Nonclinical Working Group
Industry Insights

IMI and eTOX an overview

Phil Drew

PhUSE EU CSS – June 19th – 20th, 2017

What is IMI?

Innovative Medicines Initiative:

Part of the €3.2BN EU Horizon 2020 Project:

What is eTOX

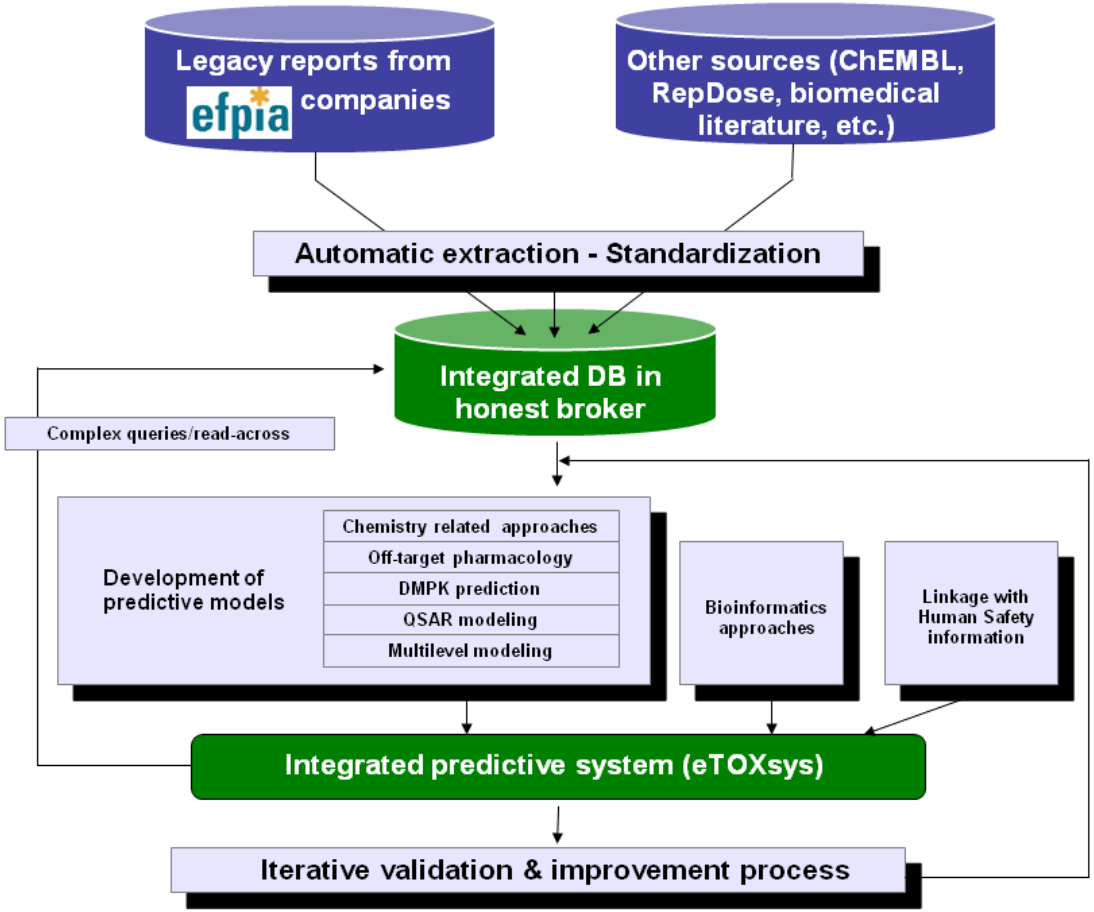
The eTOX consortium/project has developed a huge drug safety database from pharmaceutical industry legacy toxicology reports and from publicly available toxicology data. It has also developed innovative in silico strategies and novel software tools to better predict the toxicological profiles of small molecules in early stages of the drug development pipeline.

This IMI project is devoted to large-scale sharing, data-basing and mining of industry legacy tox reports and includes predictive models and has a central tool for database querying and running the predictive models. Furthermore, a series of purpose-fit ontologies were developed which are supported by the OntoBrowser (a project deliverable).

There is significant collaboration (partnering and shared funding) with ‘efpia’ Companies:

efpia
European Federation of Pharmaceutical Industries and Associations

Roadmap


Legacy reports from efpia companies
Other sources (ChEMBL, RepDose, biomedical literature, etc.)
Automatic extraction - Standardization
Integrated DB in honest broker
Complex queries/read-across
Development of predictive models
Chemistry related approaches
Off-target pharmacology
DMPK prediction
QSAR modeling
Multilevel modeling
Bioinformatics approaches
Linkage with Human Safety information
Integrated predictive system (eTOXsys)
Iterative validation & improvement process

TOOLS (free)

eTOXsys

An Integrated Platform (central tool) for Profiling and Data Mining Across Multiple Databases and Prediction Models (Lhasa Ltd).

Human Outcomes Module

A webservice that allows the search of similar marketed drugs from a simple SMILES code and retrieves adverse events data associated to these drugs from three sources currently implemented: literature (PubMed), EURETOS and LAERTES. The eTOXsys architecture was adapted to activate the eTOX Human Outcomes application directly from the eTOXsys GUI. The SMILES selected in eTOXsys is automatically passed on to the eTOX Human Outcomes prototype by means of a button that is shown in the new eTOXsys front-end.

TOOLS (free)

LIBRARY

Supplies easy access to a chronological selection of recommended Articles relevant to data related to drug toxicity, a list of selected Journals extracted from ISI Web of Knowledge with Toxicology filtering, and finally, a section of interesting Links, where all these resources are presented in alphabetic order and organized by categories. The information is updated in a monthly basis.

eTOXlab

A flexible modelling framework for supporting models predicting the biological properties of chemical compounds (e.g. QSAR models) in production environments. The eTOXlab-VM constitutes a portable, self-contained prediction engine that can be easily distributed and installed.

TOOLS (free)

ONTOBROWSER

The primary objective of OntoBrowser is to provide an easy-to-use online collaborative solution for subject matter experts to map reported terms to preferred ontology (or code list) terms and facilitate ontology evolution. Additional features include web service access to data, visualization of ontologies in hierarchical/graph format and a peer review/approval workflow with alerting.

Availability and implementation: The source code is freely available under the Apache v2.0 license. Source code and installation instructions are available at http://opensource.nibr.com. This software is designed to run on a Java EE application server and store data in a relational database.

TOOLS (free)

LiMTox

The LiMTox system is a text mining approach that extracts associations between compounds and toxicological end points at various levels of granularity and evidence types, all inspired by the content of toxicology reports. It integrates direct retrieval of associations between compounds and hepatotoxicity through combination of heterogeneous complementary strategies from term co-occurrences, linguistic rules, and text patterns to machine learning based text classification. It also returns relations of compounds and CYPs and compounds and biochemical markers (e.g. liver enzymes).

TOOLS (free)

Collector

Collector allows extraction from the Open PHACTS Discovery platform and the eTOX project series of compounds annotated with experimental data that can be used directly for building QSAR predictive models. Collector implements a flexible collection of filters for selecting the target, the biological data ranges and the physicochemical properties of the compounds.

Collector is installed locally and implements a local database for keeping record of the data extracted. The application provides a web based user a command line user interface.

TOOLS (free)

Standardiser

The Standardiser is a tool designed to provide a simple way of standardising molecules as a prelude to e.g. molecular modelling exercises. The tool is available on GitHub and is released under the Apache 2.0 license.

Verification

The verification of the models developed within the eTOX project comprises an assessment of (i) the quality of data used to build (and test) the model, (ii) whether the predictions generated by the model when executed in eTOXsys are consistent, and (iii) the completeness of the documentation accompanying the model. This process has been semi-automated by implementing the required checks into a KNIME workflow and results in a more efficient and standardised verification process.

Further Information

- This was a five-year project which started in 2010. It was extended for a further year and concluded in Late November 2016.
- Because data were (initially) taken from study reports, the eTOX database contains mainly summary data (i.e. group means/incidences).
- Digitizing report data proved to be slow, expensive and inaccurate.
- The majority of submitted studies were old and not particularly interesting (mainly because of fears around sharing company-sensitive information).
- More recently, SEND-formatted data have been imported and summarized. The original data are available as an output.
- Many data issues were encountered along the way which were not adequately predicted at the outset leading to a reduced number of studies. This has informed the design and planning of IMI2/NexGETS.
- Data contained in the eTOX database are automatically made available to efpia member companies and will be used as a seed resource in NexGETS.

Thank You

pds consultants