Nothing to report?

An *in silico* fragmentation methodology for making explicit negative predictions of skin sensitisation

Martyn Chilton
Scientist
Martyn.Chilton@lhasalimited.org
Overview

• Background
  • Structural alerts for skin sensitisation
  • Making negative predictions within a knowledge-based system
  • Can negative predictions be made for skin sensitisation?

• Results
  • Constructing the reference dataset
  • Performance of the negative predictions model
  • The importance of expert review

• Conclusions
Structural alerts for skin sensitisation

- Structure-activity relationships for skin sensitisation have been known for many years
- Lhasa Limited have built up a knowledge base containing structural alerts for skin sensitisation over the past 25 years
- The 90 alerts in Derek Nexus include information about:
  - the **scope** of the alert
  - the suspected **mechanism** of sensitisation
  - literature **references** supporting the alert
  - examples of chemicals with **toxicity data** that fire the alert
  - the **performance** of the alert against validation datasets
Making negative predictions within a knowledge-based system

For mutagenicity *in vitro*:

- Reactivity-based mechanism (DNA adduct formation)
- Fragments can capture reactivity well
- Sufficiently large dataset available
Making negative predictions within a knowledge-based system

Reference Ames dataset

$\text{n} = 10243$

<table>
<thead>
<tr>
<th>TP</th>
<th>FP</th>
</tr>
</thead>
<tbody>
<tr>
<td>FN</td>
<td>TN</td>
</tr>
</tbody>
</table>

Derek

Classify fragments

Fragment library

Query compound

Alerting

Concerning fragments $\rightarrow$ misclassified
Unknown fragments $\rightarrow$ unclassified

R Williams et al., It’s difficult, but important, to make negative predictions, *Regul. Toxicol. Pharmacol.* 2016, 76, 79-86
Can negative predictions be made for skin sensitisation?

**Opportunities**

- The mechanism of skin sensitisation is also largely reactivity driven (protein adduct formation)
- There is a good alert coverage for this endpoint

**Challenges**

- Data comes from multiple species/assays
- Predicting an *in vivo* adverse outcome rather than the result of an *in vitro* assay
- How much data is available?
# Constructing the reference dataset

### Literature
- Human data: $n = 315$
- Animal data: $n = 2994$

### Human data
- Baskette potency categories (1-6)
- BgVV potency categories (A-C)

### Animal data
- Local Lymph Node Assay (LLNA)
- Guinea Pig Maximisation Test
- Non-standard LLNA
- Non-radioactive LLNA
- Adjuvant tests
- Optimisation tests
- Epicutaneous tests
- Mouse ear swelling test
- Buehler test
- Draize test

### Reference dataset
- $n = 2768$

### Flowchart
1. Literature
2. Human data
3. Combine and curate
4. Categorise assays
5. Assign overall call
6. Reference dataset

### Distribution
- Human: 11%
- Standard animal: 78%
- Non-standard animal: 9%
- Other animal: 2%
Performance of the negative predictions model

Negative Predictivity (%)

- Non-alerting:
- Non-sensitiser:
- Non-sensitiser but contains misclassified features
- Non-sensitiser but contains unclassified features
**The importance of expert review**

- **Misclassified feature:** how similar to the query chemical are the non-alerting sensitisers found in the reference dataset?

- **Unclassified feature:** are there other compounds with sensitisation data that contain the same unclassified feature as the query chemical?
The importance of expert review

NON-SENSITISER but contains missclassified feature(s)

Longifolene
CAS = 475-20-7
LLNA positive
EC3 = 1.8% (moderate)
Probable prehapten
The importance of expert review

NON-SENSITISER but contains unclassified feature(s)

9-[bis(2-Hydroxyethyl) amino]-5-[(3-methylphenyl) amino]-7-phenyl-benzo[a] phenazinium chloride
LLNA negative
Maximum Tested Dose = 20% as a suspension in 7:3 EtOH/H₂O (not usable at higher concentrations)
Conclusions

• A fragmentation approach has been applied to make *in silico* negative predictions of skin sensitisation
  
  - Built on an *in vivo* dataset containing >2500 chemicals
  - Better performance than simply considering a lack of structural alerts as a negative prediction
  - Enables expert review by highlighting any previously unseen fragments, as well as those of possible concern

• Planning to validate the approach through data sharing
  
  - Recently 8 companies have shared LLNA data with us
  - Aiming to produce a joint publication with the results
  - Interested in talking to anyone else who has sensitisation data (guinea pig, mouse or human) that they may be willing to share
Acknowledgements

• Donna Macmillan
• Rich Williams
• Susanne Stalford
• The various Lhasa members who have recently shared skin sensitisation data with us
Thank you for your attention

Any questions?