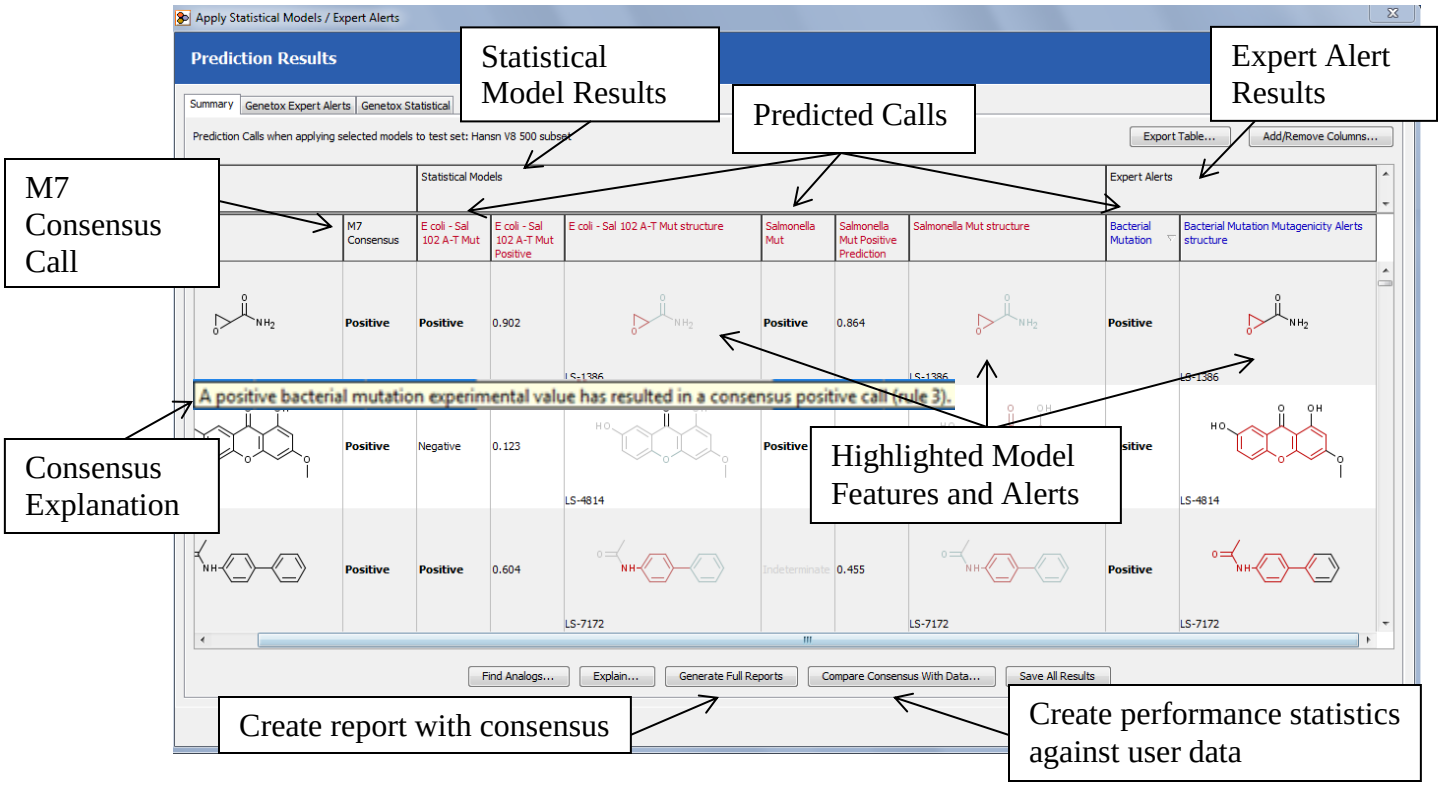
C-4b. Applying Expert Alerts

The Genetox Expert Alerts were added as of version 1.8 of the Leadscope Model Applier. The Expert Alerts are included in the Genetox Expert Alerts Suite under the Expert Alerts section of the hierarchy (versus the Statistical Models section). Expert Alerts may be selected for application to a test set in the same manner as statistical models. In fact, both may be applied at the same time to a given test set. Version 2.2 of the Leadscope Model Applier includes an update to the Genetox Expert Alerts Suite (version 4).

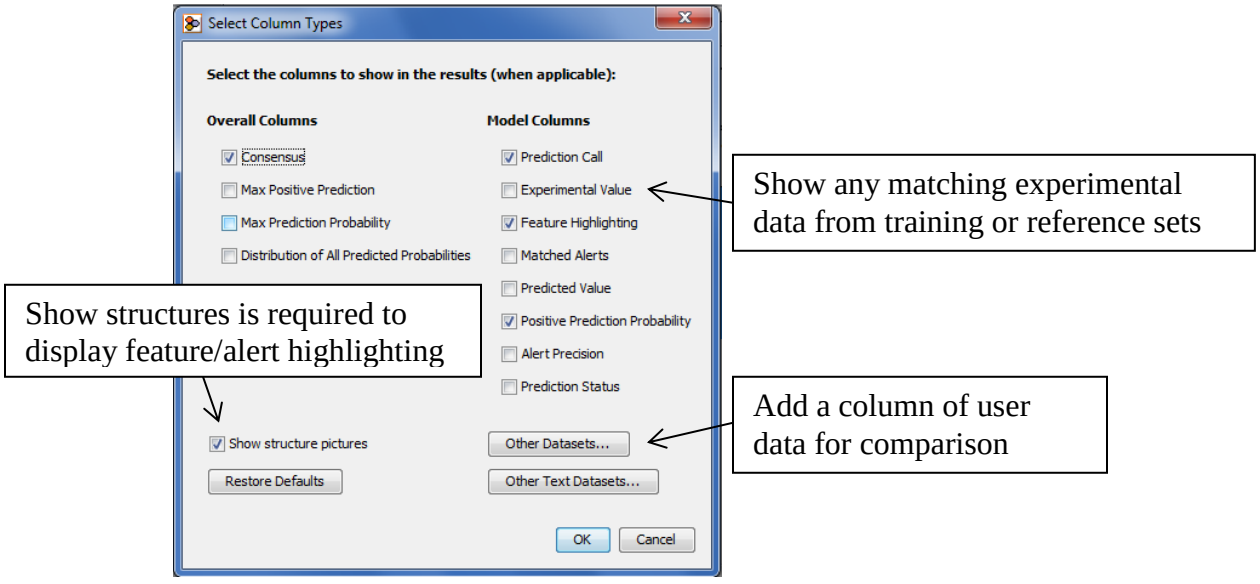
To help with ease of use, a "Use ICH M7 Settings" button has been created. This button will select all the appropriate expert alerts and statistical models necessary for analysis of impurities in accordance with the ICH M7 guidance on that subject as well as set any appropriate settings. In this example below, the expert alerts for Bacterial Mutation were selected as well as statistical models for Salmonella Mutagenicity and E. coli / Salmonella TA102 Mutagenicity.



The results when applying both alerts and statistical models are displayed as different tabs in the results. As previously shown, a summary view provides an overall picture of the results. However, now the statistical model and alerts are shown in different sections as two different methodologies were used in the creation of their results, respectively. In each section, results are displayed for experimental and predicted results. Highlighting of the test compounds with positive and negative models features is shown for the statistical models and alerts are highlighted in the Expert Alerts section. In the figure below we seen that the epoxide is highlighted as a pertinent model feature in both the Salmonella and E. coli model predictions (in the Statistical Models sections) and it is also found as an alert (in the Expert Alerts section). Predictions, alerts, and experimental data (not shown for brevity) are used together as evidence for creating an overall M7 consensus call displayed in the M7 Consensus column. This consensus is only shown if the ICH M7 button is used in running the prediction.



Reports containing the consensus call with reasoning may be generated using the Generate Full Reports button. Performance against a validation set can be calculated using the Compare Conensensus with Data button. Additional columns may be added or removed using the Add/Remove Columns button at the top right of the dialog. Note that experimental values found in the model training sets or alert reference set may be displayed as an individual column or part of the Prediction column. User data, may also be added as a column to the spreadsheet using the Other Datas sets button.



Hovering the mouse over the M7 Consensus call displays the reasoning used in calculating the consensus. This information is also displayed in the report and when selecting the Explain button from the Summary Sheet. The set of rules used in caclulating the consensus are shown below. Note that experimental data takes precedence over QSAR-only prediction calls. Positive experimental data results in a positive consensus. Negative experimental data can over-ride a positive prediction for a given endpoint. Lacking experimental data for a test compound, a predicted positive call for any model or alert results in a positive consensus prediction. Rules handling combinations of negative and out-of-domain (or indeterminate) predictions are more subtle and depend on the weight of evidence available. Sometimes there is not enough evidence to support a conensus prediction and the result is inconclusive. As of version 2.2 the consensus values for rules 12 and 14 where changed from negative to inconclusive (reflecting the cases when one technology was out-of-domain) and rules 14b, and 14c were added reflecting the cases when both model and alert predictions were negative.

The table or ordered rules used for creating the consensus prediction is shown below.

Rule	*Salmonella* G:C Statistical Prediction	*Salmonella* G:C Laboratory Result	*E. coli* / *Salmonella* A:T Statistical Prediction	*E. coli* / *Salmonella* A:T Laboratory Result	Bacterial Mutation Alert Identification	Bacterial Mutation Laboratory Result	Recommended Consensus Conclusion
10	-	-	-	-	-	Negative	Negative
1	-	Positive	-	-	-	N/A	Positive
2	-	Negative or N/A	-	Positive	-	N/A	Positive
3	-	Negative or N/A	-	Negative or N/A	-	Positive	Positive
4	-	Negative	-	Negative	-	N/A	Negative
8	-	Negative or N/A	-	N/A	Positive	N/A	Positive
11	-	N/A	-	Negative or N/A	Positive	N/A	Positive
6	Positive	N/A	-	Negative or N/A	-	N/A	Positive
7	-	Negative or N/A	Positive	N/A	-	N/A	Positive
5	-	Negative	-	Negative or N/A	-	N/A	Negative
9	-	N/A	-	Negative	Negative	N/A	Negative
13b	Out of Domain	N/A	Negative or Out of Domain	Negative	Out of Domain	N/A	Inconclusive
14b	Negative	N/A	Out of Domain	N/A	Negative	N/A	Negative
14c	Out of Domain	N/A	Negative	N/A	Negative	N/A	Negative
14	Out of Domain	N/A	Out of Domain	N/A	Negative	N/A	Inconclusive
12	Negative	N/A	-	Negative or N/A	Out of Domain	N/A	Inconclusive
13	Out of Domain	N/A	Negative	N/A	Out of Domain	N/A	Inconclusive
15	Out of Domain	N/A	Out of Domain	N/A	Out of Domain	N/A	Inconclusive

Positive predictions

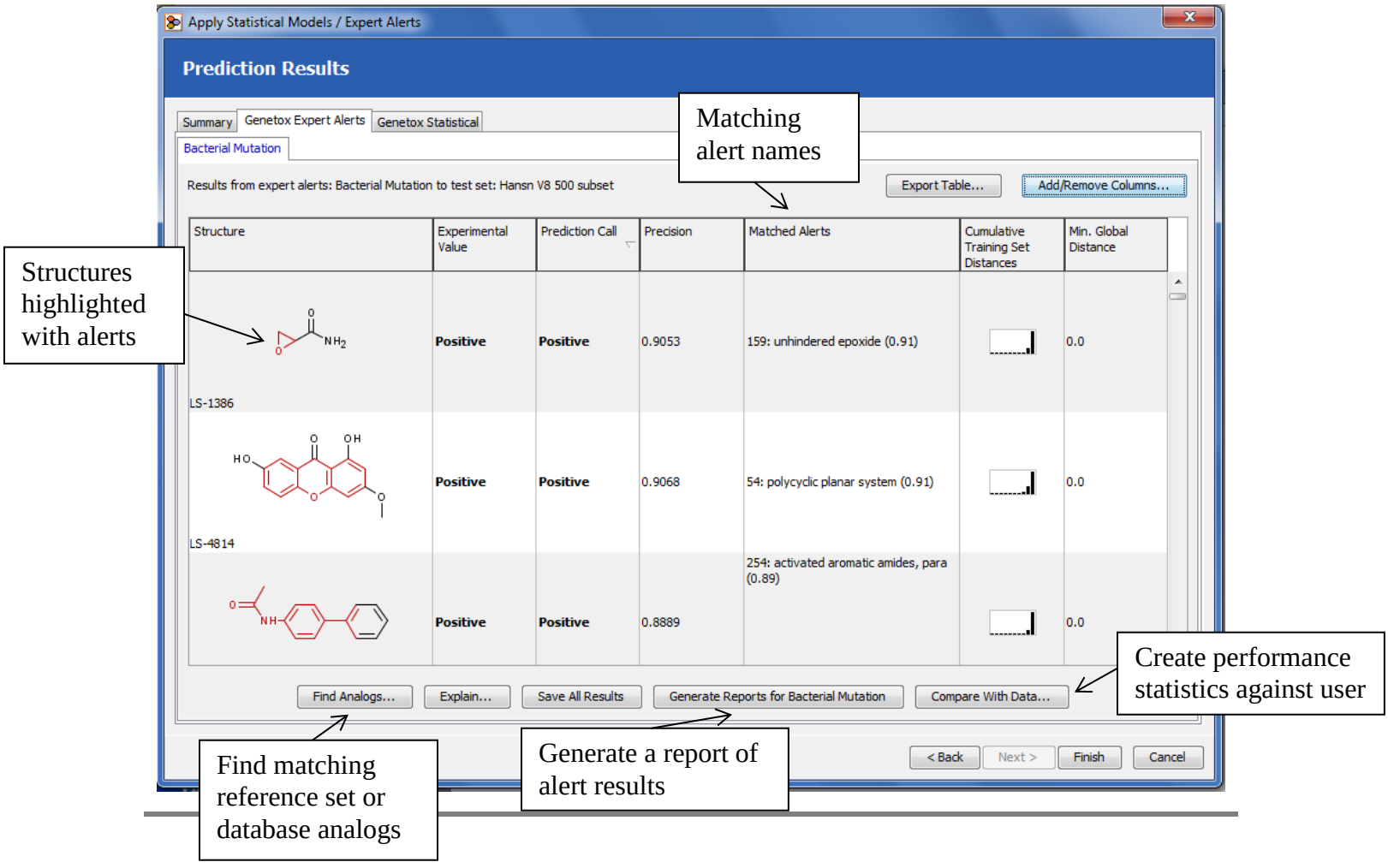
Negative predictions

Positive experimental data

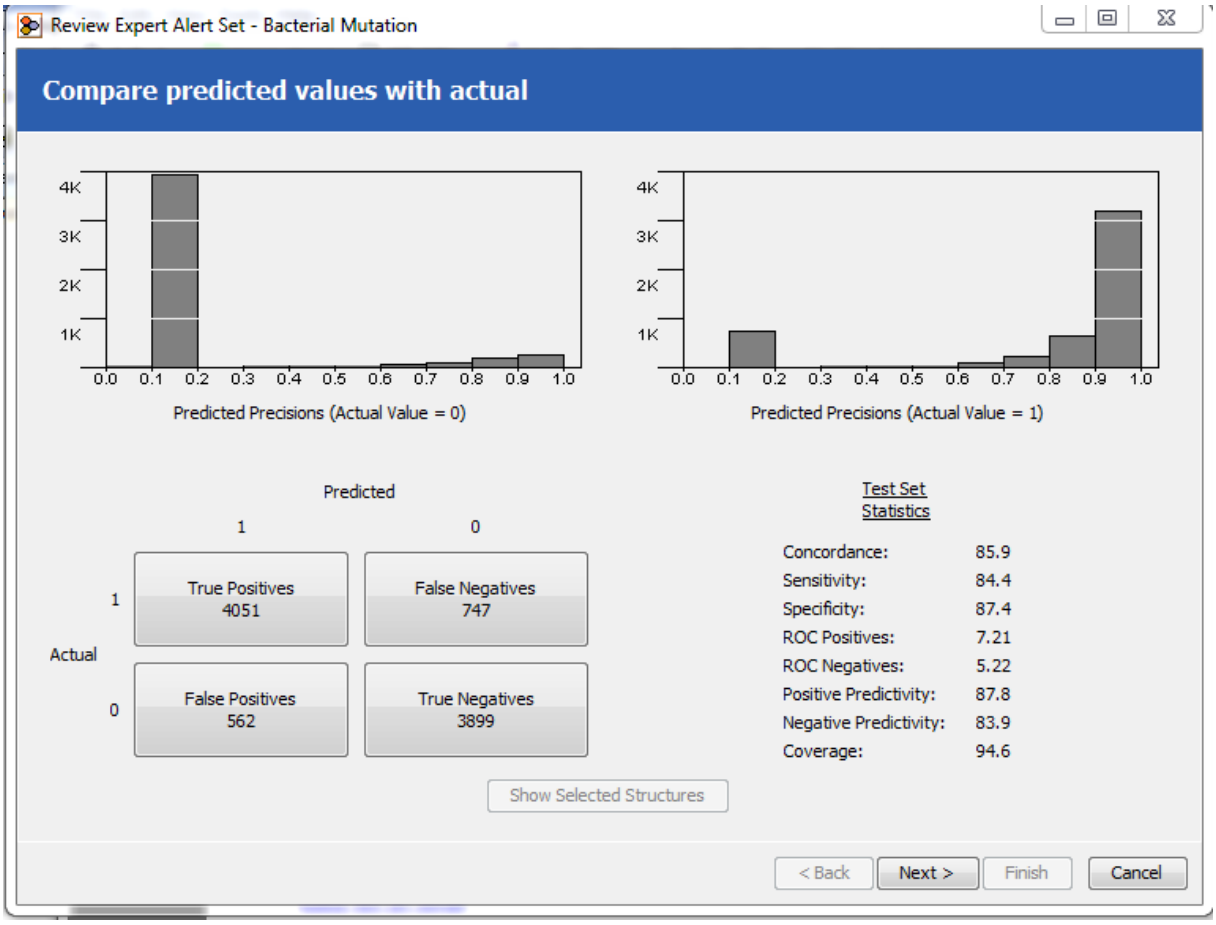
Negative experimental data

Not enough evidence

Selecting the Gentox Expert Alerts tab on the summary sheet displays (only) the results from applying the expert alerts against the test set. The alerts (if any are found) are highlighted on the test structures in the first column. The names of all matched alerts are displayed in a separate column. The alert precision column displays the mean bacterial mutation value for all reference compounds having the matching alert. In the case where multiple alerts match, this mean value corresponds to the alert with the highest mean value. The Generate Reports button will generate a report for the matching alerts. However, since this view is specific to only alert results, the report will not contain a consensus prediction nor any results from statistical model predictions. Performance against a validation set can be calculated using the Compare with Data button. Note that this uses (Q)SAR predictions and any experimental data when comparing results with an external data set. Additional columns may be added or removed using the Add/Remove Columns button at the top right of the dialog. The Find Analogs button allows searching for known compounds similar to the test compound in either the reference set for the alerts (that was used for qualifying the alerts) or against a Leadscope databases containing genotoxic and carcinogenicity experimental data. See section C-7 for further details.



The result of Comparing with Data is shown below. This feature compares the prediction results against a user-specified property calculating the Cooper statistics for the result. The user-specified property values are considered the "actual values" for purposes of the comparison. For comparison of alert results against actual values, the histograms display the precisions of the matching alerts when run against a reference set. This provides a visualization of the distribution of the predictability of the alerts against a validation set. Note that for negative predictions, where no alert is found and the test compounds are in the applicability domain, that the precision is about 14%. This indicates that for the reference set, the set of all alerts accounts for 85% of the positive examples in that set. And that 14% of the positive reference compounds did not have an associated alert that could explain its activity.



C-5. Explaining Prediction Results

A new capability added into version 1.3 of the Model Applier is the ability to visually inspect how the software arrived at a particular prediction. This is accomplished by coloring the test structure's atoms and bonds red when representing positive probability prediction contributions ("hot spots") and green/blue when representing detoxifying features ("cold spots"). Deeper shades represent a stronger contribution. The coloring is supported by displaying the partial probability contribution value of each feature to the overall predicted probability value.

There are three ways to explain prediction results in the model applier:

1) A single prediction on a single structure can be explained in detail. This visualization allows contributions of individual features to the overall prediction to be examined.
2) Multiple-model predictions on a single structure can be compared with this feature. This comparison allows visual inspection to see whether the same features are significant in predictions from several models. That is, are the same features responsible for toxicity?
3) A single model predicting values for multiple structures can be compared. For a single model, this allows a series of structures to be compared to see how changing structural features affects the prediction results.

This "Explain" feature is accessed through the "Explain" button that appears on different prediction results pages.

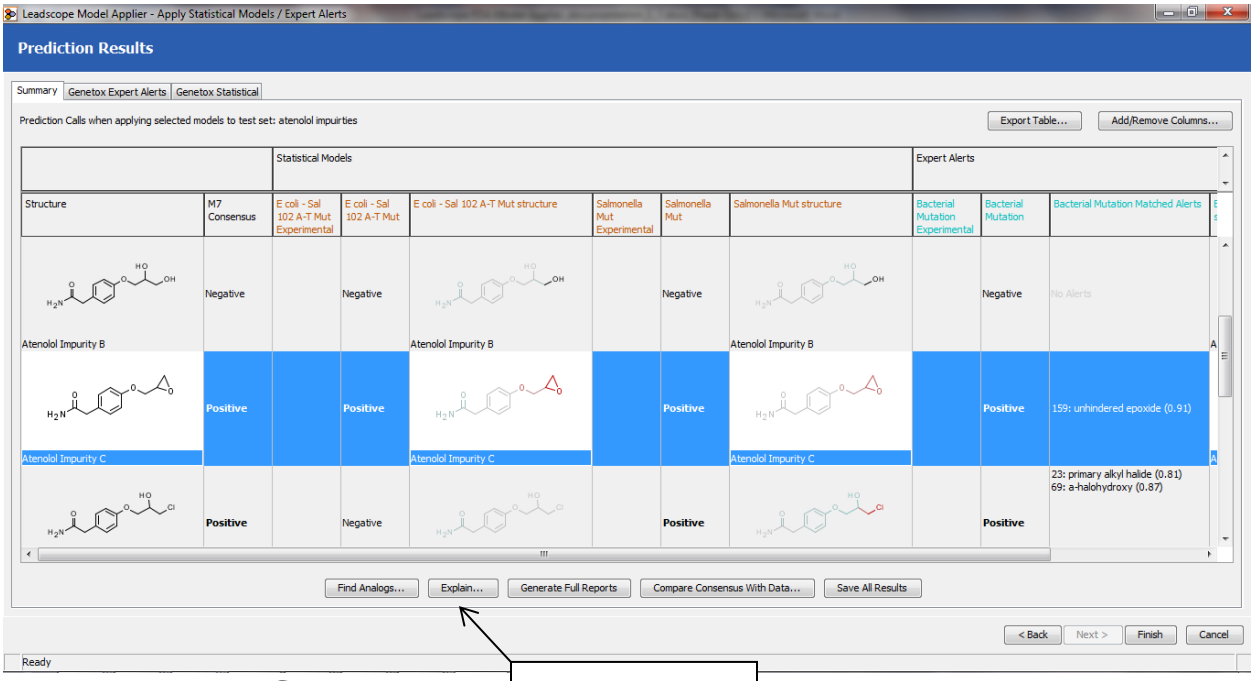


Figure C-5.1

1. Explaining the results of a single model prediction on a single test structure

The detailed results of the predictions will be displayed as in figure C-5.2. The test structure is displayed at the top of the page and is repeated three times. The features matching the structure (14 in this case) are highlighted together (i.e. unioned) on the picture on the left. Weights are calculated for each atom and bond in the picture based on the summation of the contributions from each feature. Both positively and negatively contributing features are summed together to color the structure – darker shades of red represent increasing positive contributions and darker shades of blue/green increasing negative contributions. Black indicates where no features matched a portion of the test structure. In order to more easily distinguish the positive and negative contributions, two additional pictures are displayed, one for only positive contributions (red) and one for only negative contributions (blue/green).

In this example only one model was used; however, it is possible that multiple submodels could be used in the prediction and the average predicted value would be displayed. In this example, the total number of features (from a total training set of 3,974 structures) was 369 with 14 features plus 7 properties used to make this prediction. The features present contributed 87.49% to the predicted value with the properties contributed the remaining amount (12.51%). The "All features" structure display shows the highest contribution to the prediction coming from the ring system colored red. The table of features shows the top two positive features (as it is sorted by default by the %partial contribution column). Both the features contribute positively: 0.283 (28.73%) from the first and 0.2696 (27.37%) from the second.

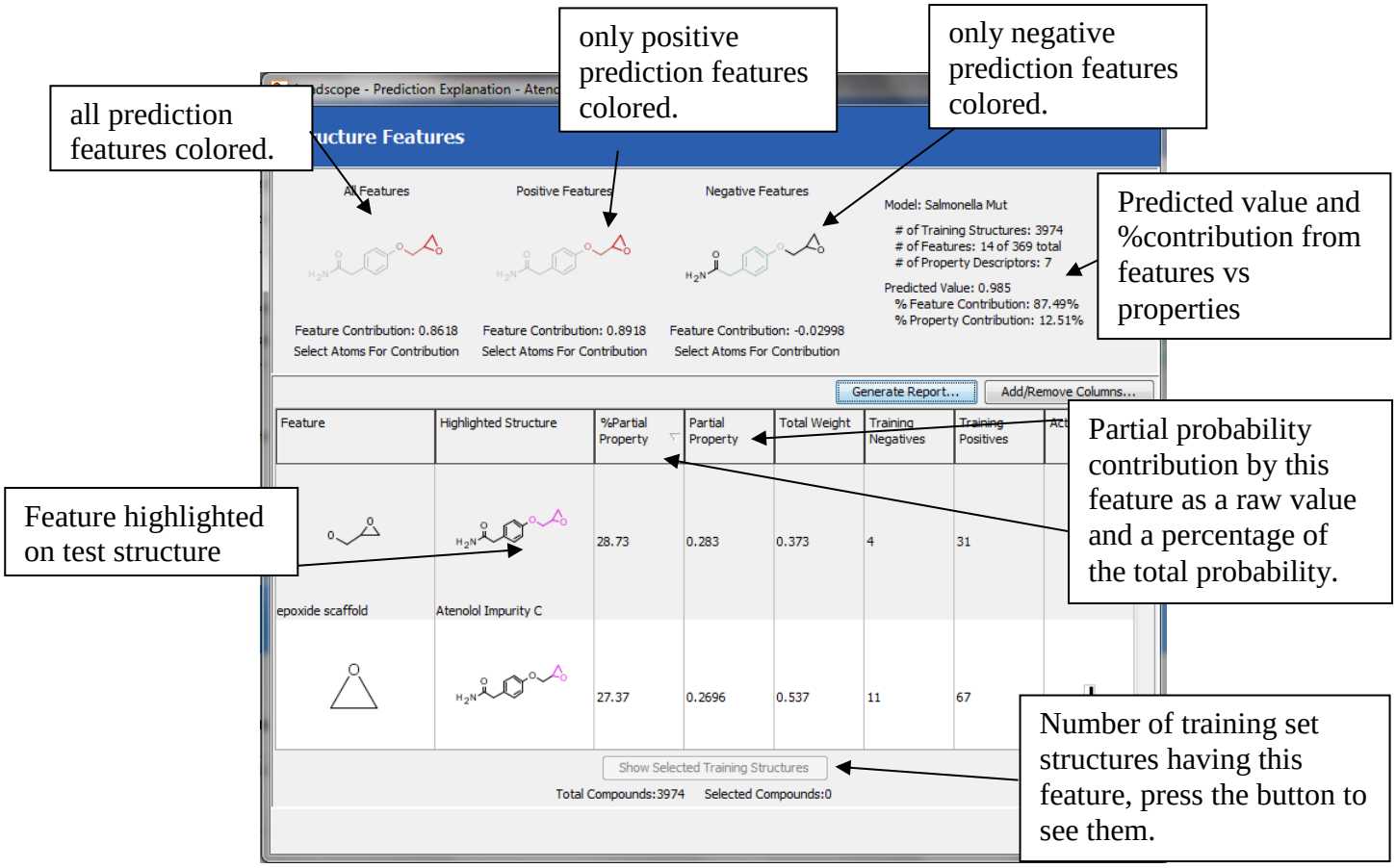


Figure C-5.2

The total individual atom contributions (as summed from all overlapping features) can be calculated when the user selects the atoms via the mouse (lasso or control-click). This provides an easy way to find the contributions of features which may overlap individual atoms or bonds (and to redefine a feature definition). The purpose of this approach is to easily understand the underlying matrix contributions being made towards the overall predicted probability without examining a matrix of numbers. It is the sum of all the positive and negative features contributions for those features matching the test structure. Since it is unlikely that all the features will consist of independent atoms and bonds, the structure displayed at the top of the page combines these together. Since it may be difficult to determine the positive and negative contributions from one picture, these have been separated out into individual pictures.

This display can be useful in determining the major features contributing to an overall positive prediction when the contribution can be confined to a reasonably small substructure of the test structure. It can also be insightful in cases where both highly positive features and highly negative features exist. This may indicate conflicting information in the model (which may reduce the user confidence of this particular prediction), or it may indicate a mitigating effect of

a feature, such as a detoxifying alert in the presence of a primary toxic alert. In any case, the user knows exactly the information that went into the prediction and quantitatively.

The user has the option to examine the training sets matching any selected model feature. When the "Show Selected Training Structures" button is pressed, the training set structures having the selected features are displayed with the feature highlighted.

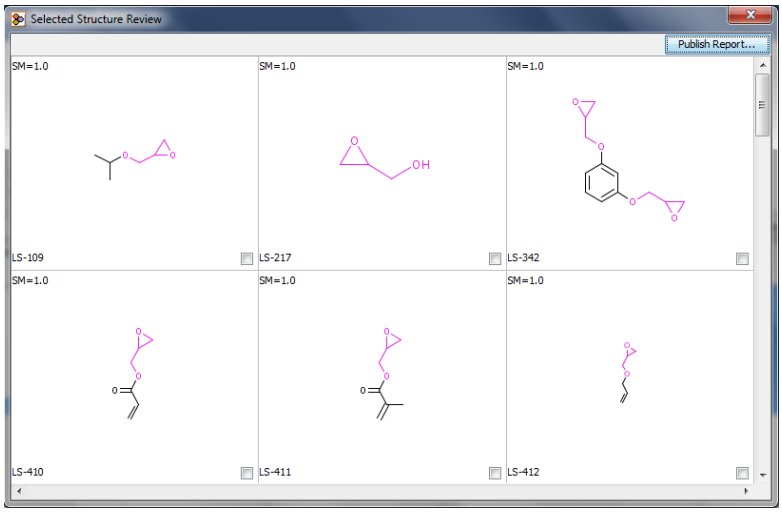


Figure C-5.3

The results in Figure C-5.2 also allow the user to examine the prediction results from individual submodels where they have been used. As each submodel is constructed with different training subsets, different features may arise as the most prominent features for that submodel prediction. In some cases, one submodel prediction may be significantly different (e.g. lower) than the others. That submodel may be missing significant features that were used in the other predictions. If the user can determine this by examination, then the contribution made by this submodel may be validly dismissed.

2. Explaining the results of multiple model predictions on a single test structure

The second type of explanation is to compare the results from several model predictions for a single test structure. If a single row (i.e. structure) is selected from the Summary page (see figure C-5.1), a visual comparison of the all the predictions is possible. The user first selects the model results for comparison. All prediction results may be selected, however, it is most useful to compare positive prediction results to identify common substructures responsible for the prediction. All selected positive predictions could be selected by pressing the "Select Positives Only" button.

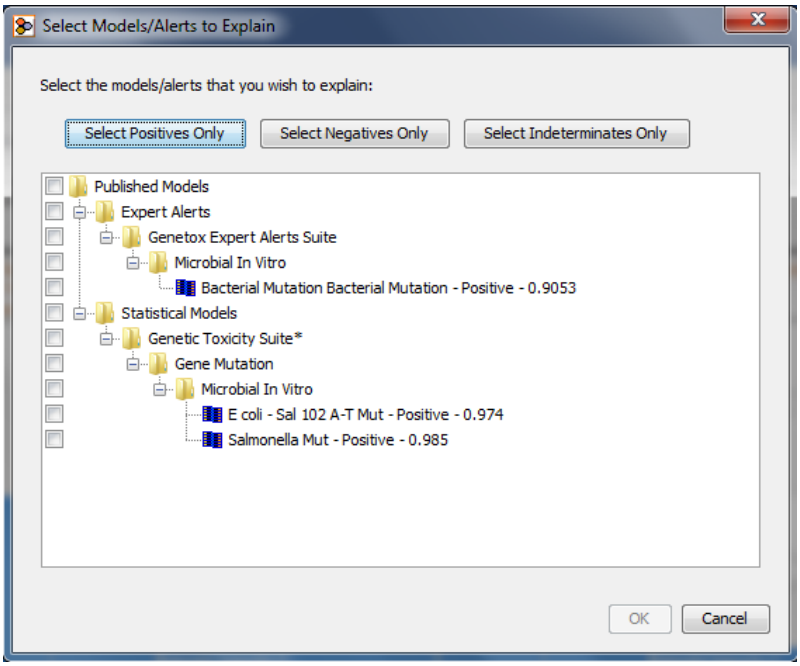


Figure C-5.4

The resulting presentation shows the highlighting of significant features on the same structure for two models. The coloring scheme is absolute, such that the same significance can be assigned the same colored features across all the model results displayed. An explain button is available on each result. It allows further drilling-down into individual results for a more detailed analysis.

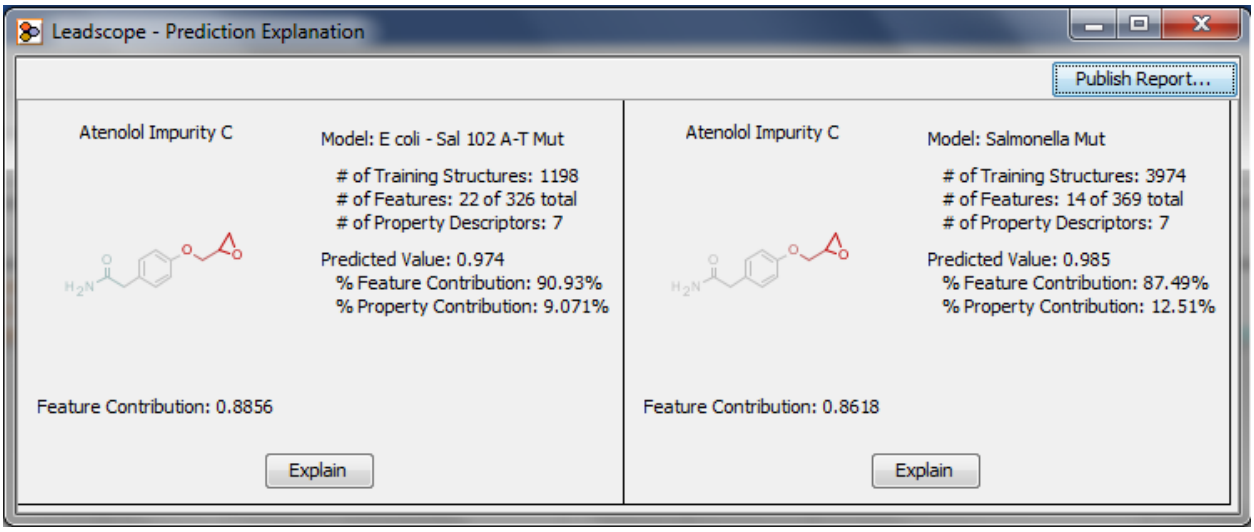


Figure C-5.5

3. Explaining the results of a single model prediction on multiple test structures

The third type of explanation is to compare the prediction results from a single model prediction for multiple test structures. If a more than one row (i.e. structures) are selected from the prediction results for a single model (see Figure C-5.2) then the resulting presentation shows the highlighting of significant features on the series of structures for a single model.

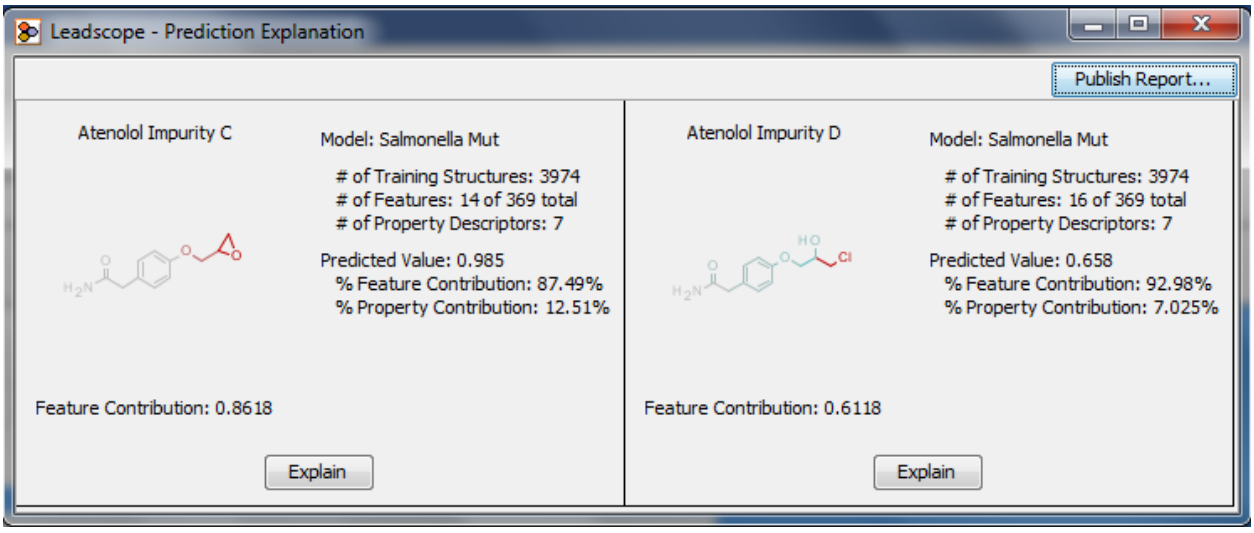


Figure C-5.6

Again, the coloring is absolute, so that differences in contributions (e.g. toxicity) from different features in a series of compounds can be compared.

The explain button, again, is available on each result and allows further drilling-down into individual results for a more detailed analysis.

C-5b: Explaining Alert Results:

It is possible to explain the expert alerts results in a similar manner to the (Q)SAR model. The initial explanation view lists any matching alerts in the test chemical. The "All Alerts" structure will highlight in red all active or indeterminate alerts matching the test compound (as long as the alert has not been deactivated). Any deactivating features that negates an alert will be shown in blue/green (with the alert being deactivated in gray). The "Primary Alert Template" view will only show the alerts, with the "Deactivating Templates" only showing structural features responsible for deactivation of an alert. Where one or more alerts are identified (with no deactivation), each alert is summarized in the main table along with the number of chemicals in the reference set (broken down by positive and negative examples). The precision of the alert is shown corresponding to the proportion of positives examples alongside a small histogram summarizing this proportion. In this example, there were no deactivated alert which would have been shown under a second tab "Deactivated Alerts".

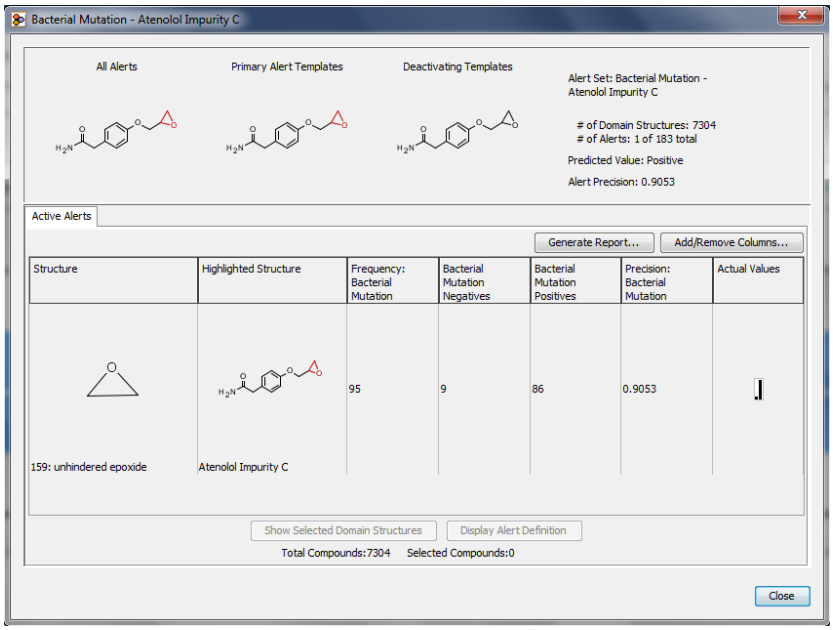


Figure C-5.7

Examples (illustrated in Figure C-5.8) from the reference set can be displayed from this explain view by double clicking on an alert. The definition for the alert can also be shown by clicking on the "Display Alert Definition" button, with the results shown in Figure C-5.9.

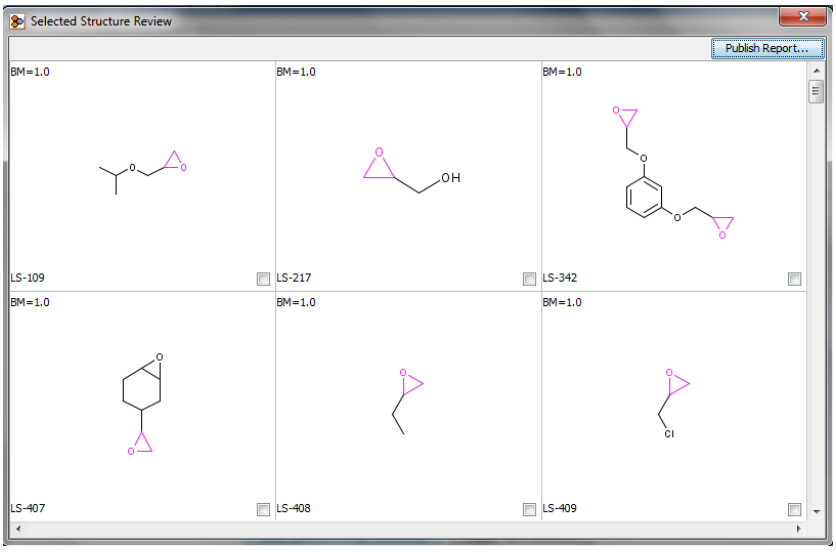


Figure C-5.8

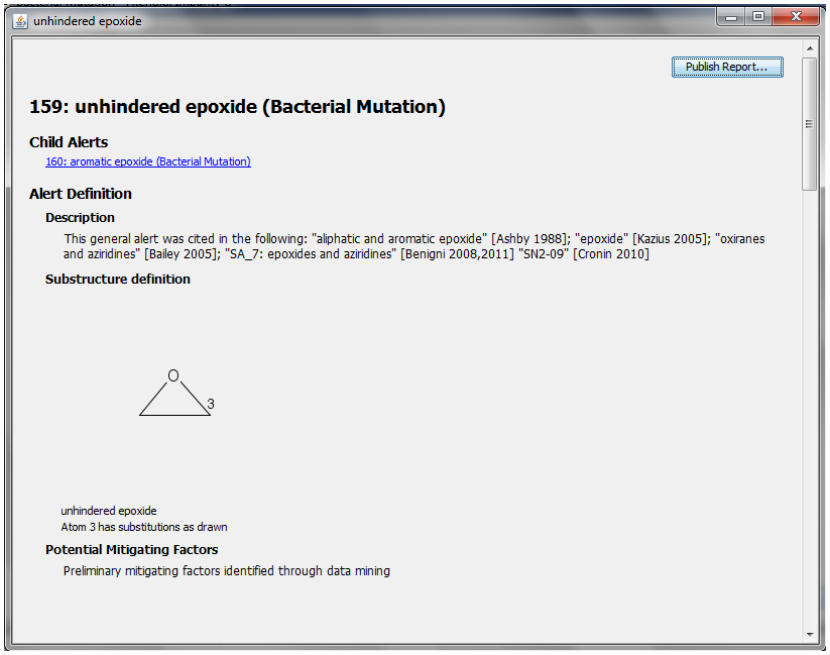


Figure C-5.9

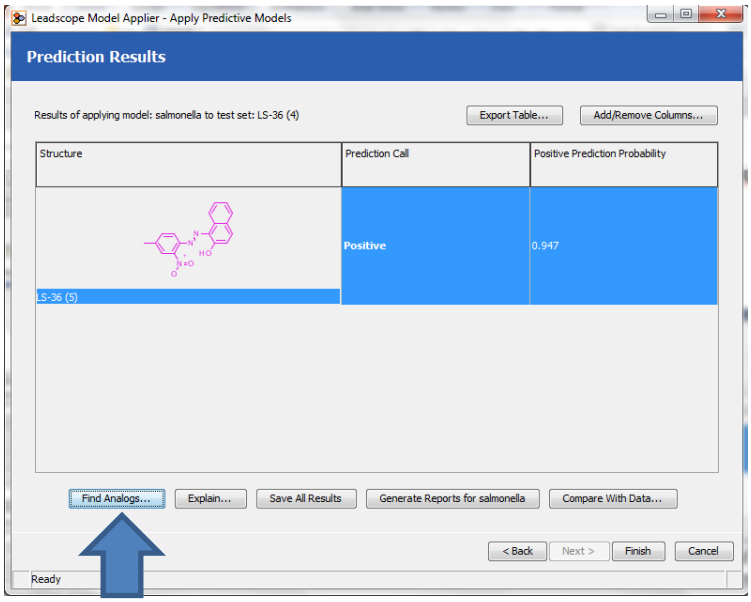
C-6. Reviewing Prediction Results

The Review Prediction Results button on the main page of the Leadscope Model Applier allows for review of previously applied model/alert results. All of the same functionality described in earlier sections of this chapter is available and may be applied. This button was added primarily for pay-as-you-go customers so that they may save their prediction results and review them at a later time without being charged again for making a prediction.

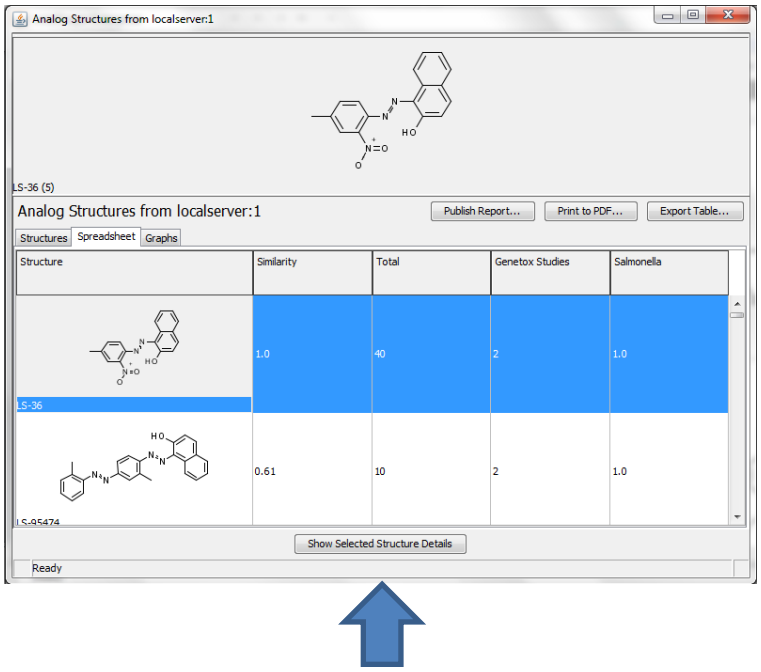
C-7. Browsing Analogs in Leadscope Databases

Once a prediction has been made the user has option of searching companion databases for similar compounds containing experimental data. The Leadscope Genetox and Carcinogenicity Databases are now provided with the purchase of those model suites respectively. This includes overall call data for tested compounds as well as the underlying study-level calls, test-level calls, tests, and literature references for the data. This is in addition to searching analogs in the model training sets which is available under the Explain capability.

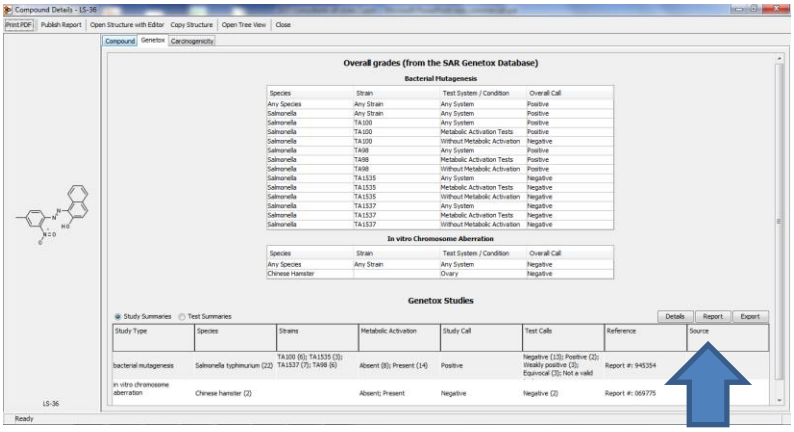
After a prediction is made, the "Find Analogs" button is present as part of the prediction results.



Next the database analogs similar to the test compound are displayed in a spreadsheet. This sheet includes the similarity score of the analogs, overall call data for genotoxicity and carcinogenicity (if available) as well as the study count for each analog.

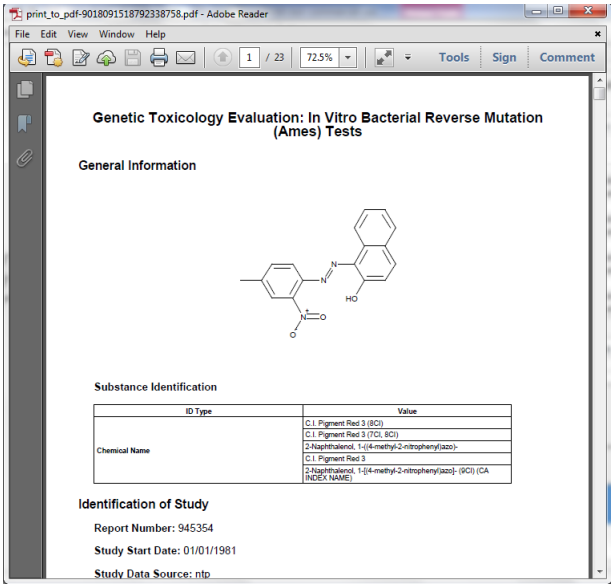

Analog Structures from localserver:1
LS-36 (5)
Analog Structures from localserver:1
Publish Report...
Print to PDF...
Export Table...
Structures
Spreadsheet
Graphs
Structure
Similarity
Total
Genetox Studies
Salmonella
1.0
40
2
1.0
LS-36
0.61
10
2
1.0
Show Selected Structure Details
Ready

Any analogs of interest may be highlighted with the mouse and selected to display all the known database information about that compound.



Tabs at the top of the screen separate the data into a summary of all overall calls, and tabs specifically for genetox and carcinogenicity study and test data. Within these dialogs are summaries of all the studies and tests. Individual studies may be expanded into tests comprising them by selecting the "Test Studies" radio button.

Lastly, reports containing data on the analogs may be created by clicking on the "Report" button.



Further information is available for the companion databases as documentation accessible from the Help button on the main page of the Model Applier.

C-8. The Regulatory Submission Tool

In the ICH M7 guidance "*Assessment and control of DNA reactive (mutagenic) impurities in pharmaceuticals to limit potential carcinogenic risk*" the purpose is stated as:

> "… *to provide a practical framework that is applicable to the identification, categorization, qualification, and control of these mutagenic impurities to limit potential carcinogenic risk*." [30]

All actual and potential impurities or degradation products need to be identified and then a subsequent hazard assessment performed. In the absence of available laboratory data (e.g. rodent carcinogenicity study or a bacterial mutagenesis study) an *in silico* analysis is permitted to assign compounds certain to classes. These results can be supplemented with an expert opinion [31].

Leadscope has published a standard operating procedure (SOP) to complete an *ICH M7* compliant *in silico* assessment to include in a regulatory submission - including accompanying expert opinions. [32]. This SOP is implemented in the Leadscope Model Applier and Leadscope Client as a "Submission tool" (version 2) for ICH M7 that applies the defined process in a consistent manner. The software integrates tools, databases, alerts and statistical models using a step-by-step wizard to generate the necessary submission assessment. These integrated tools perform a search over public literature and on-line databases, integrate proprietary information, and execute the necessary *in silico* prediction methodologies. The tool rapidly generates expert opinions and documents any necessary risk characterization and controls in a consistent and complete manner to be included as part of a regulatory submission. The use of a standardized reporting format facilitates review by both industrial sponsors and regulatory authorities.

For further documentation on the use of the submission tool see the accompanying document: ., Amberg, et. al., Principles and procedures for implementation of ICH M7 recommended (Q)SAR analyses ", Regulatory Toxicology and Pharmacology, 77 (2016) 13-24 and the available online webinars.

Appendix A. Common Organic and Inorganic Salts

The following tables describe the organic and inorganic salts that Leadscope uses to identify parent compounds, along with all amino acids and single inorganic salt fragments:

Table 1. Common Organic Salts

Common Organic Salts	Systematic Name	CAS Registry Number
Acetate	Acetic acid	64-19-7
Aceturate	N-Acetyl-Glycine	543-24-8
Amsonate	2,2'-(1,2-ethenediyl)bis[5-amino-]Benzenesulfonic acid	81-11-8
Armstrong's acid	1,5-Napthalenedisulfonic	81-04-9
Ascorbate	D-Ascorbic acid	10504-35-5
Benzoate	Benzoic acid	65-85-0
Besylate	Benzenesulfonic acid	98-11-3
Brucine	Strychnidin-10-one,2,3-dimethoxy-	357-57-3
Butyrate	Butanoic acid	107-92-6
Camsylate	7,7-Dimethyl-2-oxo-bicyclo[2.2.1]heptane-1-methanesulfonic acid	3144-16-9
Caproate	Hexanoic acid	142-62-1
Carbamate	Carbamic acid	463-77-4
Carbanilate	Phenyl-Carbamic acid	501-82-6
Cinchonidine	2-Quinuclidinemethanol,alpha-4-quinolyl-5-vinyl-(-)	485-71-2
Cinchonine	2-Quinuclidinemethanol,alpha-4-quinolyl-5-vinyl-(+)	118-10-5
Cinnamate	3-phenyl-2-Propenoic acid	621-82-9
Citrate	2-hydroxy-1,2,3-Propanetricarboxylic acid	77-92-9
Clofibrate	2-(4-chlorophenoxy)-2-methyl-Propanoic acid	882-09-7
Cyclamate	cyclohexyl-Sulfamic acid	100-88-9
Cyclohexanamine	Hexahydrobenzenamine	108-91-8
Cypionate	Cyclopentanepropanoic acid	140-77-2
Decanoate	Decanoic acid	334-48-5
Diethanolamine	Ethanol, 2,2'-iminobis-	111-42-2
Diethylamine	N-ethylethanamine	109-89-7

Common Organic Salts	Systematic Name	CAS Registry Number
Dimethylamine	Methanamine, N-methyl-	124-40-3
Diphenylacetic acid	Benzeneacetic acid, alphaphenyl-	117-00-8
Edetate	N,N'-1,2-ethanediylbis[N-(carboxy-methyl)-Glycine	60-00-4
Edisylate	1,2-Ethanedisulfonic acid	110-04-3
Enanthate	Heptanoic acid	111-14-8
Ephedrine	Benzenemethanol, alpha-(1-(methylamino)ethyl)-	299-42-3
Estolate	Dodecyl sulfate (mono salt)	151-41-7
Esylate	Ethanesulfonic acid	594-45-6
Ethanol	Ethyl Alcohol	64-17-5
Ethanolamine	Ethanol, 2-amino-	141-43-5
Formate	Hydrogencarboxylic acid	64-18-6
Fumarate	2-Butenedioic acid (2E)	110-17-8
Gluceptate	D-glycero-D-gulo-Heptonic acid	87-74-1
Gluconate	D-Gluconic acid	133-42-6
Glucose	Glucose	50-99-7
Glutamate	L-Glutamic acid	617-65-2
Glycinate	Aminoacetic acid	56-40-6
Glycolate	Hydroxyacetic Acid	79-14-1
Guanidine	Guanidine	113-00-8
Hexanoic acid	Hexanoic acid	142-62-1
Hippurate	N-benzoyl-Glycine	495-69-2
Isethionate	2-hydroxy-Ethanesulfonic acid	107-36-8
Isopropylamine	2-Propanamine	75-31-0
Lactate	2-Hydroxy-Propanoic acid	50-21-5
Lactobionate	4-O-ß-D-galactopyranosyl-D-Gluconic acid	96-82-2
Laurate	Dodecanoic acid	143-07-7
Maleate	2-Butenedioic acid (2Z)	110-16-7
Malonic acid	Propanedioic acid	141-82-2
Mandelate	a-hydroxy-Benzeneacetic acid	90-64-2
Mesylate	Methanesulfonic acid	75-75-2
Methiodide	Iodomethane	74-88-4
Methyl Bromide	Bromomethane	74-83-9
Methyl carbamate	Carbamic acid methyl ester	598-55-0
Monomethyl sulfate	Sulfuric acid monomethyl ester	75-93-4
Morpholine	Tetrahydro-2H-1,4-oxazine	110-91-8
Mucate	Galactaric acid	526-99-8

Common Organic Salts	Systematic Name	CAS Registry Number
Myristate	Tetradecanoic acid	544-63-8
Naphthalenesulfonic acid	1-Naphthalenesulfonic acid	85-47-2
Napthylethanamine	1-Naphthaleneethanamine	4735-50-6
Nicotinate	3-Pyridinecarboxylic acid	59-67-6
Oleate	9-Octadecenoic acid (9Z)	112-80-1
Oxalate	Ethanedioic acid	144-62-7
Palmitate	Hexadecanoic acid	57-10-3
Pamoate	4,4'-methylenebis[3-hydroxy-2-Naphthalenecarboxylic acid	130-85-8
Pentetate	N,N'-bis[2-[bis(carboxymethyl)amino]-Glycine	67-43-6
Phenpropinate	Phenylpropanoic acid	501-52-0
Phenylacetate	Phenylacetic acid	103-82-2
Phosphanilate	(4-aminophenyl)-Phosphonic acid	5337-17-7
Phthalate	1,2-Benzenedicarboxylic acid	88-99-3
Picrate	2,4,6-trinitro-Phenol	88-89-1
Picric acid	2,4,6-Trinitrophenol	29663-11-4
Piperazine	Piperazine	110-82-2
Piperidine	Hexahydropyridine	110-89-4
Pivalate	2,2-dimethyl-Propanoic acid	75-98-9
p-Nitrobenzoic acid	4-Nitrobenzoic acid	62-23-7
Probenate	4-[(dipropylamino)sulfonyl]-Benzoic acid	57-66-9
Propionate	Propanoic acid	79-09-4
p-Salicylic acid	4-Hydroxybenzoic acid	99-96-7
Pyridine	Azabenzene	110-86-1
Pyrrolidine	Pyrrolidine	123-75-1
Salicylate	2-hydroxy-Benzoic acid	69-72-7
Sorbic acid	2,4-Hexadienoic acid	110-44-1
Stearate	Octadecanoic acid	57-11-4
Stinoprate	N-Acetyl-L-cysteine	616-91-1
Suberic acid	Octanedioic acid	505-48-6
Succinate	Butanedioic acid	110-15-6
Sulfamate	Aminosulfonic acid	5329-14-6
Tartrate	2,3-dihydroxy-Butanedioic acid	87-69-4
Terephthalate	1,4-Benzenedicarboxylic acid	100-21-0
Tetrabutanaminium	1-Butanaminium, N,N,Ntributyl-	10549-76-5

Common Organic Salts	Systematic Name	CAS Registry Number
Tetraethanaminium	Ethanaminium, N,N,N-triethyl-	66-40-0
Tosylate	4-methylbenzenesulfonic acid	104-15-4
Triethanamine	Ethanamine, N,N-diethyl-	121-44-8
Triethanolamine	2,2',2''-Nitrilotriethanol	102-71-6
Trifluoro acetate	Trifluoroacetic Acid	76-05-1
Trimethyl acetate	Trimethyl acetic acid	75-98-9
Tropate	3-Hydroxy-2-phenylpropanoic acid	529-64-6
Undecylenate	10-Undecenoic acid	112-38-9
Urea	Carbamimidic acid	57-13-6
Valerate	Pentanoic acid	109-52-4
Xanthanoic acid	Xanthene-9-carboxylic acid	82-07-5

TABLE 2. Common Inorganic Salts

Common Inorganic Salts	CAS Registry Number
Aluminum (III) salt	7429-90-5
Ammonium salt	14798-03-9
Antimony salt	7440-36-0
Barium (II) salt	7440-39-3
Beryllium salt	7440-41-7
Bismuth(III) salt	
Borate	11113-50-1
Bromide	24959-67-9
Cadmium (II) salt	7440-43-9
Calcium(II) salt	7440-70-2
Carbonate	463-79-6
Cerium salt	7440-45-1
Cerium (III) salt	
Cerium (IV) salt	
Cesium salt	7440-46-2
Chlorate	7790-93-4
Chloride	16887-00-6
Chlorite	7790-93-4
Chromium(III) salt	7440-47-3
Cobalt salt	7440-48-4
Cobalt(II) salt	7440-48-4
Cobalt(III) salt	7440-48-4
Copper salt	7440-50-8
Copper(I) salt	7440-50-8

Common Inorganic Salts	CAS Registry Number
Copper(II) salt	7440-50-8
Diphosphate	14000-31-8
Dysprosium(III) salt	7429-91-6
Erbium(III) salt	7440-52-0
Europium(III) salt	7440-53-1
Tetrafluoroborate	16872-11-0
Gadolinium(III) salt	7440-54-2
Gallium(III) salt	7440-55-3
Hexafluorophosphate	16919-18-9
Holmium(III) salt	7440-60-0
Hydrate	
Hydriodide	10034-85-2
Hydrobromide	10035-10-6
Hydrochloride	7647-01-0
Hydrofluoride	7664-39-3
Hydrogen cyanide	74-90-8
Hydrogen phosphate	
Hydrogen sulfide	
Hypochlorite	14380-61-1
Indium(III) salt	7440-74-6
Iodide	20461-54-5
Iron salt	7439-89-6
Iron(II) salt	7439-89-6
Iron(III) salt	7439-89-6
Lanthanum(III) salt	7439-91-0
Lead salt	7439-92-1
Lead(II) salt	7439-92-1
Lead(IV) salt	7439-92-1
Lithium salt	7439-93-2
Lutetium(III) salt	7439-94-3
Magnesium(II) salt	7439-95-4
Manganese(II) salt	7439-95-4
Mercury salt	7439-97-6
Mercury(I) salt	7439-97-6
Mercury(II) salt	7439-97-6
Neodymium(III) salt	7440-00-8
Nickel(II) salt	7440-02-0
Nitrate	7697-37-2
Nitrite	7782-77-6
Palladium(II) salt	

Common Inorganic Salts	CAS Registry Number
Perchlorate	7601-90-3
Phosphate	14265-44-2
Phosphonic Acid	13598-36-2
Plutonium(VI) salt	
Potassium salt	
Praseodymium(III) salt	7440-10-0
Rhodium(II) salt	7440-16-6
Rubidium salt	7440-17-7
Ruthenium salt	7440-18-8
Samarium(III) salt	7440-19-9
Selenium(II) salt	7782-49-2
Silicate	7699-41-4
Silver salt	7440-22-4
Sodium salt	7440-23-5
Strontium salt	7440-24-6
Sulfate	7664-93-9
Sulfite	14265-45-3
Sulfurous acid	7782-99-2
Terbium(III) salt	7440-27-9
Thallium salt	7440-28-0
Thallium(I) salt	7440-28-0
Thallium(III) salt	7440-28-0
Thiosulfate	
Thorium salt	7440-29-1
Thulium(III) salt	7440-30-4
Tin salt	7440-31-5
Tin(II) salt	7440-31-5
Tin(IV) salt	7440-31-5
Titanium(IV) salt	7440-32-6
Uranium salt	7440-61-1
Vanadium salt	7440-62-2
Ytterbium(III) salt	7440-64-4
Yttrium(III) salt	7440-65-5
Zinc salt	7440-66-6
Zinc(II) salt	7440-66-6
Zirconium(IV) salt	7440-67-7

D. References

1 Benz, R.D. and Irausquin, H. (1991) Priority-based Assessment of Food Additives Database of the U.S. Food and Drug Administration Center for Food Safety and Applied Nutrition. *Environmental Health Perspectives* 96:85-89.

2 Contrera, J.F. (1998) Transgenic Animals: Refining the Two Year Rodent Carcinogenicity Study. *Laboratory Animals* 27:30-34.

3 Contrera, J.F. and DeGeorge, J.J. (1998) *In Vivo* Transgenic Bioassays and the Assessment of the Carcinogenic Potential of Pharmaceuticals. *Environmental Health Perspectives* 106(Sup1):71-80.

4 Contrera, J.F., Jacobs, A.C., DeGeorge, J.J., Chen, C.H., Choudary, J.B., DeFelice, A.F., Fairweather, W.R., Farrelly, J.G., Fitzgerald, G.G., Goheer, A.M., Jordan, A.W., Kelly, R.E., Lin, D., Lin, K.K., Meyers, L.L., Osterberg, R.E., Prasanna, H.R., Resnick, C.A., Sheevers, H.V. and Sun, J. (1997) Carcinogenicity Testing and the Evaluation of Regulatory Requirements for Pharmaceuticals. *Regulatory Toxicology and Pharmacology* 25:130-145.

5 Contrera, J.F., Matthews, E.J., and Benz, R.D. (2003) Predicting the Carcinogenic Potential of Pharmaceuticals in Rodents Using Molecular Structural Similarity and E-State Indices. *Regulatory Toxicology and Pharmacology* 38:243-259.

6 Contrera, J.F., Matthews, E.J., Kruhlak, N.L., and Benz, R.D. (2004) Estimating the Safe Starting Dose in Phase I Clinical Trials and No Observed Effect Level Based on QSAR Modeling of the Human Maximum Recommended Daily Dose. *Regulatory Toxicology and Pharmacology* 40:185-206.

7 Contrera, J.F., Matthews, E.J., Kruhlak, N.L., and Benz, R.D. (2005) In Silico Screening of Chemicals for Bacterial Mutagenicity Using electrotopological E-state Indices and MDL QSAR Software. *Regulatory Toxicology and Pharmacology* 43:313-323.

8 Contrera, J.F., MacLaughlin, P., Hall, L.H., and Kier, L.B. (2005) QSAR Modeling of Carcinogenic Risk Using Discriminant Analysis and Topological Molecular Descriptors. *Current Drug Discovery Technologies* 2:55-67.

9 Contrera, J.F. and Woodcock, J. (1996) Pilot Project for Electronic Submissions. *The Regulatory Affairs Journal* 7(3):256-257.

10 Dearfield, K.L. and Benz, R.D. (1999) Can the New Genetic Toxicology Tests be Used for Regulatory Safety Decisions? *Environmental and Molecular Mutagenesis* 33:91-93.

11 DeGeorge, J.J. and Contrera, J.F. (1996) A Regulatory Perspective on the Utility of Two Rodent Species in Carcinogenicity Testing. In *Proceeding, Third International Conference on*

Harmonization, Yokohama, 1995 (P.F. D'Arcy and D.W.G. Harron, Eds.), pp. 274-277, Greystone Books Ltd., N. Ireland.

12 DeGeorge, J.J., Myers, L.L., Takahashi, M. and Contrera, J.F. (1999) The Duration of Non-rodent Toxicity Studies for Pharmaceuticals. *Toxicology Science* 49:143-155.

13 Matthews, E.J., Benz, R.D., and Contrera, J.F. (2000) Use of Toxicological Information in Drug Design. *Journal of Molecular Graphics and Modeling* 18(December):605-614.

14 Matthews, E.J. and Contrera, J.F., (1998) A New Highly Specific Method for Predicting the Carcinogenic Potential of Pharmaceuticals in Rodents Using Enhanced MCASE QSAR-ES Software. *Regulatory Toxicology and Pharmacology* 28:242-264.

15 Matthews, E.J., Kruhlak, N.L., Benz, R.D., and Contrera, J.F. (2004) Assessment of the Health Effects of Chemicals in Humans: I. QSAR Estimation of the Maximum Recommended Therapeutic Dose (MRTD) and No Effect Level (NOEL) of Organic Chemicals Based on Clinical Trial Data. *Current Drug Discovery Technologies* 1:61-76.

16 Matthews, E.J., Kruhlak, N.L., Cimino, M.C., Benz, R.D., and Contrera, J.F. (2006) An Analysis of Genetic Toxicity, Reproductive and Developmental Toxicity, and Carcinogenicity Data: I. Identification of Carcinogens Using Surrogate Endpoints. *Regulatory Toxicology and Pharmacology* 44:83-96.

17 Matthews, E.J., Kruhlak, N.L., Cimino, M.C., Benz, R.D., and Contrera, J.F. (2006) An Analysis of Genetic Toxicity, Reproductive and Developmental Toxicity, and Carcinogenicity Data: II. Identification of Genotoxicants, Reprotoxicants, and Carcinogens Using In Silico Methods. *Regulatory Toxicology and Pharmacology* 44:97-110.

18 Stavitskaya, L., et. al., Development of Improved *Salmonella* Mutagenicity QSAR Models UsingStructural Fingerprints of Known Toxicophores, *FDA Center for Drug Evaluation and Research*, Soc. of Tox. 2013.

19 Stavitskaya, L., et. al., Development of Improved QSAR Models for Predicting A-T Base Pair Mutations, Genetic Toxicity Association. 2013.

20 Matthews, E.J., Kruhlak, N.L., Weaver, J.L., Benz, R.D., and Contrera, J.F. (2004) Assessment of the Health Effects of Chemicals in Humans: II. Construction of an Adverse Effects Database for QSAR Modeling. *Current Drug Discovery Technologies* 1:243-254.

21 Matthews, E.J., Fried, A.; Prediction of Drug-Related Cardiac Adverse Effects in Humans: A: Creation of a Database of Effects and Identification of Factors Affecting Their Occurance, in-press, accepted manuscript, *Regulatory Toxicology and Pharmacology*.

22 Fried, A.; Matthews, E.J., Prediction of Drug-Related Cardiac Adverse Effects in Humans: B: Use of QSAR Programs for Early Detection of Drug-Induced Cardiac Toxicities, in-press, accepted manuscript, *Regulatory Toxicology and Pharmacology*.

23 Leadscope Genetox Expert Alerts, White Paper, Leadscope, (May 2017).

24 FDA CDFS DDSR liver database:
http://www.fda.gov/Cder/Offices/OPS_IO/adverse_effect_database.htm

25 Hansen K. et al., J. Chem. Inf. Model. (2009), 49:2077-2081

26 Kruhlak, N. L, Enhanced (Q)SAR Models for Prediction Rodent Carcinogenicity, Society of Toxicology Meeting, San Diego, CA, 2015.

27 Hewes, K.P., Construction and Application of (Q)SAR Models to Predict *in vitro* Chromosome Aberrations, Society of Toxicology Meeting, San Diego, CA, 2015.

28 Zhu, X., Kruhlak, N. L, Construction and analysis of a human hepatotoxicity database Suitable for QSAR modeling using post-market safety data, Toxicology 321 (2014) 62–72.

29 Kruhlak, N. L., et al, Development of QSAR Models for Prediction of Hepatotoxicity Using Post-Market Data from US FDA's AERS Database, Society of Toxicology Meeting, Phoenix, AZ, 2014, abstract 2273h.

30 http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Multidisciplinary/M7/M7_Step_4.pdf

31 Powley MW, (Q)SAR assessments of potentially mutagenic impurities: A regulatory perspective on the utility of expert knowledge and data submission , Regulatory Toxicology and Pharmacology 71 (2015) 295–300.

32 Amberg, A., Beilke, L.D., Bercu, J., Bower, D.A., Brigo, A., Cross, K.P., Custer, L., Dobo, K., Dowdy, E., Ford, K.A., Van Gompel, J., Harvey, J., Kemper, R., Kenyon, M., Kruhlak, N.L., Glowienke, S., Jolly, R., Leavitt, P., Miller, S.A., Muster, W., Nicolette, J., Quigley, D.P., Reddy, Plaper, A., Sprikl, H., Stavitskaya, L., Teasdale, A., Weiner, S., White, A.T., Wichard, J., Myatt, G.J., Principles and procedures for implementation of ICH M7 recommended (Q)SAR analyses , Regulatory Toxicology and Pharmacology, 77 (2016) 13-24 freely available.

33 Yoo, J.W., et. Al, Development of Improved (Q)SAR Models for Predicting the Outcome of the in vivo Micronucleus Genetic Toxicity Assay, Society of Toxicology Meeting, San Diego, CA, 2015.

34 Moore, M.M., et.al., Mouse lymphoma thymidine kinase gene mutation assay: Meeting of the International Workshop on Genotoxicity Testing, San Francisco, 2005, recommendations for 24-

h treatment, Mutation Research/Genetic Toxicology and Environmental Mutagenesis 627(1) (2007), 36–40.

35 Cross, K.P., el. al, Bacterial Mutation (Q)SAR Alerts using Limited-Ames and Tester Strain Data, Society of Toxicology Meeting, New Orleans, LA, 2016.