Developing predictive *in silico* models for liver toxicity endpoints from *in vivo* histopathology data

Pinto-Gil K, Gregori-Puigjane E, Pasamontes-Funez I, Gómez-Tamayo JC, Stela B, Sanz F and Pastor M.

Research Programme on Biomedical Informatics (GRIB), FIMIM, Barcelona, Spain.

**Facts & Figures**

- **Start date:** 01/01/2010
- **End date:** 31/12/2016
- **Contributions**
  - **IMI funding:** 6,910,018 €
  - **EFPIA in kind:** 10,157,590 €
  - **Other:** 1,719,500 €
  - **Total Cost:** 18,877,109 €
- **Project website:** www.etoxproject.eu

**Social media:** www.youtube.com/watch?v=fMWpDUaOdNc

**Challenge**

- **The eTOX project compiled a collection of nearly 9000 of such studies.**
- **In this work we describe how these collected data can be exploited for the development of predictive models for *in vivo* toxicity endpoints.**

**Approach & Methodology**

- **Normalization**
  - Using existing as well as ad-hoc developed ontologies

- **Filtering**
  - Selecting a consistent, comparable set of studies

- **Aggregation**
  - Merging studies to characterize compound properties

- **Scoring**
  - Using finding profiles to characterize organ toxicity endpoints

**How can we walk this path?**

- **DEG**
  - **Positive compounds:** 164
  - **Negative compounds:** 168

- **INF**
  - **Positive compounds:** 94
  - **Negative compounds:** 164

- **PRO**
  - **Positive compounds:** 82
  - **Negative compounds:** 164

**Representation of the confusion matrix, sensitivity and specificity for the best qualitative models obtained using conformal Random Forest Classifier.**

**Quality parameters of conformal model predictions.**

<table>
<thead>
<tr>
<th>Endpoint</th>
<th>Internal prediction (cross-validated)</th>
<th>External prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coverage</td>
<td>MCC</td>
</tr>
<tr>
<td>DEG</td>
<td>0.47</td>
<td>0.55</td>
</tr>
<tr>
<td>INF</td>
<td>0.74</td>
<td>0.64</td>
</tr>
<tr>
<td>PRO</td>
<td>0.64</td>
<td>0.45</td>
</tr>
</tbody>
</table>

**In general, the quality of the models is acceptable, if we consider the complexity of the *in vivo* endpoints, representing many different mechanisms of liver toxicity.**

**Value of IMI collaboration**

- **The results shown here are only an example of how the results of the eTOX project (the eTOX database and eTOXsys) are being exploited for practical purposes.**
- **The eTOX database represents a successful example of precompetitive sharing of information between 13 pharmaceutical companies, made possible thanks to the public-private partnership IMI project.**

**Impact & take home message**

- **The information present in RDT reports generated for regulatory purposes can be transformed to obtain data amenable for developing predictive models.**
- **These models exemplify the practical use of the *in vivo* data collected by the eTOX project.**
- **Precompetitive collaboration can produce valuables resources which can be exploited to obtaining safer and more effective drugs, in a faster way.**

This work has received support from the EU/EFPIA/Innovative Medicines Initiative Joint Undertaking (eTOX grant no 15002).