Predicting skin sensitisation using a decision tree integrated testing strategy with an *in silico* model and *in chemico/in vitro* assays.

Donna S. Macmillan, Steven J. Canipa, Martyn L. Chilton, Richard V. Williams and Christopher G. Barber.
Lhasa Limited, Granary Wharf House, 2 Canal Wharf, Leeds, LS11 5PS
Email: donna.macmillan @lhasalimited.org

■ Introduction

Skin sensitisation is a common occupational disease with a well-understood adverse outcome pathway (AOP)[1]. As a result of the implementation of EU regulation 1223/2009[2] prohibiting the use of animals for cosmetic ingredient safety assessments (e.g. murine local lymph node assay (LLNA) or the guinea pig maximisation test (GPMT)) a number of non-animal assays assessing different key events in the AOP have been designed. However, two main classes of compounds fall outside the applicability domain of these assays, lipophilic compounds and pre-/pro-haptens, and are usually not predicted well[3-5]. Using multiple assays, as an integrated testing strategy (ITS), may improve predictive performance by combining results from individual assays and/or use molecular descriptors to derive an overall assessment of hazard or risk. In principle, Derek Nexus is capable of modelling the entire AOP and skin sensitisation alerts can take lipophilicity and pre-/pro-haptens into account as alerts are expert-derived and have a mechanism-based domain. To this end, an integrated testing strategy (ITS) using Derek Nexus and a maximum of two *in chemico/in vitro* assays (from DPRA, KeratinoSens, LuSens, h-CLAT and U-SENS) has been developed[6].
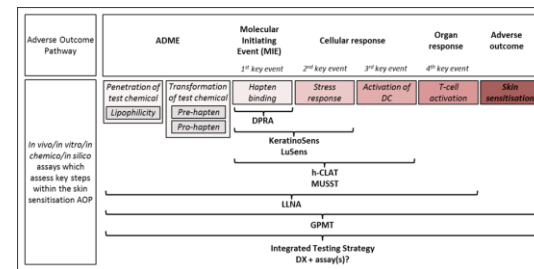
Skin sensitisation AOP. Adapted from OECD 2012[1].

■ Method

213 compounds (Urbisch *et al.*) with LLNA data and *in chemico/in vitro* test data (DPRA, 194 compounds; KeratinoSens, 187 compounds; LuSens, 78 compounds; h-CLAT, 166 compounds; U-SENS, 149 compounds) were used to design a decision tree integrated testing strategy. Assessment of *in chemico/in vitro* sensitisation potential was based on the author call, however, conflicting and/or borderline results were removed from each data set. These were then processed against skin sensitisation using Derek Nexus 4.1.0 (using Knowledge Base Derek 2014 1.0). Certain/probable/plausible/equivocal predictions were considered as sensitisers. Where no alerts were fired, compounds were considered as non-sensitisers. The predictivity of single assays or combinations were calculated according to Cooper statistics (Cooper *et al.*)[8] and the mean of each assay combination/type used for comparison.

■ Decision tree construction

A decision tree was designed by Lhasa using the LLNA, DPRA, KeratinoSens, LuSens, h-CLAT and U-SENS data collated by Urbish *et al.*, (*n* = 78-213 dependent on assay)[7], guided by the following principles:

- *Derek Nexus covers the entire skin sensitisation AOP.*
- *Reduce unnecessary in chemico/in vitro assays where possible.*
- *Concordant Derek Nexus and in chemico/in vitro results = reliable result - no additional assays required.*
- *Discordant Derek Nexus and in chemico/in vitro results = unclear result - run another assay and use majority call.*
- *Consider the in chemico/in vitro applicability domain.*
- *Maximise the negative predictivity.*

How to use the Lhasa decision tree

- Process through Derek Nexus.
- Assign domain.
  - In domain = logP < 3.5 and hapten.
  - Out of domain = logP > 3.5 and pre/pro-hapten.
- Run 1st *in chemico/in vitro* assay (DPRA, KeratinoSens, LuSens, h-CLAT, U-SENS).
- Use decision tree to assign call or run an additional assay before assigning call.
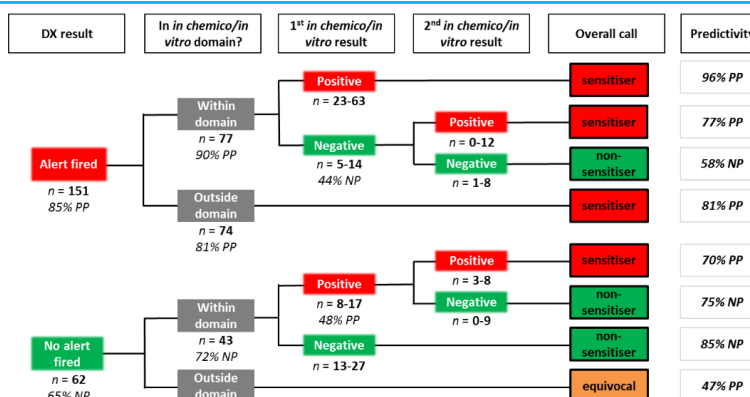
DX = Derek Nexus. n = number of compounds in each decision tree branch. n varies based on which in chemico/in vitro assay is used for the 1st and/or 2nd assay result. See Supporting Information in Macmillan et al. for individual combinations.

■ Performance of decision tree against other ITS

The performance of the Lhasa decision tree was compared to the mean of the following: single assay results, a 1 out of 2 conservative call, a 2 out of 3 weight of evidence call and a 3 out of 3 concordance call using all combinations of Derek Nexus, DPRA, KeratinoSens, LuSens, h-CLAT and/or U-SENS.
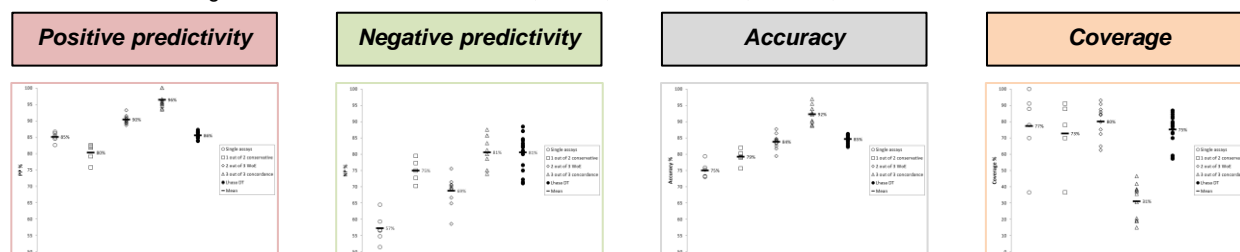
| Assay type	PP	NP	Acc	Cov
Single assays	85	57	75	77
1 out of 2 conservative	80	75	79	73
2 out of 3 WoE	90	69	84	80
3 out of 3 concordance	96	81	92	31
Lhasa DT	**86**	**81**	**85**	**75**

Positive predictivity (TP/[TP + FP]*100); negative predictivity (TN/[TN + FN]*100); accuracy: (TP + TN/[TP + TN + FP + FN]*100); coverage: (number of compounds which can be predicted by assay or ITS/total number of compounds in data set). The mean value of all assay combinations were used to calculate the metrics.

■ Decision tree significantly reduces number of assays required for overall call

We compared the number of assays required for the Lhasa decision tree ITS against a 2 out of 3 weight of evidence ITS using **Derek Nexus, DPRA** as the first assay and **KeratinoSens** as the second assay.

- 74 test chemicals fired an alert but were outside the *in chemico/in vitro* assay domain - assigned as sensitisers.
- 90 test chemicals were within the assay domain and were assigned an overall call based on the concordant Derek Nexus and DPRA result.
- 19 chemicals within this data set had a discordant Derek Nexus and DPRA result, and required a KeratinoSens result to assign an overall call.
- 13 compounds were non-alerting in Derek Nexus and were outside the assay domain - these were assigned as equivocal.
- 17 compounds could not be evaluated as they lacked data for either DPRA or KeratinoSens.

For this example, between 426 and 639 assays would be required for a 2 out of 3 weight of evidence ITS, whereas the Lhasa decision tree ITS requires only 128 assays (1 assay x 90 compounds, 2 assays x 19 compounds). This reduction in assays required would have a significant impact on resources such as time and cost. Other combinations of Derek Nexus, DPRA and KeratinoSens are shown in the graph.
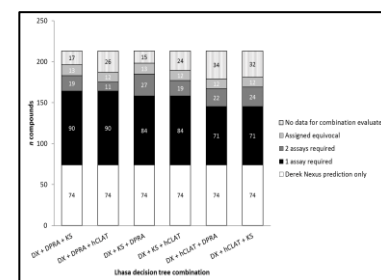
DX = Derek Nexus. KS = KeratinoSens.

■ Conclusions and further work

- This ITS displays high concordance with the LLNA and the positive predictivity is high (86%) and has superior negative predictivity (81%) when compared to the other ITS evaluated in this study.
- The use of an *in silico* model in an ITS can extend the applicability domain to include chemicals poorly predicted by *in chemico/in vitro* assays and may help address some of their limitations e.g. poor prediction of metals, insolubility of some test chemicals, lack of metabolic capacity within assays.
- The Lhasa decision tree ITS is a time- and cost-efficient method to assess skin sensitisation.

References: 1. OECD 2012, The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins Part 1: Scientific Evidence. 2. European Union, 2013, Off. J. Eur. Union, 56, 34–66. 3. OECD, 2015, Test No. 442C: In Chemico Skin Sensitisation: Direct Peptide Reactivity Assay (DPRA). 4. OECD, 2015, Test No. 442D: In Vitro Skin Sensitisation: ARE-Nrf2 Luciferase Test Method. 5. OECD, 2014, Draft Guideline: In Vitro Skin Sensitisation: human Cell Line Activation Test (h-CLAT). 6. Macmillan *et al.*, 2016 Reg. Toxicol. Pharmacol., 76, 30-38. 7. Urbisch *et al.*, 2015, Regul. Toxicol. Pharmacol., 71, 337–351. 8. Cooper, J.A., Saracci, R., Cole, P., 1979, Br. J. Cancer, 39, 87–89.