Getting under the skin of an in silico approach to predicting dermal sensitisation

Donna S. Macmillan & Martyn L. Chilton
Granary Wharf House, 2 Canal Wharf, Leeds, LS11 5PS

Abstract no. 68

Introduction

Derek Nexus is an expert toxicity prediction tool established by Lhasa Limited in 1983. It uses structure-activity relationships and reasoning rules developed by Lhasa experts to predict over 70 different toxicity endpoints including genotoxicity, carcinogenicity, skin sensitisation, HEIR channel inhibition, hepatotoxicity and irritation. Skin sensitisation has been a key endpoint over the past few years, in part due to the implementation of EU regulation 1233/2009 which prohibits the sale and marketing of any cosmetics and cosmetic ingredients which have been tested on animals, alongside REACH4 and CLP5 which state that non-animal methods must be used to consider the use of animal tests. The use of in silico tools such as Derek Nexus is increasingly popular due to their rapid toxicity assessment of chemicals, transparent predictions and access to the wealth of toxicity data and mechanistic information provided.

Method

Skin sensitisation alerts - The performance of the Derek Nexus skin sensitisation alerts from 2014-2018 was assessed using an in-house dataset of 1243 sensitisers and 1300 non-sensitisers (n = 2543) with murine local lymph node assay (LLNA) and/or guinea pig data (consensual call applied when results for both were present). Chemicals activating an alert with a likelihood of equivocal or above were classified as sensitisers. Chemicals activating an alert with a likelihood of improbable or non-alerting chemicals were classified as non-sensitizers. The following metrics were calculated: sensitivity (Se), (TP/[TP+FP]*100); specificity (Sp), (TN/[TN+FN]*100); positive predictivity (PP), (TP/[TP+FP]*100); negative predictivity (NP), (TN/[TN+FN]*100); accuracy (Acc), ([TP+TN]/[TP+FP+FN+TN]*100).

EC3 prediction model – a k-Nearest Neighbours model was developed based on a curated in-house database of over 1000 publicly available, LLNA studies. The model was validated using 103 previously unseen chemicals with LLNA data4.

Integrated testing strategy (ITS) - a previously published dataset of 213 compounds with LLNA data and in chemico/in vitro data (DPRA, n = 194; KeratinoSens, n = 187; LLUs, n = 78; h-CLAT, n = 186; U-SENS, n = 149) was used, with Derek predictions and physicochemical parameters, to develop a decision tree for ITS-1 and ITS-2.

Conclusion

Skin sensitisation alerts in Derek currently perform with an accuracy of 76%, a sensitivity of 80% and a specificity of 72%. The alerts are reviewed regularly, particularly if new public data is sourced, leading to alert refinement and/or the development of new alerts e.g. alerts 867, 878-879 and 882 (Figure 2). When proprietary data is donated, the data are anonymised and used to expand alert coverage, allowing all Derek members to benefit from improved alert predictivity and/or new alerts.

Skin sensitisation alerts

Derek alerts use structure-activity relationships (SAR) created by Lhasa experts to predict the toxicity of a given chemical. The predictions are supported by a graphical explanation of the SAR, mechanistic rationale, toxicity data of known compounds within the SAR and key references. Public, proprietary and regulatory data are used to build the alerts thereby providing extensive coverage of chemical space. The number and performance of the skin sensitisation alerts has increased steadily from 2014-2018 (Figure 1). The alerts are reviewed regularly, particularly if new public data is sourced, leading to alert refinement and/or the development of new alerts e.g. alerts 867, 878-879 and 882 (Figure 2). When proprietary data is donated, the data are anonymised and used to expand alert coverage, allowing all Derek members to benefit from improved alert predictivity and/or new alerts.

Case study: Refinement of substituted phenol alert

The scope of alert 439, substituted phenol or precursor, was investigated for any improvements to increase its predictivity. It was found that substituted phenols produced mixed positive and negative toxicity data in mouse, guinea pig and human assays, whereas a specific sub-class of vinyl or allylic anilines is consistently positive in the same assays. Consequently, a new alert, with enhanced predictivity, was created for these anilines (alert 867) and alert 439 refined to exclude these, leading to a slight improvement in performance (Table 1).

EC3 prediction model validation

The EC3 prediction model was evaluated using an external validation dataset (n = 103), consisting of LLNA data donated to Lhasa Limited by a number of members for the specific purpose of providing a test set of unseen compounds. The experimental EC3 values in this dataset were compared to the EC3s predicted by the model and were judged as correct when the experimental (Exp) and predicted (Pred) values were within a factor of 3 of each other, even if the data were not classified as a GHS category (Figure 3). For all predictions (Exp=Pred), the predicted EC3 was the same as the GHS category (same boxplot, Figure 3). The alerts are reviewed regularly, particularly if new public data is sourced, leading to alert refinement and/or the development of new alerts e.g. alerts 867, 878-879 and 882 (Figure 2). When proprietary data is donated, the data are anonymised and used to expand alert coverage, allowing all Derek members to benefit from improved alert predictivity and/or new alerts.

Table 1. Refinement of alert 439 by excluding mixed positive and negative anilines.

<table>
<thead>
<tr>
<th>Alert</th>
<th>Se (%)</th>
<th>Sp (%)</th>
<th>Acc (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>439</td>
<td>80</td>
<td>90</td>
<td>85</td>
</tr>
<tr>
<td>867</td>
<td>60</td>
<td>95</td>
<td>75</td>
</tr>
<tr>
<td>882</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Comparison of integrated testing strategies (ITS) defined approaches (DA)

Comparison of integrated testing strategies (ITS)defined approaches (DA)

It is generally accepted that no in chemico or in vitro assay can be used as a standalone method to replace animal models for the prediction of skin sensitisation potential. The focus has instead turned to combining multiple assays and/or molecular descriptors to derive a more accurate assessment of hazard or risk. These ITS are also known as DA and are key elements within integrated approaches to testing and assessment (AAT4) for skin sensitisation, used for regulatory decision-making. Lhasa’s first approach (ITS-1) used a Derek prediction, assigned the test chemical as in or out of the applicability domain of the in chemico/in vitro assay (based on physicochemical parameters), alongside up to two in chemico/in vitro assays to predict sensitiser/non-sensitiser (S/Ns). Lhasa’s subsequent, as yet unpublished, ITS-2 is similar to ITS-1 (Figure 6) but (1) utilises even more information from Derek, including alert likelihood and the new negative prediction functionality and (2) additional reactivity properties are also considered when defining the applicability domain. Finally, (3) the hazard (S/Ns) is predicted alongside an estimate of GHS category, based on data in the EC3 prediction model. ITS-2 improves significantly upon the specificity of ITS-1, with only a minor reduction in sensitivity (Table 2).

Outcome - Detailed overview of skin sensitisation potential of the query chemical including alerting features, toxicity data, mechanistic rationale, EC3 prediction, and how Derek can be used effectively in an ITS