Leadscope Model Applier and the ICH M7 Impurities Guidelines
Frequently Asked Questions

The following frequently asked questions (FAQs) are based on Leadscope Model Applier version 2.2

August 2017
1. Why do I need to run an expert rule-based system or a statistical-based model for impurities or degradants?

**Answer:** The ICH M7 guideline entitled “Assessment and control of DNA reactive (mutagenic) impurities in pharmaceuticals to limit potential carcinogenic risk” states that expert rule-based and a statistical-based models ((Q)SAR models) can be used to predict the outcome of a bacterial mutagenicity assay to support hazard assessment. In certain cases these predictions can avoid having to test impurities or degradants for bacterial mutagenicity.


2. The ICH M7 guidelines state: “Two (Q)SAR prediction methodologies that complement each other should be applied. One methodology should be expert rule-based and the second methodology should be statistical-based.” What methodologies does Leadscope use?

**Answer:** The Leadscope Model Applier contains both statistical-based models as well as expert rule-based models satisfying the requirement for having the two specified methodologies.

3. Does the Leadscope model applier provide any help in selecting the expert alerts and statistical models to use in supporting ICH M7 as well as generating an overall call for ICH M7?

**Answer:** Yes. The Leadscope Model Applier v2.2 user interface includes a button “Use ICH M7 settings” that will select the appropriate expert rule-based and statistical-based models for ICH M7 as well as the appropriate settings. The results from the two methodologies are presented alongside any available experimental data on the test compounds. An overall assessment (M7 Consensus Call) is also calculated based on all information, both predicted and experimental.

4. Which software applications does the U.S. FDA use in a regulatory capacity to support M7 submissions?

**Answer:** The U.S. Food and Drug Administration, Center for Drug Evaluation and Research, accepts M7 submissions containing (Q)SAR results generated by models that are consistent with the OECD Principles for the Validation of (Q)SARs [http://ihcp.jrc.ec.europa.eu/our_labs/predictive_toxicology/background/oecd-principles](http://ihcp.jrc.ec.europa.eu/our_labs/predictive_toxicology/background/oecd-principles). They consider (Q)SAR bacterial mutagenicity models from several vendors sufficiently validated and acceptable for use as part of M7 submissions to regulate impurities. Examples of some acceptable models include those contained in the Leadscope Model Applier, Lhasa’s Derek Nexus, and MultiCASE’s CASE Ultra. All predictions should be performed using an up-to-date version of the software and models. Leadscope is committed to delivering the most current versions of the QSAR models commercially available to customers as soon as they are released.

5. In the Leadscope Model Applier, which expert alerts should be used to support the ICH M7 guidelines i.e. to “predict the outcome of a bacterial mutagenicity assay”?

---
Answer: The Bacterial Mutation expert alerts (from the Genetox expert alerts suite) should be run to support ICH M7.

6. Within the Leadscope Model Applier, the Leadscope Genetic Toxicity Statistical QSAR Suite contains around 30 different models. Which ones do I need to run to support the ICH M7 guidelines i.e. to “predict the outcome of a bacterial mutagenicity assay”?

Answer: The Gene Mutation-Microbial In Vitro models - *Salmonella Mut* and *E. coli – Sal 102 A-T Mut* should be the only prediction results used to support the ICH M7 guidelines.

7. I understand the Leadscope Genetic Toxicity Statistical QSAR suite contains models other than these models. Is it possible to run a QSAR analysis and not generate results for other models?

Answer: Yes. You have the option to select individual predictive models within a Model Suite to apply to your test set. The Leadscope Model Applier will then only make predictions for those models selected.

8. How should a positive result be interpreted in another model from the Leadscope Genetic Toxicity Statistical QSAR Suite?

Answer: The Gene Mutation-Microbial In Vitro models - *Salmonella Mut* and *E. coli – Sal 102 A-T Mut* should be the only statistical QSAR prediction results used to support the ICH M7 guidelines.

9. Can you explain how the Leadscope expert rule-based methodology works?

Answer: Leadscope worked with expert genetic toxicologists in the development of this alerts system. The Leadscope Genetox Expert Alerts are based on well-defined mutagenicity structural alerts from the literature. Alongside this list of alerts, deactivating factors as well as active subclasses are also encoded, along with details on the source of the alerts and the biological mechanisms. These alerts have been validated against a large database of approximately 10,000 chemicals with Ames data (referred to as the reference set). In addition, these alerts have been validated against proprietary compounds from pharmaceutical companies and other organizations. Leadscope has used a SAR fingerprint methodology to improve the predictive performance of specific classes, such as aromatic amines. This methodology extracts structural knowledge related to factors that activate or deactivate aromatic amine mutagenicity. A positive prediction is made where one or more alerts are present with no defined deactivating factor. In addition, the software determines whether the test compound is similar enough to known classes of chemicals such that it is not trying to extrapolate to areas of chemistry the system has never seen. A positive prediction is only made when the compound is within this applicability domain. A negative prediction is made for chemicals that are within the applicability domain that either contains no alert or when the alert is deactivated.

See the Leadscope Genetox Expert Alerts white paper for more details

10. Can you explain how the Leadscope’s statistical-based methodology works?

**Answer:** The models were built using the Leadscope Predictive Data Miner software and the training data sets were compiled at the U.S. Food and Drug Administration (FDA) by the Division of Applied Regulatory Science (DARS). The QSAR models are implemented with molecular descriptors that include structural features and 7 calculated properties. The structural features include a set selected from Leadscope’s 27,000 pre-defined structural features, predictive scaffolds (larger structural features that show association or lack of association to the toxicity endpoint) and structural alerts identified from the literature or through an analysis for larger databases. The seven calculated properties used are: parent molecular weight, aLogP, polar surface area, hydrogen bond acceptors, hydrogen bond donors, number of rotational bonds and Lipinski score (rule violation). The QSAR models are built using the structural feature and properties as descriptors also described as x- or independent variables. The models encode the relationship between these descriptors and the toxicity endpoint, such as the results of the bacterial mutagenesis assay (i.e., y-variable or response variable). The modeling technique used to generate these models is referred to as partial logistic regression.

When a prediction is made on a new chemical, the same structural features and properties in the model are calculated for the test compounds. These descriptors are then used with the models to calculate a probability of a positive result, as long as the test compound is within the applicability domain of the model.

See appendix A for more details.

11. Do the Leadscope Model Applier methodologies follow the validation principles set forth by the Organisation for Economic Co-operation and Development (OECD)?

**Answer:** Yes. Accompanying each model is a QMRF (QSAR Model Reporting Format) report detailing how each model and alert adheres to these principles.

12. I understand the statistical based QSAR models are derived from a training set of historical results. How many chemicals does it include, where do they come from, and what sort of chemistries do they cover?

**Answer:** The Salmonella training set was developed at the U.S. FDA using a non-proprietary set of 3,979 compounds from FDA approval packages and the published literature. This includes a recent addition of 445 chemicals, 247 which are drug molecules marketed between 1970 and 2011. Others were added to specifically fill data gaps within the training set that were identified using structural features derived from known toxicophores. These include 155 new examples containing functional groups such as azides, amine oxides, hindered epoxides, propriolactones, quinones, amine halides, diazines, azo compounds, diazoniums, sulfates, aziridine chlorides, nitrites, hydrazines, nitriles, isocyanates, and sulfur mustards and 40 compounds containing previously unmodeled atoms such as boron, silicon, selenium, and tin.
The *E. coli* – *Sal 102 A-T Mut* model was constructed to predict A-T base pair mutations and to improve performance and coverage over the previous *E. coli* models. The training set of 1,198 compounds was composed of non-proprietary data from publicly available U.S. FDA approval packages and the published literature for *E. coli* WP2 uvrA, *E. coli* WP2 uvrA (pKM101), and *S. typhimurium* TA102. The training set was expanded to include molecular features from more recently marketed drugs as well as targeting areas of chemical space where previous models were known to have weaknesses.

Both this model and the Salmonella model can be used to support the ICH M7 guideline.

13. **Who developed the Leadscope statistical models and expert alerts?**

**Answer:** All the genotoxicity QSARs were constructed with the U.S. Food and Drug Administration (FDA) Center for Drug Evaluation and Research, Division of Applied Regulatory Science (DARS), under a Research Collaboration Agreement.

Leadscope scientists along with Dr. Errol Zeiger and Dr. Ronald Snyder developed the Leadscope Model Applier - Genetox Expert Alerts system.

14. **What sorts of organizations use the Leadscope Model Applier?**

**Answer:** The Leadscope Model Applier is used by pharmaceutical (large and small), biotech, chemical, cosmetic and food product companies. It is also used by regulatory agencies in Europe, United States and Canada. Users can obtain the same results for their compounds as are obtained by the regulatory agencies.

In addition, Leadscope has over 25 toxicology consulting firms participating in our toxicity consulting program. Those consultants utilize the Leadscope Model Applier in performing predictions for their clients. Predictions include analysis of impurities in accordance with the ICH M7 draft guidelines. Results of the consultants’ predictions made using Leadscope Model Applier have been used in regulatory submissions.

15. **I understand that the Leadscope Model Applier calculates a probability of a positive prediction for the statistical QSAR models – what cut-off should I use to determine a positive or negative result?**

**Answer:** For mutagenicity predictions, it is generally accepted to use a cut-off of greater than or equal to 0.6 for a positive call, less than 0.4 for a negative call and assign calls between 0.4 and 0.6 as equivocal. These are the cutoffs presently used by the U.S. FDA in evaluating M7 QSAR submissions using Leadscope.

16. **How do I assign confidence to a prediction result?**

**Answer:** The Leadscope Model applier will calculate a probability of a positive outcome for the QSAR models. This value, along with information used to explain the results (such as the structural analogs from the training set) can be used to assess the weight of the evidence. The genetox expert alerts are
accompanied by a detailed explanation, including the structural definition as well the literature source for any matched alert, a description of the mechanistic rationale for the alert as well as data on historically tested chemicals that also contained the alert.

17. I have a positive result what should I do now?

**Answer:** If negative bacterial mutagenicity data is available for the compound then this information would override a positive prediction. The draft guidelines mention that if the (Q)SAR is positive, but the alerting substructure is contained in the API (or related substance) in the same chemical environment and it has tested negative, then the alert or (Q)SAR prediction could be refuted. In addition, the draft guideline also allows for an expert review that could refute the positive prediction with sufficient supporting evidence (see Amberg et al., 2016). Otherwise, the compound would be predicted positive and should be controlled at or below acceptable limits (generic or adjusted TTC) or a bacterial mutagenicity test should be performed.

18. Does the Leadscope Model Applier automatically generate reports? What sort of information is available in these reports?

**Answer:** Yes. The report for the expert alerts contains a picture of the test chemical(s) highlighting the matched alert(s), the structural definition and source(s) for the alert, chemicals with historical Ames data containing the alert as well as any structural analogs with Ames data from the reference set. The statistical QSAR report includes a picture of the chemical structure(s), a description of the models that were used, a table of the results, an explanation of how the results were calculated (including a presentation of the structural feature responsible for the positive or negative prediction), and a listing of the structural analogs from the training set. In addition, it is also possible to include information on the consensus M7 call along with the reasoning for the consensus call.

Leadscope also includes a regulatory submission tool to generate a report that could be included in a regulatory submission. This tool integrates public and proprietary data, generates predictions from any necessary (Q)SAR models, incorporates tools to support the generation of expert opinions that may support or refute any prediction results. The tools generate a Word or PDF document containing a table of the results as well as supporting information such as information on any laboratory tests used in the assessment and any expert opinions. A series of demos are provided on the Leadscope website.

19. What documentation do I need to submit to regulatory agencies supporting the (Q)SAR impurities or degradants assessment?

**Answer:** The regulatory submission tool will automatically generate a complete report that outlines the materials and methods used, a summary of the results, and expert reviews. It also includes appendices containing full in silico reports and reports on the experimental data used in the assessment.

20. What is applicability domain? And how is the domain defined by Leadscope?

**Answer:** In (Q)SAR modeling the applicability domain represents the physico-chemical, structural or biological space of the predictive model. It is important that your test set fit in this space in order for a reliable prediction to be made. The Leadscope Model Applier evaluates the fit of your test set to a QSAR
model domain using two criteria - 1) the test chemical must have one of the predictive model chemical descriptors (as well as all of the property descriptors) and 2) the test chemical must have at least one structural analog in the model training set. If both of these criteria are not met, no prediction is generated and the test chemical is labeled as “Not in Domain”. The alert system also assesses the applicability domain based on structurally similar compounds in the alert’s reference set (using step 2 above).

21. **On rare occasions the Leadscope Model Applier does not calculate a prediction and reports “Out of Domain” or “Indeterminate”. How should I interpret this for assessing impurities and degradants? Can I turn off domain assessment and what are the implications?**

**Answer:** In the model applier, an overall M7 consensus call is calculated based on the expert alerts system and QSAR model results as well as available experimental data. A series of consensus rules were developed with input from the FDA under a Research Collaboration Agreement which weighs negative, positive, out of domain, and indeterminate calls in calculating an overall consensus call. Alternatively, a conservative approach would be to treat an out-of-domain or indeterminate as positive; however, supporting evidence such as the results from an analog search of the Leadscope SAR Genetox database could be used to assign it as negative. Such evidence could also be used to form an expert review supporting or refuting a consensus prediction. (see Amberg et al., 2016).

It is not recommended to turn off the domain assessment since there is limited information in the QSAR training set or the alerts’ domain to support a prediction of any chemicals identified as out-of-domain.

22. **How do I interpret seemingly conflicting results, for instance what if the Salmonella Mut model predicts positive but the E. coli – Sal 102 A-T model predicts either negative or out-of-domain?**

**Answer:** A positive result in either the Salmonella Mut or the E. coli – Sal 102 A-T Mut models would result in a positive call (see question 17 and 18 for handling out-of-domain predictions). If a Salmonella negative and E. coli negative or E. coli out-of-domain (or equivocal) predictions were made, there is more evidence to assign the result as an overall negative. The M7 consensus call takes into account potentially conflicting prediction results as well as the availability of experimental data to arrive at an overall call. In all cases, expert analysis weighing all available experimental and predicted evidence should be performed. Supporting evidence from analog searching of the Leadscope SAR Genetox database, training sets, and customer experimental data can be used to further support an assignment.

23. **Can the Leadscope Model Applier support the need to provide “…additional supportive evidence...”?**

**Answer:** Yes. Included with the Leadscope Model Applier is a high quality toxicity data base that can be used to identify structural analogs that provide additional supportive evidence. In addition, there is an explain capability that identifies what portions of the compound were used to calculate positive or negative results. The expert alerts are accompanied by an explanation of the mechanistic basis for each alert.

24. **How much does the Leadscope Model Applier cost?**
**Answer:** The Leadscope Model Applier – Genetic Toxicity Statistical QSAR Suite can be licensed either via an unlimited use license or a pay-per-compound license. The unlimited license Genetox Statistical QSAR Suite is $20,000 per year, for unlimited users and for unlimited predictions. It is truly an organization wide unlimited license. The pay-per-compound license for the Genetic Toxicity or Rodent Carcinogenicity Suites is $150 per compound, per suite.

As an additional feature, an organization can include an analog search for their compounds when they make a prediction with the Genetox Suite using either the unlimited license or the pay-per-compound license. The analog searching function utilizes the Leadscope SAR Genetox database.

The Leadscope Model Applier – Genetox Expert Alerts Suite can be licensed either via an unlimited use license or a pay-per-compound license. The unlimited license for the expert alerts is $20,000 per year, for unlimited users and for unlimited predictions. It is truly an organization wide unlimited license. The pay-per-compound license for the Genetox Expert Alerts Suite is $150 per compound.

A 10% discount is provided if a customer licenses both the Genetox Statistical QSAR Suite and the Genetox Expert Alerts Suite through an unlimited license ($36,000 annual for both after discount). The discounted pay-per-compound license for both the Genetox Statistical and Genetox Expert Alerts is $250 per compound.

25. I have no training in using (Q)SAR models or alerts. How easy is the Leadscope Model Applier to operate?

**Answer:** Once you have drawn a chemical structure or have access to a SMILES string for the chemical, running the software is easy. Simply paste in a chemical structure, select the models and alerts you want to run (or click on the “Use ICH M7 Setting” button), and then view the results in a customizable table. Reports can be generated directly from this results screen.

26. Does Leadscope Inc. cooperate with Toxicology consultants who could run the Leadscope Model Applier over a list of impurities and degradants and who could assist in interpreting my prediction results and preparing a report for regulatory authorities?

**Answer:** If a company chooses not to run their own predictions, Leadscope can run the predictions for companies for impurities and degradants in accordance with the ICH M7 guidelines. Reports will be provided by Leadscope and can be included in regulatory submissions. In addition, the Leadscope Model Applier includes QMFR reports recommended for REACH submissions. The Model Applier system can generate both the QRMF and QRPF reports automatically.

Leadscope also partners with over 25 consulting firms that utilize the Leadscope Model Applier in making predictions. Our consultants also have the ability to provide toxicology and chemistry interpretation of the results of predictions. Companies have utilized the reporting from Leadscope Model Applier in submissions to the U.S. FDA and other regulatory agencies.

27. How well do the Leadscope’s Genetox Expert Alerts work?
The Leadscope version 4.0 Genetox Expert Alerts have been validated against two data sets: the expert alerts reference set of approximately 10,000 compounds (chemicals from the Leadscope SAR genetox database and the training sets described in question 12) and the Hansen set of over 3,700 compounds (described in question 28).

The results against the reference set are:

<table>
<thead>
<tr>
<th>Metric</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concordance</td>
<td>86%</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>85%</td>
</tr>
<tr>
<td>Specificity</td>
<td>87%</td>
</tr>
<tr>
<td>Positive Predictivity</td>
<td>88%</td>
</tr>
<tr>
<td>Negative Predictivity</td>
<td>85%</td>
</tr>
<tr>
<td>Coverage</td>
<td>92%</td>
</tr>
</tbody>
</table>

Table 1.

The results against the Hansen set are presented below and compared with those from Derek Nexus run against the same test set. The U.S. FDA presented the results of an external validation of Derek Nexus in a poster presented at the SOT in 2013 (L. Stavitskaya, B. L. Minnier, R. D. Benz, N. L. Kruhlak, FDA Center for Drug Evaluation and Research, SOT 2013 “Development of Improved Salmonella Mutagenicity QSAR Models Using Structural Fingerprints of Known Toxicophores”).

<table>
<thead>
<tr>
<th>Model</th>
<th>Leadscope Genetox Alerts v4.0</th>
<th>Lhasa Derek Nexus v3.0.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concordance</td>
<td>83%</td>
<td>73%</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>92%</td>
<td>75%</td>
</tr>
<tr>
<td>Specificity</td>
<td>70%</td>
<td>69%</td>
</tr>
<tr>
<td>Positive Predictivity</td>
<td>81%</td>
<td>77%</td>
</tr>
<tr>
<td>Negative Predictivity</td>
<td>85%</td>
<td>67%</td>
</tr>
<tr>
<td>Coverage</td>
<td>94%</td>
<td>96%</td>
</tr>
</tbody>
</table>

Table 2.

28. Have the Leadscope Statistical QSAR models been externally validated and what are the results?

Answer: The U.S. FDA presented the results of an external validation of the Salmonella Mut model in a poster presented at the SOT in 2013 (L. Stavitskaya, B. L. Minnier, R. D. Benz, N. L. Kruhlak, FDA Center for Drug Evaluation and Research, SOT 2013 “Development of Improved Salmonella Mutagenicity QSAR Models Using Structural Fingerprints of Known Toxicophores”). Two data sets were used: the Hansen
set and a set from a commercial database (Leadscope external validation). The Hansen set is comprised of public data collected by Hansen et al. from the scientific literature. The entire set contained 6,512 compounds; however, 2,680 were in the training set or were stereo or geometric isomers of structures already in the training set. A further 132 were perceived duplicates within the test set or un-modelable structures, which were removed, leaving a total of 3,700 compounds in the final test set. The Leadscope external validation set is comprised of non-proprietary data harvested from FDA approval packages and the published literature. The entire set contained 3,005 compounds; however, 719 were in the training set or were stereo or geometric isomers of structures already in the training set, or were perceived duplicates within the set. These 719 structures were removed leaving a total of 2,286 compounds in the final test set. The following table from the 2013 SOT poster below summarizes the Leadscope QSAR performance statistics:

<table>
<thead>
<tr>
<th></th>
<th>Leadscope Enterprise</th>
<th>Hansen Validation</th>
<th>Leadscope Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coverage</td>
<td>N/A</td>
<td>87%</td>
<td>86%</td>
</tr>
<tr>
<td>Specificity</td>
<td>89%</td>
<td>66%</td>
<td>79%</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>79%</td>
<td>82%</td>
<td>78%</td>
</tr>
<tr>
<td>Concordance</td>
<td>84%</td>
<td>76%</td>
<td>78%</td>
</tr>
<tr>
<td>Positive Predictivity</td>
<td>84%</td>
<td>78%</td>
<td>76%</td>
</tr>
<tr>
<td>Negative Predictivity</td>
<td>85%</td>
<td>73%</td>
<td>81%</td>
</tr>
</tbody>
</table>

Table 3.

The *E. coli* – *Sal 102 A-T Mut* model does not have an external validation set for validation. However, results from the leave 10% out cross-validation were performed by the FDA and presented at the 2013 Genetic Toxicology Association meeting (L. Stavitskaya, B. L. Minnier, R. D. Benz, N. L. Kruhlak, FDA Center for Drug Evaluation and Research, “Development of Improved QSAR Models for Predicting A-T Base Pair Mutations”, GTA 2013b poster). The performance is summarized below*:

<table>
<thead>
<tr>
<th></th>
<th>Leadscope (%)</th>
<th>CASE Ultra (%)</th>
<th>Sarah Nexus (%)</th>
<th>Derek Nexus (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coverage</td>
<td>92.7</td>
<td>84.0 ± 3.5</td>
<td>88.8</td>
<td>[96.3]</td>
</tr>
<tr>
<td>Specificity</td>
<td>87.2 ± 0.1</td>
<td>79.8 ± 7.2</td>
<td>83.9</td>
<td>[79.1]</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>72.7 ± 0.2</td>
<td>68.5 ± 10.4</td>
<td>63.1</td>
<td>71.4</td>
</tr>
<tr>
<td>Concordance</td>
<td>80.9 ± 0.1</td>
<td>75.0 ± 6.8</td>
<td>74.9</td>
<td>[75.9]</td>
</tr>
<tr>
<td>Positive Predictivity</td>
<td>81.5 ± 0.1</td>
<td>72.2 ± 8.8</td>
<td>74.7</td>
<td>70.6</td>
</tr>
<tr>
<td>Negative Predictivity</td>
<td>80.6 ± 0.1</td>
<td>77.2 ± 6.4</td>
<td>75.1</td>
<td>[79.7]</td>
</tr>
</tbody>
</table>
The cross-validation studies are carried out using slightly different methodologies developed by each software provider and therefore are not directly comparable.

29. What are the performance statistics of combining a Leadscope Statistical QSAR prediction with another system?

**Answer:** The following table reports the performance statistics for the latest Leadscope statistical QSAR models and genetox expert alerts as well as the M7 consensus call for the Hansen test set. Note that the M7 consensus call includes rules for combining alerts, salmonella, and E. coli predictions as well consideration of any experimental data found in Leadscope databases:

<table>
<thead>
<tr>
<th></th>
<th>Leadscope Genetox Expert Alerts</th>
<th>Leadscope Statistical Salmonella QSAR</th>
<th>M7 Consensus Call</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concordance</td>
<td>83%</td>
<td>76%</td>
<td>92%</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>92%</td>
<td>82%</td>
<td>97%</td>
</tr>
<tr>
<td>Specificity</td>
<td>70%</td>
<td>68%</td>
<td>84%</td>
</tr>
<tr>
<td>Positive Predictivity</td>
<td>81%</td>
<td>78%</td>
<td>90%</td>
</tr>
<tr>
<td>Negative Predictivity</td>
<td>85%</td>
<td>73%</td>
<td>96%</td>
</tr>
<tr>
<td>Coverage</td>
<td>94%</td>
<td>87%</td>
<td>99%</td>
</tr>
</tbody>
</table>

The following table, presented by the U.S. FDA at the 2013 SOT meeting summarizes the performance of combining the Leadscope Statistical QSAR models (LSE) with Derek Nexus alerts from Lhasa Limited (DX) and CASE Ultra from MultiCASE (CU) using the external validation sets discussed in question 28 (L. Stavitskaya, B. L. Minnier, R. D. Benz, N. L. Kruhlak, *FDA Center for Drug Evaluation and Research, SOT 2013 “Development of Improved Salmonella Mutagenicity QSAR Models Using Structural Fingerprints of Known Toxicophores”*).
<table>
<thead>
<tr>
<th></th>
<th>Hansen External Validation</th>
<th>Leadscope External Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DX and LSE (%)</td>
<td>DX and CU (%)</td>
</tr>
<tr>
<td>Coverage</td>
<td>93%</td>
<td>94%</td>
</tr>
<tr>
<td>Specificity</td>
<td>57%</td>
<td>49%</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>89%</td>
<td>90%</td>
</tr>
<tr>
<td>Concordance</td>
<td>76%</td>
<td>73%</td>
</tr>
<tr>
<td>Positive Predictivity</td>
<td>75%</td>
<td>71%</td>
</tr>
<tr>
<td>Negative Predictivity</td>
<td>79%</td>
<td>77%</td>
</tr>
</tbody>
</table>

Table 6.

30. What options are available for training or support?

**Answer:** Leadscope provides unlimited training and support for the product, including on-site or online training, tutorials and telephone/email support at no additional cost.
Appendix A: Development and Analysis of the Leadscope *Salmonella* QSAR Model for Analyzing Pharmaceutical Impurities

**Development and Analysis of the Leadscope *Salmonella* QSAR Model for Analyzing Pharmaceutical Impurities**

*Leadscope, Inc.*
*April 2013*

**U.S. FDA and ICH Regulatory Context**

- The *Salmonella* mutagenicity assay is used in the drug review process to assess the genotoxic potential of APIs, as well as impurities, metabolites, and degradants.
- Current FDA/CDER guidance and EU guidelines state that, in some instances, (Q)SAR model predictions of *Salmonella* mutagenicity may be used in place of an experimental assay.
- Once finalized, the ICH M7 guidance will supersede FDA/CDER and EU guidance and define more explicit terms how a (Q)SAR analysis may be used to qualify a drug impurity.
- Under the ICH M7 guidance draft (step 2), public comment language, sponsors may submit (Q)SAR analyses performed using commercially available models or in-house proprietary models.
- If a sponsor submits the results of a (Q)SAR analysis as part of a drug application under ICH M7 (step 2), they must provide sufficient supporting documentation to assure regulators that the analyses were adequately performed and appropriately interpreted.

*L. Stavitskaya, O. L. Minner, P. D. Benz, N. L. Kruthiek,*
*FDA Center for Drug Evaluation and Research, SOT 2013*
**QSAR Model Building Process**

- **Building a model**
  - Construct training set
  - Generate Descriptors
  - Select Descriptors
  - Generate Model
  - Refine
  - Internal Validation
  - External Test set
  - Validation, Results, and Explanation

- **Using the model**

---

**Salmonella QSAR Model (version 3)**

- Construction of the QSAR Training data set
  - Training set contains 3979 non-proprietary compounds.
  - Data were harvested by the FDA from internal and external sources.
  - Overall *Salmonella* calls were assigned by the FDA.
  - The positive-to-negative ratio of calls in the training sets is 43%-to-57%.
Salmonella QSAR Model (version 3)

- Construction of the QSAR Model Descriptor Set
  - Substructural features from the Leadscope native feature set of 27,142 medicinal chemistry building blocks were scored for predictivity against the training set and the most predictive features were selected\(^1\).
  - Scaffolds of predictive larger features were algorithmically assembled from the Leadscope features training set\(^2\).
  - A set of published mechanistic-based genotoxic and carcinogenicity alerts were added to the feature set.
  - The feature set was automatically pruned and similar features removed.
  - The final feature set was manually pruned resulting in 369 features (193 positive and 176 negative features).

\(^1\)Ning, C. et al, J. Med. Chem. 2004, 47, 5984-5994

**Leadscope Salmonella QSAR Model**

**Selected Positive Features**

Legend:
- Frequency
- Z-score
- Sensitivity
- Precision

Leadscope *Salmonella* QSAR Model (version 3)

- **Generation of the QSAR Model**
  - Partial Logistic Regression of the training set created a model for prediction of overall binary *Salmonella* calls
  - 369 substructural features and 7 continuous calculated properties were included in the model
  - 5 features were extracted
  - 10 fold cross-validation (leave 10% out) was run multiple times to validate the model, select the best factorcount, and avoid over-training the model
  - Additional validation of the model was performed by FDA using two separate external validation sets and its performance assessed against other models by other vendors constructed using the same training set

**Definition of the External Validation Sets used to Validate the *Salmonella* QSAR Model**

- The **Hansen set** is comprised of non-proprietary data collected by Hansen *et al.* from the scientific literature. The entire set contained 6512 compounds; however, 2680 were in the training set or were stereo or geometric isomers of structures already in the training set. A further 132 were perceived duplicates within the test set or un-modelable structures, which were removed, leaving a total of 3700 compounds in the final test set.

- The **Leadscope external validation set** is comprised of non-proprietary data harvested from FDA approval packages and the published literature. The entire set contained 3005 compounds; however, 719 were in the training set or were stereo or geometric isomers of structures already in the training set, or were duplicates or perceived duplicates within the set. These 719 structures were removed leaving a total of 2286 compounds in the final test set.

L. Stavitskaya, B. L. Mihler, R. D. Benz, N. L. Krutikak
FDA Center for Drug Evaluation and Research, SOT 2013
### Salmonella QSAR Model (version 3)

Cross-validation and External Validation Set Performance of Individual Vendor Models

<table>
<thead>
<tr>
<th></th>
<th>Leadscape Enterprise</th>
<th>CASE Ultra</th>
<th>Derek Nexus</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cross-validation</td>
<td>Hansen</td>
<td>Leadscape</td>
</tr>
<tr>
<td>Coverage</td>
<td>N/A</td>
<td>87%</td>
<td>86%</td>
</tr>
<tr>
<td>Specificity</td>
<td>89%</td>
<td>66%</td>
<td>79%</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>79%</td>
<td>82%</td>
<td>76%</td>
</tr>
<tr>
<td>Concordance</td>
<td>84%</td>
<td>76%</td>
<td>76%</td>
</tr>
<tr>
<td>Positive Predictiv</td>
<td>84%</td>
<td>78%</td>
<td>76%</td>
</tr>
<tr>
<td>Negative Predictiv</td>
<td>85%</td>
<td>73%</td>
<td>81%</td>
</tr>
</tbody>
</table>

- Hansen external validation set of 5700 compounds (43% negatives, 57% positives)
- Leadscape external validation test of 2220 compounds (54% negatives, 46% positives)

Leadscape outperforms Case-Ultra and Derek Nexus using both Hansen and Leadscape external validation sets. Positive and Negative Predictivity statistic comparisons are highlighted.

### Salmonella QSAR Model (version 3)

Cross-validation and External Validation Set Performance of Combined Vendor Models

<table>
<thead>
<tr>
<th></th>
<th>Hansen External Validation</th>
<th>Leadscape External Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DX and LSE (%)</td>
<td>DX and CI (%)</td>
</tr>
<tr>
<td>Coverage</td>
<td>93%</td>
<td>54%</td>
</tr>
<tr>
<td>Specificity</td>
<td>57%</td>
<td>49%</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>89%</td>
<td>90%</td>
</tr>
<tr>
<td>Concordance</td>
<td>76%</td>
<td>74%</td>
</tr>
<tr>
<td>Positive Predictiv</td>
<td>75%</td>
<td>71%</td>
</tr>
<tr>
<td>Negative Predictiv</td>
<td>78%</td>
<td>77%</td>
</tr>
</tbody>
</table>

LSE – Leadscape, DX – LHASA Derek Nexus, FDA Center for Drug Evaluation and Research, SOT 2013

Using Leadscape in combination with Derek outperforms Case-Ultra with Derek. Positive and Negative Predictivity statistic comparisons are highlighted. "Any positive as a consensus positive" results in improvements in consensus sensitivity at the expense of specificity and positive predictivity due to a corresponding increase in false positives.
Analog Browsing Leadscope Databases

- Leadscope Genetox and Carcinogenicity Databases are now provided with the purchase of those model suites respectively in Leadscope Model Applier and Leadscope Enterprise.
- Now two variations of searching analogs of a compound structure:
  - Searching for analogs in the training set (via the Explain feature)
  - Searching for analogs in the database(s) (via the Browse Analogs feature)
  - Not all training set structures are in the Leadscope database and vice-versa.
  - All known Leadscope database and FDA training set data are integrated together for compounds in common.