Statistics for the final eTOX database release v2016.3

Katharine A Briggs, Grace Kocks & David S Wilkinson.

Lhasa Limited, Leeds, United Kingdom

Contents of the eTOX database

The latest release of the eTOX database includes 1,464 non-confidential structures and 483 confidential structures associated with 8,196 studies. Data is captured for each dose and sex group evaluated in the study plus the control group and all time points where measurements or observations were taken, whether they show an increase, decrease or are unchanged.

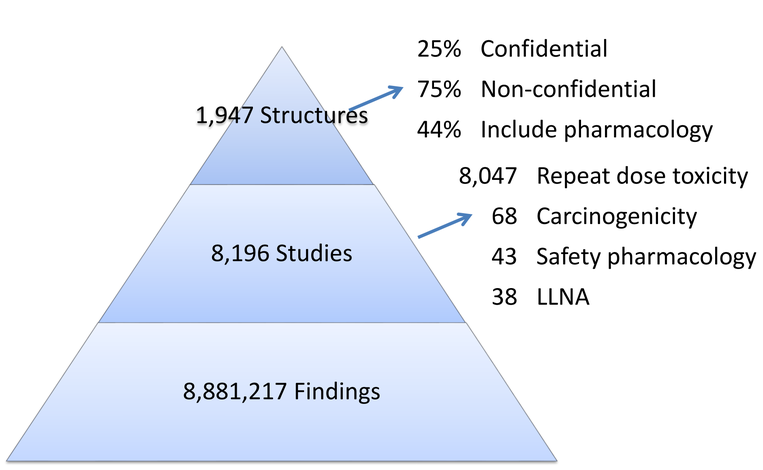

Figure 1: Contents of the eTOX database version 2016.3

Breakdown by species, route and study duration

The pie charts illustrate how the data is broken down in terms of species (i.e. mostly rat), administration route (i.e. mostly oral) and duration of study (i.e. mostly short term <20 days duration).

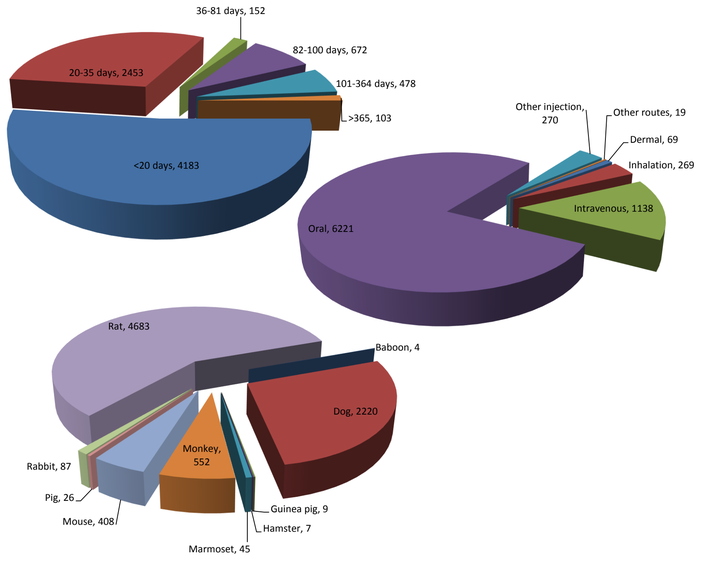

Figure 2: Breakdown by species, administration route and study duration (number of studies)

Chemical space representative of approved drugs

The graphic below compares the chemical space of the eTOX non-confidential structures [represented in blue] with approved drugs [represented in green] included in DrugBank (v 5.0.3) using StarDrop (v6.3.2) from Optibrium. Exact, substructure and chemical similarity queries are possible allowing data mining for structurally similar compounds.

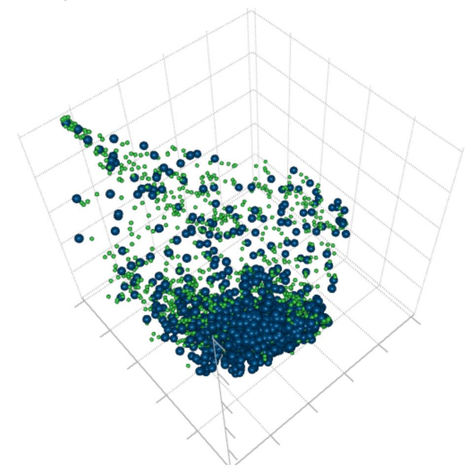

Figure 3: Chemical space of eTOX non-confidential data (blue) compared to approved drugs included in DrugBank v5.0.3 (green)

Explore target related effects

Around 44% of the substances in the eTOX database have been annotated to include their pharmacological target. In the latest release this data has been aggregated and the table below summarises the top 10 aggregated targets in the database. Inclusion of this data allows users to query the database for identical or related targets.

Aggregated target	Structures
Membrane receptor - GPCR - chemokine	55
Enzyme - kinase - SER/THR kinase	48
Enzyme - oxidoreductase	43
Ion channel	41
Membrane receptor - GPCR - serotonin receptor	41
Enzyme - kinase - receptor tyrosine kinase	36
Transcription factor - nuclear receptor	33
Transporter	31
Enzyme - transferase	30
Membrane receptor - GPCR - adrenergic receptor	29

Figure 4: Top 10 aggregated pharmacology targets included in the database

Explore treatment related findings

The graph here shows the top 10 organs affected by treatment related histopathology findings. Positive structures are those that have treatment related findings. Treatment related refers to the expert call generated and encoded within the original report and not one generated by a retrospective analysis of the raw data. Organ and histopathology findings are being captured as is using the verbatim terms in the reports and are then mapped retrospectively to the eTOX ontology which has also been developed within the project.

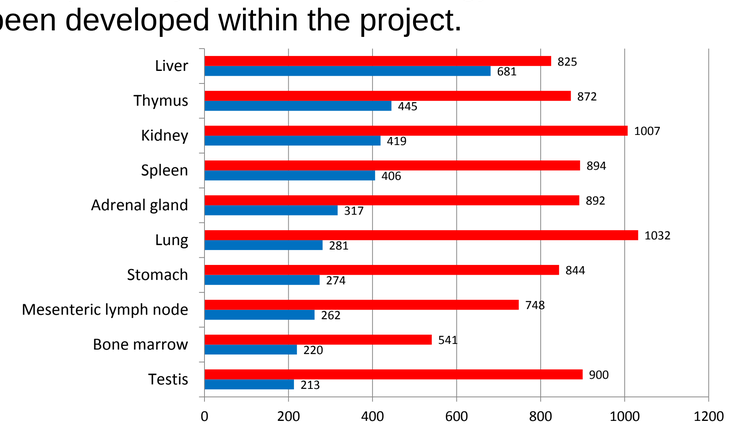

Figure 5: Top 10 organs with histopathology findings (blue = non treatment related & red = treatment related)

Source of historical control data

As data is captured for all dose groups including control groups the eTOX database also provides a useful resource for analysing historical control data. The box plot below illustrates the range of values reported for alanine aminotransferase (ALT) within control groups in three different strains of rat; Fischer 344, Sprague-Dawley and Wistar.

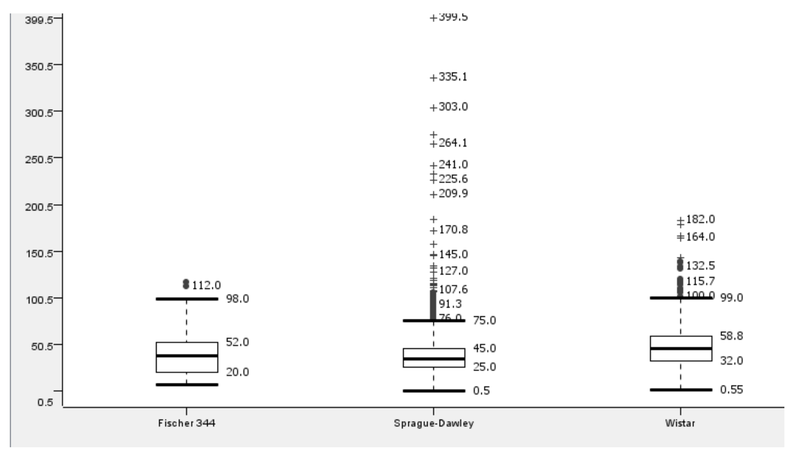

Figure 6: Box plot of alanine aminotransferase (ALT) values in control groups of 3 rat strains

Compare effects in different species

It is possible to compare effects observed in different species. The graph below illustrates the differences in treatment related histopathology reported in the liver for substances with studies conducted in both dog and rat. Within this dataset it is clear that the dog is less susceptible to hepatocyte hypertrophy.

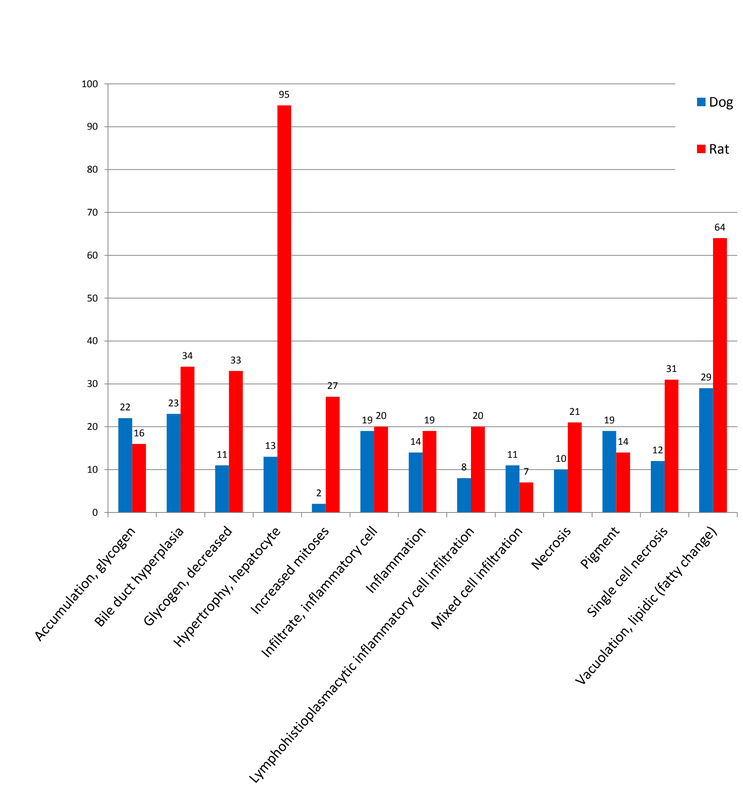

Figure 7: Incidence (number of substances) of treatment related liver histopathology findings reported for substances tested in both dog and rat

Compare study types

It is also possible to look at differences in the findings reported with studies of different duration. For instance treatment related histopathology reported in the liver for substances with studies of either short (20-35 days) or long (82-100 days) duration. Within this dataset the number of substances exhibiting hepatocyte hypertrophy is almost doubled in the longer studies.

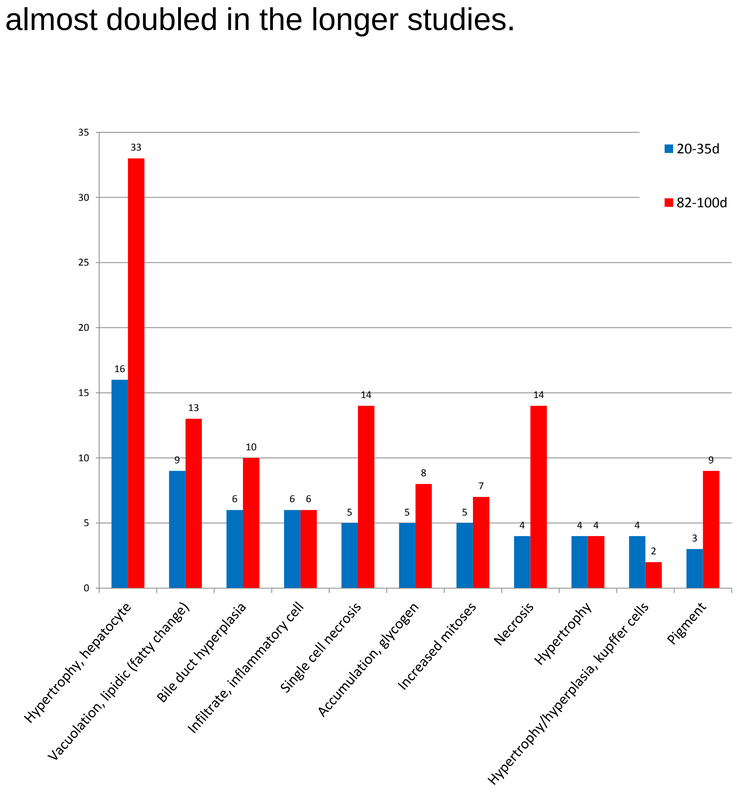

Figure 8: : Incidence (number of substances) of treatment related liver histopathology findings reported for substances tested in studies of 20-35 days and 82-100 days duration

Datasets for model building or validation

Sustainability plans have been developed to ensure that this valuable resource will continue to be supported and maintained post project. A large subset of the final eTOX database release will be made available through eTOXsys in early 2017 as an integrated but modular package. Providing access to highly relevant proprietary data that is not available elsewhere.

Access to other databases, such as RepDose from Fraunhofer, will be technically enabled as well as access to the ontologies developed during the project and the interface will allow data mining of the federated databases.

Well documented and verified predictive models developed during the project will also be available and users will be able to choose the ones they want.

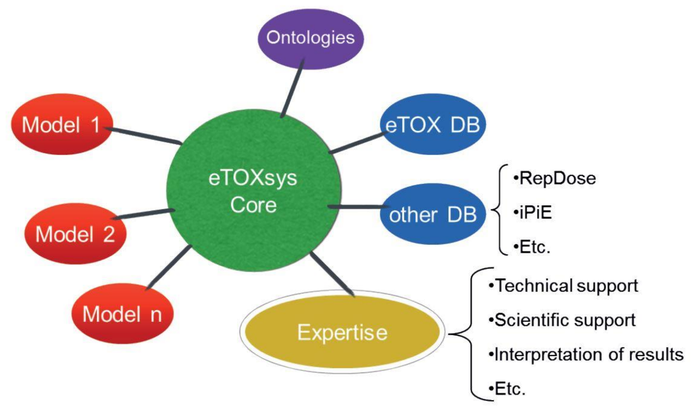

Figure 9: Modular approach for eTOXsys sustainability

Acknowledgement

The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement nº 115002 (eTOX), resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contributions.