How to apply expert review to EC3 predictions in Derek Nexus
Skin sensitisation

Defined as the process of a chemical, known as a hapten, causing an allergic reaction which leads to allergic contact dermatitis (ACD)\(^1\).

ACD develops in 2 stages:

- **Stage 1**, induction where the hapten forms a complex with the skin protein and initiates a cascade ending with proliferation of allergen specific T cells.
- **Stage 2**, elicitation, occurs after subsequent contact with the same allergen. This leads to the hapten-skin complex triggering the same allergen specific T cells, inducing inflammatory cytokines and ACD.

This mechanism of skin sensitisation has been studied extensively and has a well-defined Adverse Outcome Pathway (AOP)\(^2\).

An AOP is the sequence of events that leads to a particular in vivo outcome of interest and the following AOP (Figure 1), published by the OECD in 2012, covers each event in the pathway of skin sensitisation caused by covalent binding of chemicals to skin proteins in detail.

There are 4 key events. The 1\(^{\text{st}}\) key event, the molecular initiating event (MIE), is the binding of the chemical hapten to skin protein. The 2\(^{\text{nd}}\) and 3\(^{\text{rd}}\) key events are cellular responses by keratinocytes and dendritic cells. The 4\(^{\text{th}}\) key event is defined as the organ response where T-cells are activated and then finally the adverse outcome of skin sensitisation is reached.

![Image of AOP](image.png)

*Figure 1. Skin sensitisation Adverse Outcome Pathway (adapted from OECD 2012).*
In vivo assays

The murine local lymph node assay (LLNA)\(^1\) and guinea pig maximisation test (GPMT)\(^4\) are the most widely used in vivo assays to identify sensitising chemicals. However, some chemicals are intrinsically more potent than others and it has been suggested that factors such as chemical reactivity and hydrophobicity may play a role in determining the potency of specific classes. Consequently, differentiating only between sensitising and non-sensitising chemicals can lead to suboptimal risk assessments, which in some cases may be overprotective or too lax.

In this respect, the LLNA offers a significant advantage over the GPMT. While the GPMT is typically used only to identify chemicals that represent a sensitising hazard, the LLNA provides information on the potency of an identified sensitiser, thereby permitting an assessment of both hazard and potency. Potency is inferred from the LLNA using the test concentration required to cause a 3-fold increase in the stimulation index (SI) relative to vehicle controls which is referred to as the EC3 value\(^5\). EC3 values have proven reliable measures of potency and several studies have shown a good correlation between EC3 values and known sensitisation potency in humans\(^6-8\).

EC3 classifications

Two EC3 classification schemes are in common use (Table 1 and Table 2): the European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC) and a Globally Harmonized System of Classification and Labelling (GHS). Both of these classification schemes assign categories to chemicals based on their potency (EC3 value) to aid risk assessment.

<table>
<thead>
<tr>
<th>ECETOC Classification</th>
<th>GHS Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EC3 (%)</strong></td>
<td><strong>Category</strong></td>
</tr>
<tr>
<td>≥10</td>
<td>Weak sensitiser</td>
</tr>
<tr>
<td>≥1-&lt;10</td>
<td>Moderate sensitiser</td>
</tr>
<tr>
<td>≥0.1-&lt;1</td>
<td>Strong sensitiser</td>
</tr>
<tr>
<td>&lt;0.1</td>
<td>Extreme sensitiser</td>
</tr>
</tbody>
</table>

Table 1 and Table 2. ECETOC and GHS EC3 classification systems.
Non-animal assays and ITS/DIP

Despite the success of the LLNA to detect skin sensitisers, regulatory changes and animal welfare concerns have resulted in a significant drive to develop non-animal alternatives. As of early 2017, there were three in chemico/in vitro assays for skin sensitisation validated by ECVAM and associated Test Guidelines from the OECD published:

- **DPRA**: an in chemico assay which assesses the 1st key event, the molecular initiating event, in the skin sensitisation AOP
- **KeratinoSens™**: an in vitro assay which assesses the 2nd key event in the skin sensitisation AOP
- **h-CLAT**: an in vitro assay which assesses the 3rd key event in the skin sensitisation AOP

However, it is generally accepted that no single non-animal assay will fully replace in vivo tests and the focus has now turned to combining results from multiple in chemico and in vitro assays and in silico tools in Integrated Testing Strategies (ITS, also known as Data Integration Procedures (DIP)) (Figure 2). Lhasa Limited has published on using Derek Nexus as part of an in silico/in chemico/in vitro ITS to predict skin sensitisation hazard (sensitiser/non-sensitiser) and is working to improve its skin sensitisation endpoint by developing in silico predictions of skin sensitisation potency.

![Figure 2. Schematic of ITS/DIP (Integrated Testing Strategies/Data Integration Procedures).](image-url)
**In silico EC3 predictions using Derek**

An ideal in silico tool would provide reliable predictions of skin sensitisation hazard but also be able to accurately predict potency. With this goal in mind, an EC3 prediction model has been developed based on publicly available LLNA studies and mechanistic (alert) domains as provided by Derek Nexus, an expert knowledge base system for toxicity predictions developed by Lhasa Limited.

The model utilises a k-Nearest Neighbours (kNN) approach. Initially, both compounds (query and dataset) must activate the same skin sensitisation structural alert in Derek to be considered as a Nearest Neighbour. These Nearest Neighbours are then arranged by Tanimoto similarity generated by an in-house radial fingerprinting method. The minimum number of neighbours required is 3, otherwise no prediction is given, and the most similar neighbours, up to $10^{th}$ place, are considered (Figure 3). The predicted EC3 value is the weighted average of all the valid neighbours. This approach has been validated using an internal test set ($n = 45$) and an external data set of proprietary member data ($n = 103$)\textsuperscript{13}.

**Expert assessment of EC3 predictions**

Expert assessment can help improve the quality of EC3 predictions given by Derek Nexus.

The model is built on a dataset consisting of over 1000 high quality LLNA studies.

The initial prediction given by the model may be improved by applying expert review for the following reasons:

- There is significant variability associated with the LLNA (approx. 3-fold).
- The alert domain used to calculate the alert may have a wide spread of EC3 values.
- An outlier may have a disproportionate effect on the EC3 prediction.
- A subset of the alert domain may be more relevant to a prediction of a specific query compound.

![Figure 3. Example of 10 Nearest Neighbours used in a typical EC3 prediction.](image)

This workbook provides a number of worked examples of performing expert assessment on EC3 predictions in Derek Nexus. All expert calls made on the examples are solely the opinion of experts at Lhasa Limited and are made here for guidance purposes only.
\( (E)-1-(\text{cyclohex-1-en-1-yl})\text{but-2-en-1-one} \)

SMILES | O=C(/C=C/C)C1=CCCCC1
---|---
logP | 3.13 (BioByte Corp, version 5.9)
MW | 150.22 (Lhasa Limited, version 1.0)
Alert | 480: alpha,beta-Unsaturated ketone
Predicted EC3 | 0.0048%
Look at EC3 data in the alert activated by the query compound

13 data points in alert space.

EC3 values dispersed over multiple ECETOC and GHS categories.

Predicted LLNA EC3: 0.0048% (extreme sensitiser) - [Derek EC3 Model - 1.0.6]
Number of similar compounds used in the calculation: 10/13

Highlight similar/dissimilar compounds and/or outlier(s)

Majority of compounds in alert space have an EC3 value between 1% - 10%.

One compound in particular is significantly more potent than any other compound in the alert space (EC3 = 3x10^{-4}%).
Analyse compound(s) of interest

This compound has a cyclopropenone group, not present in any other compound in this alert space. Small carbocycles like this have high ring strain due to bond angle deviation from the standard tetrahedral 109.5°. This increases the reactivity of this alpha,beta-unsaturated ketone substantially and may explain the potency of this compound compared to others within the same alert space.

Furthermore, this compound has only been tested in one LLNA study - result may not necessarily be reproducible.

Suggest addition/removal of compounds based on chemical and mechanistic knowledge

Due to:

• the potency of this compound and

• the potential for increased reactivity compared to others within the same alert space

removal of this compound should give a more accurate EC3 prediction.
Assess expert-reviewed EC3 prediction

Expert-reviewed EC3 prediction more in line with other compounds in the same alert space.

Original EC3 prediction: 0.0048%

Expert-reviewed EC3 prediction: 2.7%

Predicted EC3 is relatively close to GHS boundary.

User may prefer to use conservative approach and treat chemical as 1A sensitiser.

Notes:

Number of similar compounds used in the calculation: 9/13

Predicted LLNA EC3: 2.7% (moderate sensitiser) - [Derek EC3 Model - 1.0.6]
3-(4-ethylphenyl)propanal

![Chemical structure](image)

<table>
<thead>
<tr>
<th>SMILES</th>
<th>C1=C(C=CC(=C1)CC)CCC=O</th>
</tr>
</thead>
<tbody>
<tr>
<td>logP</td>
<td>2.9 (BioByte Corp., version 5.9)</td>
</tr>
<tr>
<td>MW</td>
<td>162.23 (Lhasa Limited, version 1.0)</td>
</tr>
<tr>
<td>Alert</td>
<td>419: Aldehyde</td>
</tr>
<tr>
<td>Predicted EC3</td>
<td>4.4%</td>
</tr>
</tbody>
</table>

- **L**: Look at EC3 data in the alert activated by the query compound
- **H**: Highlight similar/dissimilar compounds and/or outlier(s)
- **A**: Analyse compound(s)
- **S**: Suggest addition/removal from EC3 calculation based on chemical and mechanistic knowledge
- **A**: Assess new EC3 prediction
Look at EC3 data in the alert activated by the query compound

26 data points in alert space.

EC3 values mainly weak or moderate (based on ECETOC categories).

Highlight similar/dissimilar compounds and/or outlier(s)

Cluster of 5 chemicals with high Tanimoto similarity to query compound.

Small cluster of two compounds also used in default Lhasa EC3 prediction.
Analyse compound(s) of interest

**Cluster of 5 chemicals** - subset of the query compound. Alpha methyl group adjacent to the aldehyde shouldn’t impact chemical reactivity so no need to remove these from the calculation.

**Cluster of 2 chemicals** - one carbon shorter in chain length. Shouldn’t impact chemical reactivity of aldehyde.

**Alpha,beta-unsaturated ketone** - May undergo alternative mechanism due to presence of Michael acceptor.

Suggest addition/removal of compounds based on chemical and mechanistic knowledge

Keep cluster of 5 chemicals with high Tanimoto similarity to query compound as they are extremely similar.

Small cluster of two compounds kept as reduction in alkyl chain length should not impact aldehyde reactivity.

Final three used in default prediction removed:

- Alpha,beta-unsaturated ketone because of alternative mechanism.
- Final two as not required with 7 very similar compounds in EC3 prediction.
Assess expert-reviewed EC3 prediction

Expert-reviewed EC3 prediction more in line with other compounds in the same alert space.

Original EC3 prediction: 4.4%

Expert-reviewed EC3 prediction: 8.3%

Predicted EC3 is borderline between weak and moderate ECETOC categories.

User may prefer to use conservative approach and treat chemical as a moderate sensitiser.

Notes:

Assessed expert-reviewed EC3 prediction.

Expert-reviewed EC3 prediction more in line with other compounds in the same alert space.

Original EC3 prediction: 4.4%

Expert-reviewed EC3 prediction: 8.3%

Predicted EC3 is borderline between weak and moderate ECETOC categories.

User may prefer to use conservative approach and treat chemical as a moderate sensitiser.

Notes:
5-methylbenzene-1,3-diol

![Structural formula of 5-methylbenzene-1,3-diol](image)

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMILES</td>
<td>OC1=CC(C)=CC(O)=C1</td>
</tr>
<tr>
<td>logP</td>
<td>1.31 (BioByte Corp., version 5.9)</td>
</tr>
<tr>
<td>MW</td>
<td>124.14 (Lhasa Limited, version 1.0)</td>
</tr>
<tr>
<td>Alert</td>
<td>440: Resorcinol or precursor</td>
</tr>
<tr>
<td>Predicted EC3</td>
<td>2.0%</td>
</tr>
</tbody>
</table>

- L: Look at EC3 data in the alert activated by the query compound
- H: Highlight similar/dissimilar compounds and/or outlier(s)
- A: Analyse compound(s)
- S: Suggest addition/removal from EC3 calculation based on chemical and mechanistic knowledge
- A: Assess new EC3 prediction
Look at EC3 data in the alert activated by the query compound

6 data points in alert space.

EC3 values dispersed over multiple ECETOC and GHS categories.

Highlight similar/dissimilar compounds and/or outlier(s)

One compound very similar to query.

2/6 data points are experimental non-sensitisers.

One compound is slightly more potent than any other compound in the alert space (EC3 = 0.49%).
**Analyse compound(s) of interest**

**Small cluster of non-sensitisers** - Anisole motif present instead of two hydroxyl groups. May have an impact on chemical reactivity.

**Alpha,beta-unsaturated ketone** - May undergo alternative mechanism due to presence of Michael acceptor.

<table>
<thead>
<tr>
<th>Similar Compounds</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Chemical structures" /></td>
</tr>
<tr>
<td>LLNA EC3: 52% (weak sensitizer)</td>
</tr>
<tr>
<td>LLNA EC3: 58% (moderate sensitizer)</td>
</tr>
<tr>
<td>LLNA EC3: 55% (moderate sensitizer)</td>
</tr>
<tr>
<td>LLNA EC3: 34% (strong sensitizer)</td>
</tr>
</tbody>
</table>

**Suggest addition/removal of compounds based on chemical and mechanistic knowledge**

Cluster of non-sensitisers removed from EC3 prediction as anisole may be responsible for reduced activity of these compounds compared to the query compound.

Alpha,beta-unsaturated ketone removed from EC3 prediction as it may undergo an alternative mechanism due to presence of Michael acceptor.

Other three compounds in alert domain kept and used in EC3 prediction.

<table>
<thead>
<tr>
<th>Similar Compounds</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image2" alt="Chemical structures" /></td>
</tr>
<tr>
<td>LLNA EC3: 52% (weak sensitizer)</td>
</tr>
<tr>
<td>LLNA EC3: 58% (moderate sensitizer)</td>
</tr>
<tr>
<td>LLNA EC3: 55% (moderate sensitizer)</td>
</tr>
<tr>
<td>LLNA EC3: 34% (strong sensitizer)</td>
</tr>
</tbody>
</table>
Assess expert-reviewed EC3 prediction

Expert-reviewed EC3 prediction has changed GHS category from 1A to 1B.

Original EC3 prediction: 2.0%

Expert-reviewed EC3 prediction: 7.3%

Predicted EC3 is moderate sensitizer although is close to ECETOC category border.

Notes:

Assess expert-reviewed EC3 prediction

Expert-reviewed EC3 prediction has changed GHS category from 1A to 1B.

Original EC3 prediction: 2.0%

Expert-reviewed EC3 prediction: 7.3%

Predicted EC3 is moderate sensitizer although is close to ECETOC category border.

Notes:

Predicted LLNA EC3: 7.3% (moderate sensitizer) - [Derek EC3 Model - 1.0.6]
Number of similar compounds used in the calculation: 3/6
3-((2,3,4-trichlorophenyl)amino)oxetan-2-one

---

**SMILES**

C1(C(CO1)NC2=CC=C(C(=C2Cl)Cl)Cl)=O

**logP**

3.18 (BioByte Corp., version 5.9)

**MW**

266.51 (Lhasa Limited, version 1.0)

**Alert**

411: Ring-strained amide, ester, thioamide or thioester

**Predicted EC3**

1.8%

---

**L**

Look at EC3 data in the alert activated by the query compound

**H**

Highlight similar/dissimilar compounds and/or outlier(s)

**A**

Analyse compound(s)

**S**

Suggest addition/removal from EC3 calculation based on chemical and mechanistic knowledge

**A**

Assess new EC3 prediction
Look at EC3 data in the alert activated by the query compound

Only 3 data points in alert space.

EC3 values dispersed over multiple ECETOC and GHS categories.

Predicted LLNA EC3: 1.8% (moderate sensitizer) - [Derek EC3 Model - 1.0.6]
Number of similar compounds used in the calculation: 3/3

In the absence of similar compounds in the alert space:
The remaining expert review steps cannot be carried out (see below).

How much confidence can be had in an EC3 prediction based on these three Nearest Neighbours?

- Analyse compound(s) of interest
- Suggest addition/removal of compounds based on chemical and mechanistic knowledge
- Assess expert-reviewed EC3 prediction
References


10. OECD. 2015. Test No. 442D: In Vitro Skin Sensitisation: ARE-Nrf2 Luciferase Test Method. DOI: 10.1787/9789264229832-en

11. OECD. 2016. Test No. 442E: In Vitro Skin Sensitisation DOI: 10.1787/9789264264359-en

