

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

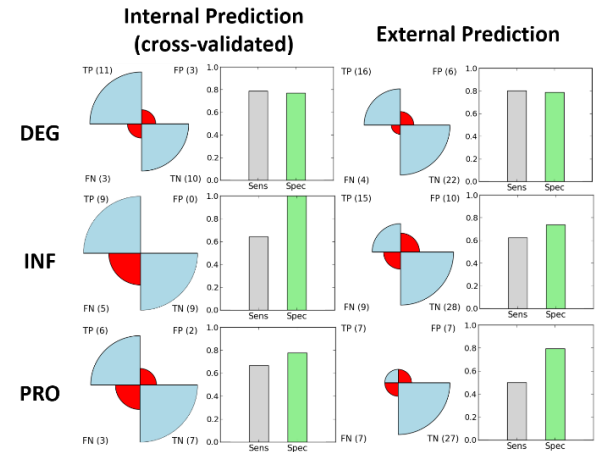
Start date:	01/01/2010
End date:	31/12/2016
Contributions	
IMI funding:	6,910,018 €
EFPIA in kind:	10,157,590 €
Other:	1 719 500 €
Total Cost:	18,877,109 €
Project website:	www.etoxproject.eu
Social media:	www.youtube.com/watch?v=fMWpDUaOdNc

Results

Distribution of qualitative hepatotoxicity scorings.

Liver <i>in vivo</i> Endpoint	Positive compounds	Negative compounds
Degeneration (DEG)	164	168
Inflammatory (INF)	94	164
Non-neoplastic Proliferative lesions (PRO)	82	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.

Endpoint	Internal prediction (cross-validation)		External prediction	
	Coverage	MCC	Coverage	MCC
DEG	0.47	0.55	0.48	0.58
INF	0.74	0.64	0.81	0.44
PRO	0.64	0.45	0.65	0.29

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.

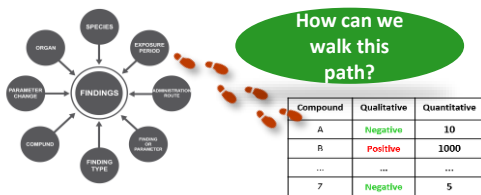
Impact & take home message

- 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 valuable resources which can be exploited to obtaining safer and more effective drugs, in a faster way.

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

