ISSN 2191-6039

19. Jahrgang • März • 1/2017

pharmazeutische medizin

1

ZUR SACHE
The eTOX Consortium: To Improve the Safety Assessment of New Drug Candidates

PHARMAZEUTISCHE MEDIZIN
Computational Methods for the Prediction of Chemical Toxicity

RECHT
Vertragsgestaltung in der klinischen Forschung – Eine „Checkliste" für den Zentrumsvertrag

DGPharMed
Deutsche Gesellschaft für Pharmazeutische Medizin e. V.

Of the Art of Sharing the Unsharable

The eTOX Consortium: To Improve the Safety Assessment of New Drug Candidates

After their incorporation into human risk assessment, reports on preclinical animal studies in most cases get buried in the archives. The wealth of these data is hardly accessible. The European Innovative Medicines Initiative project eTOX ("electronic toxicity") over the last years designed and implemented a strategy for leveraging these preclinical data and sharing them across pharmaceutical companies and academic institutions. The shared toxicological data can now be used by the participating companies to perform early safety assessments of new drug candidates and new pharmacological targets. In addition, the data are used to build in silico predictive tools for specific toxicological endpoint.

Dr. Thomas Steger-Hartmann, Bayer AG, Investigational Toxicology, Berlin
Dr. François Pognan, PreClinical Safety, Novartis Institute for Biomedical Research, Basel

1. Preclinical toxicology for the prediction of human safety

Before a new drug candidate can be administered to human for the first time, the compound has to undergo preclinical toxicological assessment. Cornerstones of this evaluation are two repeated toxicity studies, which are required by law, e. g. in Europe by Directive 2001/83/EC and its respective national laws. Repeated dose toxicity tests shall be carried out in two mammal species, one of which must be a non-rodent (in most cases in dogs, but sometimes other species such as cynomolgus monkeys or mini-pigs can be used as they are biologically more relevant). The rodent species for such tests is usually rat.

The results of the two studies determine whether the risk-benefit profile of a new drug candidate is acceptable for the first use in humans. In addition, these studies are used to derive the safe starting dose in the first clinical trial and to identify parameters and biomarkers which need monitoring during the clinical phase.

The results of these systemic toxicity studies take a back seat shortly after the clinical phase simply because human data are considered more relevant than animal data. Only in the case of unexpected adverse events in humans, the toxicologist is confronted with the question "why the finding did not occur in the animal study" or "why a finding was not considered relevant for humans".

2. How predictive are animal studies?

The above described situation of an animal–human mismatch is often cited in the lay press, and sometimes in the scientific literature, to reject animal models as useful tool in biomedical research. For example, Shanks et al. (2009) conclude in their philosophical-ethical analysis work, that "… animal models (…) fall short of being able to predict human responses". In contrast to this statement, a group of toxicologists led by Harry Olson published in the year 2000 (Olson et al. 2000) a large retrospective analysis with data from 150 drug development programs from 12 pharmaceutical companies. The authors investigated the concordance of animal toxicity with human findings and came to the conclusion that in 71% of the cases, where a human toxicity was observed, the corresponding effect was also seen in comparable target organs in the preceding animal toxicity studies. The authors also observed that some toxicities, such as hematological, gastrointestinal or cardiovascular findings, correlated better than others. A weak correlation for example was observed for skin lesions.

This pivotal and often quoted "Olson et al. study" was a strenuous organizational effort, since neither the preclinical data nor the clinical data were easily accessible through databases and had to be compiled from numerous individual studies. This situation has not changed much since that time.

While the required toxicological studies have to be performed under Good Laboratory Practice (GLP) and as such have to be archived and retrievable for at least 12 years, a statistical analysis of correlation with clinical data

across compound classes and indications of these reports strong of up to 400 pages is only possible, if the data are electronically available. Some large pharma companies undertook endeavors to transfer their reports into internal databases, but uniform database standards or commercial data base solutions have not yet come into sight.

Besides an in-depth correlation analysis 'à la Olson' the value of such a preclinical database would also lie in the following aspects:

- Toxicological effects which are rare, would not only reside in the memory of an individual pathologist diagnosing the finding, but could be easily retrieved in the database for comparative purposes.
- For a specific toxicological endpoint, the data collection could eventually lead to the establishment of structure-activity (toxicity) relationship, similar to what has been done over the last two decades for the mutagenicity endpoints. Such predictive systems could then be applied to optimize the structure of a lead compound in the early phases before even starting in vivo studies.
- Comparisons on the basis of chemical structure or searches for effects on similar targets could help to raise hypotheses on target organs and for better differentiation between on-target und off-target effects.
- If relevant data are available in the database for similar targets or structures, animal studies could be better designed or even avoided, if necessary conclusions can already be drawn from previously acquired, so-called legacy data.

The biggest value of such a database would certainly emerge if it not only contained the data of an individual pharma company but would rather allow access to study results from other researching pharmaceutical companies. But is this a realistic vision given the competition between companies, i.e., can the un-sharable be shared?

3. eTOX – the development of a preclinical toxicity database

In order to explore, whether such a vision could be implemented, thirteen pharmaceutical companies, eleven academic partners and six small and medium-sized enterprises (SMEs) of the bioinformatics sector joined forces in the eTOX project ("electronic toxicity") which started in January 2010 under the umbrella of the European "Innovative Medicines Initiative" (IMI) (http://www.etoxproject.eu). The budget of the project amounted to 17.9 million, where 7 million was provided as funding to the beneficiaries (SMEs and academia). The bigger part of the budget was provided by the pharmaceutical companies as in-kind or cash contributions. The objective of the project was not only to develop a database, but also to explore the prerequisites for data sharing in the context of patent and intellectual property protection. Furthermore, procedures for efficient data extraction from the paper or pdf reports as well as strategies for building predictive tools based on the extracted data were developed (see figure 1).

The biggest challenge experienced in the starting phase of the project was the mutual access to proprietary data. Whereas the toxicologists involved in the project were convinced of the necessity of data sharing to achieve the goals of eTOX, there was significant reservation among the patent and research departments against the idea of exchanging structures and biological data previously not retrievable in the public domain. The arguments of losing competitive advantage, of having data potentially re-interpreted and in the worst case, opening possibilities for unlawful liability suits or litigations, weighed heavily and had to be refuted. This is of particular concern because once shared, a later withdrawal of data from joint database is a barely traceable issue.

A key step in the resolution of this challenge was the nomination of an independent data broker ("honest broker"; Lhasa Ltd., Leeds, UK), responsible for the IT security aspects, for setting up a legal framework of data sharing with all partners of the

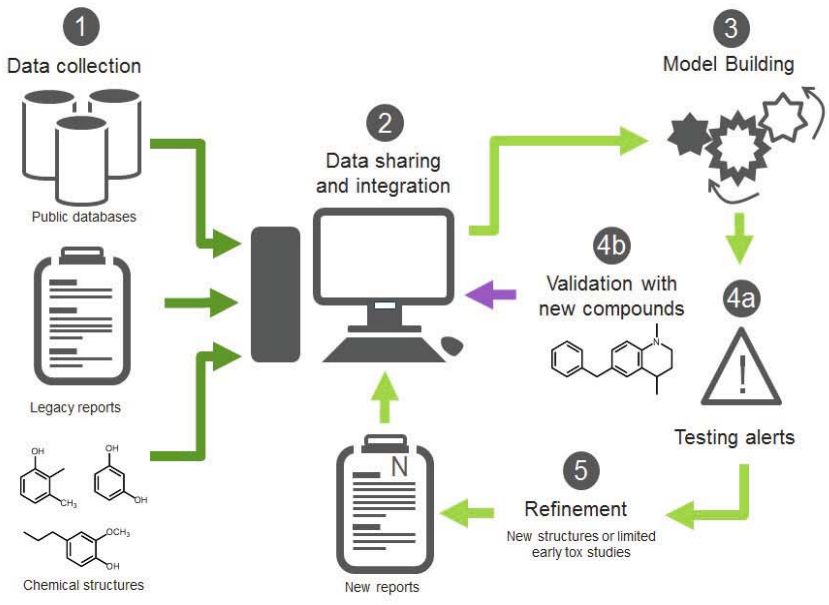

Figure 1: eTOX workflow: from data collection and extraction to data sharing and development of new in silico prediction systems.

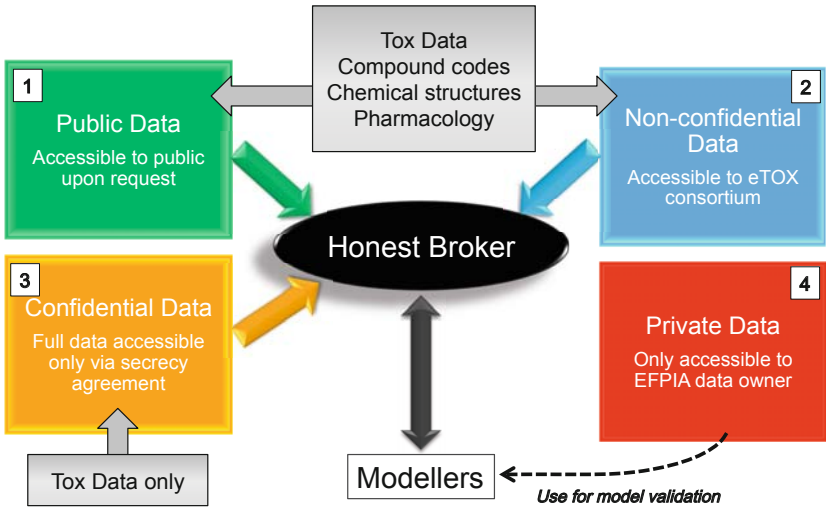

Figure 2: Concept of data sharing and access to the eTOX database.

project as well as for the governance of data access. In addition, it was agreed to leave it to the discretion of each pharmaceutical company to decide what data will be shared with all partners ("non-confidential data") and what should be available only to selected partners (e.g. experts in model building) after agreeing on additional secrecy clauses ("confidential data") or which was to be kept accessible only for the contributing company ("private data"). Besides the proprietary data from companies it was also decided to collect and integrate selected sources of complimentary public data (see figure 2).

A second challenge was the development of standardized terminologies, which was not only a prerequisite for a database search but was also a necessary prior condition for data analysis for model building. Even though the conduct of the toxicological studies is highly harmonized by international guidelines (ICH, OECD), the terminologies for the measured parameters and findings were far from being standardized.

The development of a controlled and harmonized vocabulary for some parameters such as clinical chemistry represented a manageable task. For example the measurement of the activity of a specific liver transaminase can be reported in the original study as "alanine aminotransferase", "ALAT", "ALT"; "glutamate-pyruvate transaminase", "GPT" or "serum glutamate pyruvate transaminase", SGPT, and needed to be converted in the database to one harmonized term.

However, the issue was less easily resolved for the field of pathology. Assigning just synonyms would not suffice for a search for liver toxicity which should also bring back findings in "liver lobe", "hepatobiliary system", "intrahepatic bile duct" and other related tissues and structures. To overcome this problem a logical combination of the anatomical relationship, a so-called ontology, had to be implemented in the database in order to find also the related terms both in a horizontal as well as in a vertical way. The tool developed in the course of the project to align the different terms was called ontobrowser (Ravagli et al. 2016) and is depicted in figure 3.

During the course of the eTOX project the necessity of a standardized preclinical vocabulary was also identified by regulatory agencies as a prerequisite for an in-depth data analysis. The FDA developed over the past years the "Standards of Exchange for Nonclinical Data" (SEND),

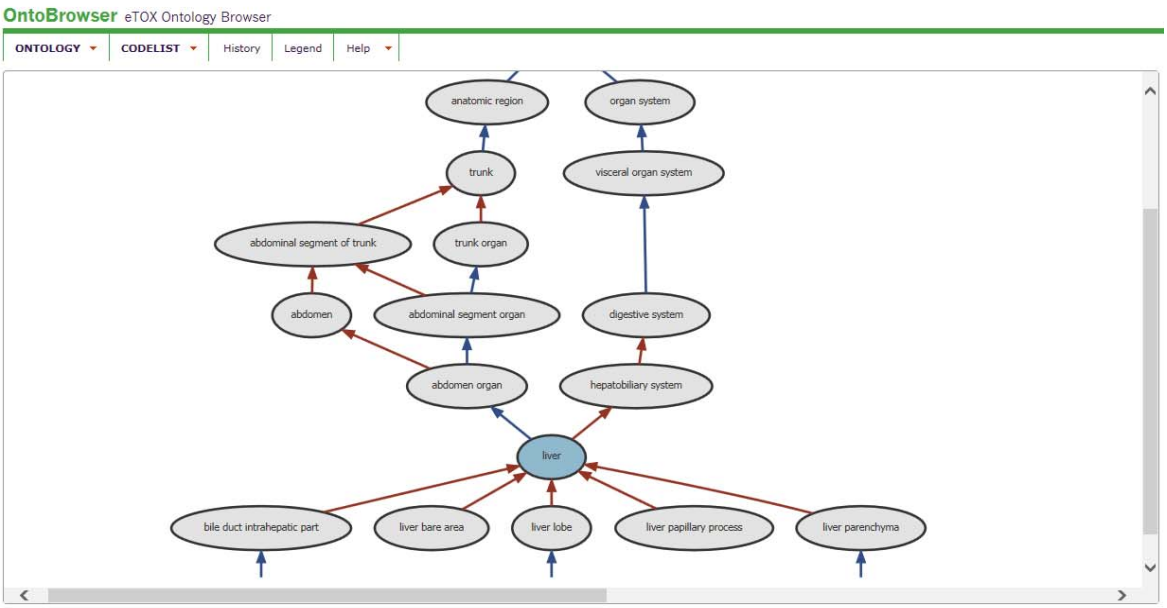

Figure 3: Screenshot of an ontobrowser section displaying how different anatomical and histological findings are logically linked to the target organ "liver". The ontobrowser was developed in the eTOX project and can be downloaded under http://www.etoxproject.eu/.

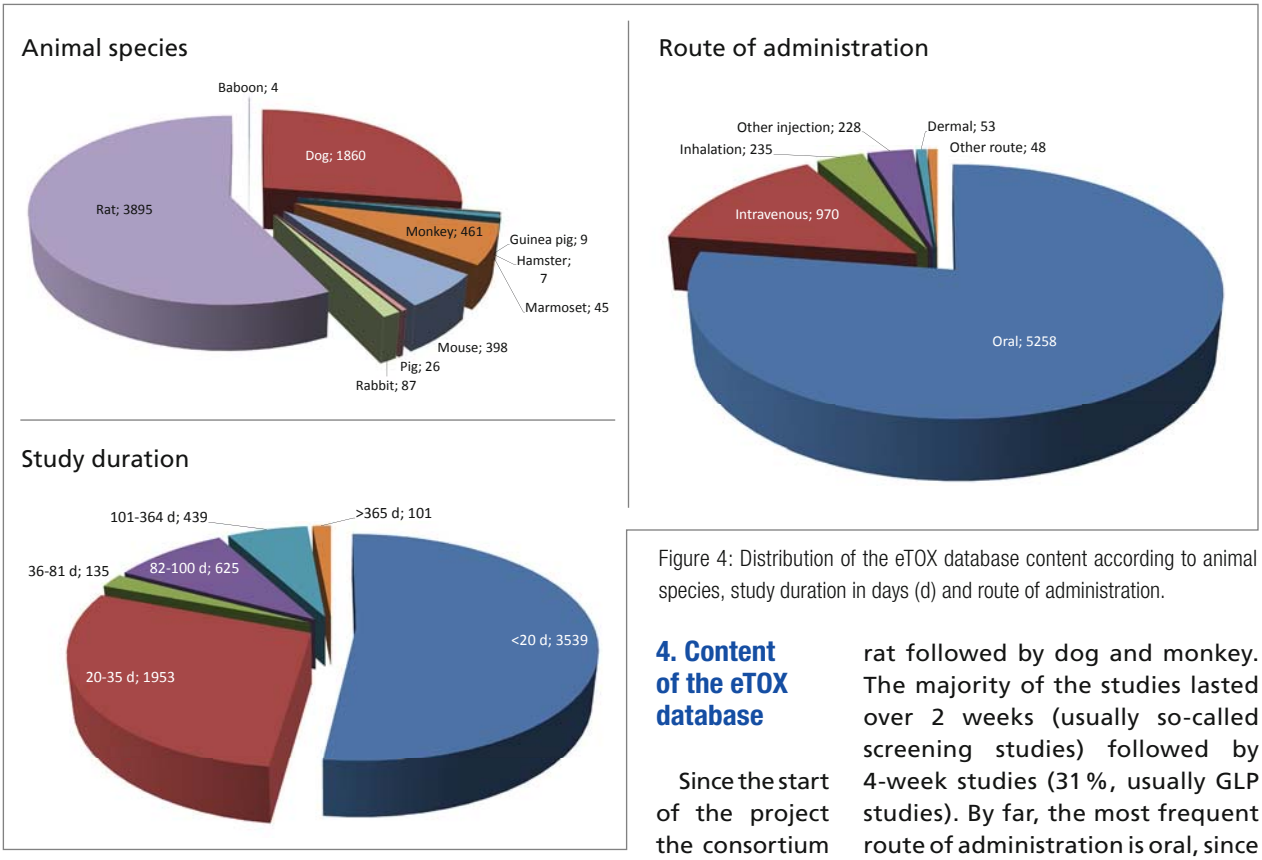

Figure 4: Distribution of the eTOX database content according to animal species, study duration in days (d) and route of administration.

4. Content of the eTOX database

Since the start of the project the consortium has collected the results of more than 8,000 systemic toxicity studies for more of 1,900 different chemical structures, which are now accessible as non-confidential data to the project partners for their searches.

Figure 4 shows an overview of the collected study types analyzed for species distribution, duration of the study and frequency of routes of administration. As expected the most commonly used species is the rat followed by dog and monkey. The majority of the studies lasted over 2 weeks (usually so-called screening studies) followed by 4-week studies (31 %, usually GLP studies). By far, the most frequent route of administration is oral, since in most cases this is also the route intended for the clinical application. Second most frequent administration is the intravenous route.

An analysis of the distribution of effects amongst organs and tissues shows that liver is the most frequently affected organ (see figure 5). Whereas this observation was not unexpected given the high doses administered in these studies, which pass through this organ and are metabolized, it came as a surprise that the second most affected organ was the thymus. It was postulated that the observed thymus effects are actually secondary effects resulting as generalized stress response from organ damage and subsequent inflammatory processes at high toxic doses – a hypothesis which needs further examination.

Simple and immediate visual data extraction as for figure 6A can already give precious information about overall frequency of events and specific sensitivity of species. Figure 6B illustrates that rat would be particularly prone to liver metabolic enzyme induction leading to hepatocyte hypertrophy, while the overall health status of the dog would be reflected through

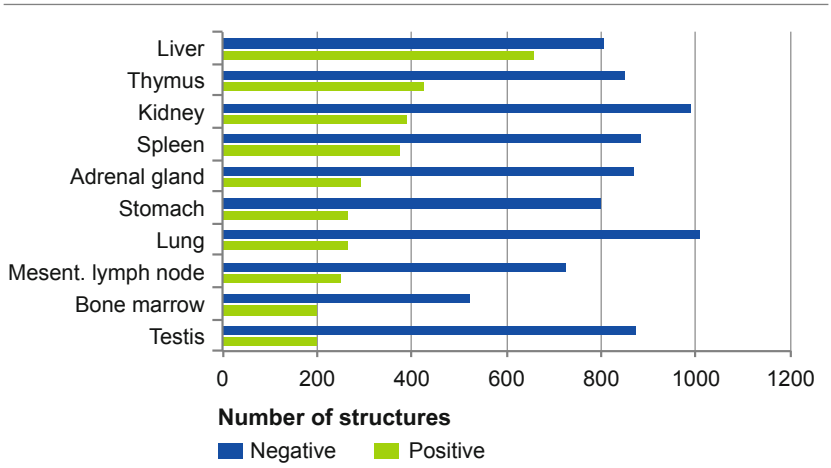

Figure 5: List of the 10 most frequent organs for which any toxicological findings were reported, with the number of compounds that have hits (positive structures – green) and compounds that have no hits (negative structures – blue) for these organs.

which is mandatory from December 2016 onwards for any new systemic toxicity study intended for a drug approval submission as set forth in the FDA Guidance for Industry "Providing Regulatory Submissions In Electronic Format – Standardized Study Data" published in December 2014. As a consequence, the developed eTOX terminologies were either provided to SEND or harmonized with this standard.

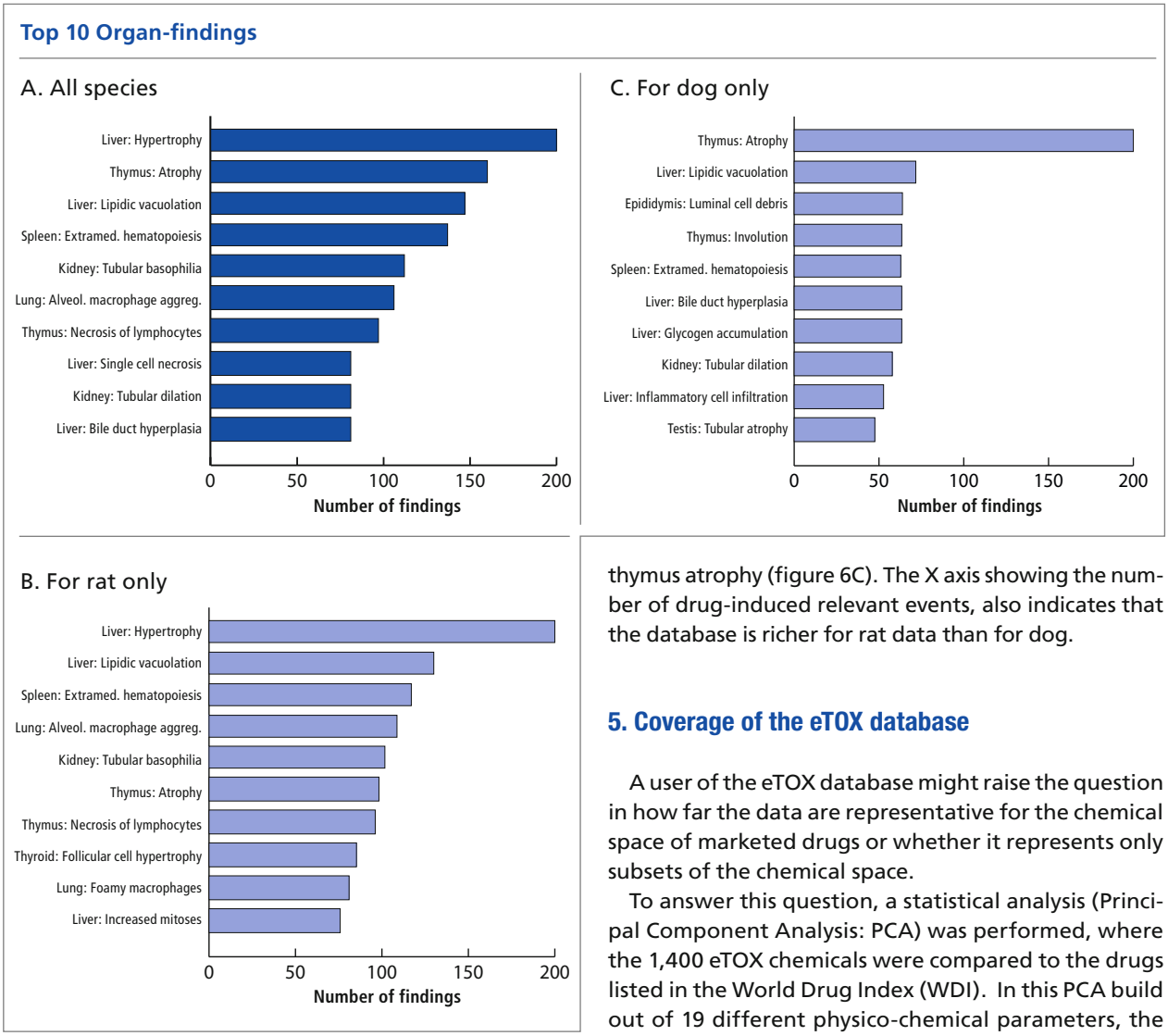

Figure 6: List of the 10 most frequent organs for which a specific histopathological finding was reported, for all species (A), for rat only (B) and for dog only (C).

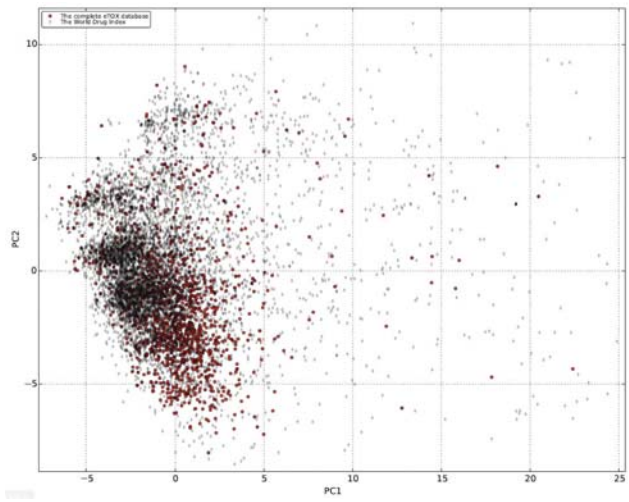

Figure 7: Principle Component Analysis (PCA) of the chemical space of eTOX and WDI (World Drug Index). Each red spot represents an eTOX chemical structure and smaller grey diamonds represent the whole druggable chemical space from the (WDI). The eTOX data shows a good overlap with the chemical space of the WDI.

thymus atrophy (figure 6C). The X axis showing the number of drug-induced relevant events, also indicates that the database is richer for rat data than for dog.

5. Coverage of the eTOX database

A user of the eTOX database might raise the question in how far the data are representative for the chemical space of marketed drugs or whether it represents only subsets of the chemical space.

To answer this question, a statistical analysis (Principal Component Analysis: PCA) was performed, where the 1,400 eTOX chemicals were compared to the drugs listed in the World Drug Index (WDI). In this PCA build out of 19 different physico-chemical parameters, the two main principal components show a dense core and a more spread space for the WDI drugs (figure 7).

The eTOX compounds follow the same pattern, allowing meta-analysis of the data within a relevant chemical space for drug development. Since the eTOX database comprises data from several marketed drugs, some of these dots are common to both databases. Furthermore, it is possible to directly identify within the eTOX database, which compounds have been marketed drugs and which of these have been withdrawn from the market for safety issues. Overall, one can conclude that the compounds in the eTOX database are covering the general "drugable" space as represented by WDI chemical structures of drugs that have ever been marketed.

6. Application of the eTOX database

6.1. Search for structural similar compounds

As shown above the analysis of the data already provides interesting insights, but the true value for early drug development lies in the knowledge discovery hidden in the database. One straightforward

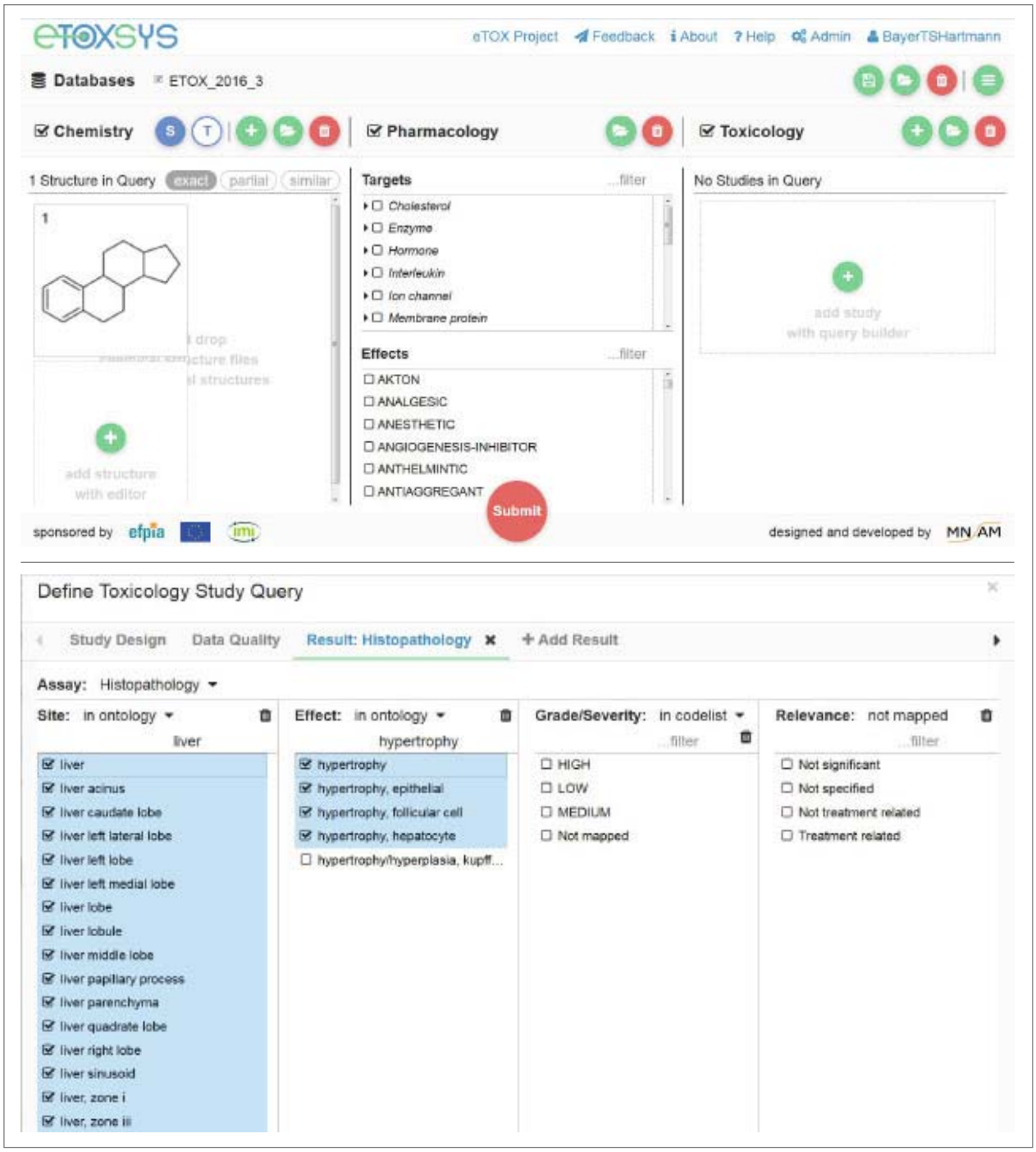

Figure 8: Screenshots of the graphical user interface: the system allows for chemistry-based (structure, substructure, trivial names, or IUPAC [International Union of Pure and Applied Chemistry] codes), pharmacology-related (targets, mode of action), or effect-based (organs, tissues, biomarkers) searches (lower picture).

approach lies in the comparison of new drug candidate structures with the collected data. The graphical user interface developed by the partner Molecular Networks (Nuremberg, Germany) of the database allows for three search strategies (see figure 8), one for structural similarity, a second for pharmacological target and mode of action and a third for similar toxicological effects, which can be combined for complex queries.

At Bayer, a process has recently been established where all new drug candidates for which a lead structure has been identified are searched for structural similar compounds in the database. Results of this search can be used to trigger an inclusion of an organ or tissue already in a screening study if there is evidence from legacy data that structurally closely related compounds cause effects to this organ. A recent example shall illustrate this application:

P2Xn receptors are a family of cation-permeable ligand gated ion channels belonging to the purinoreceptor family. Due to the fact that representatives of this receptor family play a prominent

Figure 9: Distribution chart of the different in silico model category, which are currently implemented in eTOXsys (ADME: adsorption, distribution, metabolism & excretion).

role in nociception, Bayer currently investigates antagonists for specific receptors within said family. Searching the database for structurally similar compounds resulted inter alia in the retrieval of Fluopyram, a fungicide. Reported compound-related organ effects included liver (hypertrophy) in the mid-dose and thyroid (hypertrophy) in the high dose. The effects of the fungicide Fluopyram and similar compounds on liver and thyroid are mechanistically well characterized and considered to be rodent-specific without relevance for human safety (activation of hepatic CAR/PXR receptor; Rouquie et al. 2015). It was therefore decided to include investigations of the thyroid already in the early screening studies with rats. Findings in thyroid and liver would then trigger additional mechanistic studies similar to those performed for Fluopyram, which could proof similar underlying mechanisms and thus confirm the lack of human relevance. By these means we were able to de-risk the candidate in an early project phase.

6.2. Application of in silico prediction systems

Complementary to the search capabilities in the collected legacy data, the public partners in the eTOX project developed prediction models for numerous biological processes or properties. The models are available through the same system, i.e., compounds for which the database search resulted in evidence for liver toxicity can be directly subjected to the models for drug-induced liver toxicity (DILI) or transporter inhibition.

The system currently contains 202 different in silico models ("predictive tools"). The majority of them relate to the predictions of pharmacological predictions, followed by organ toxicity prediction models ad pharmacokinetic properties (figure 9).

Figure 10 illustrates, how such models can be applied during early drug development in order to develop hypotheses regarding the chemical substituents responsible for an observed or predicted effect. Based on the characterized chemical space for a specific pharmacological property of a drug candidate, the medicinal chemist can modify the chemistry of the candidate in order to obtain optimal properties.

6.3. Search for identical or similar pharmacological targets

Another use case for the database is the search for pharmacological targets. This information is currently not yet available for all studies or structures because some pharma companies considered this information as particularly sensitive. Despite the incomplete coverage it was interesting to observe for the Bayer early drug development projects that in 17% of the queries, toxicological legacy data were available for the same target in the database. This indicates that the target was or currently is already under investigation by competitors. In further 17% of the searches similar or related targets were found (agonists vs. antagonists, downstream targets of the original target). The retrieved structure corresponding to the studies sometimes shows structural similarities to the in-house hits. In other cases they belong to a completely different chemical space. Very often the activities of the competitors are known to the in-house pharmacologists from patent searches, congress reports or publications. However, the toxicological data are rarely published in detail. The eTOX database therefore allows for a comparative safety assessment of such compounds. A result of such a comparison can be that specific toxicity findings which were observed for the own compound, are also found for the same target in the database. If the compounds for this target differ in their structural class, then the observation of overlapping findings points towards a target-related toxicity, i.e. it will be difficult to influence the target without triggering certain toxicities.

Furthermore the onset and the time course as well as the severity of the findings can be compared between the own candidate and competitor compounds. In the past, in some projects significant effort was spent to synthesize the competitor structure for performing own comparative pharmaco-toxicological studies. Using the eTOX data, it can be envisaged that synthetic capacity and animal studies can be avoided in some cases.

If toxicities are found for the own candidate but competitor structures for the same target are devoid of such effects, the conclu-

sion may be drawn that the toxicity is not target-related, which then leaves room for the structural optimization of the in-house compound.

6.4. Assessment of species-specificity of toxicological findings

As mentioned above the systemic toxicity studies in two species are a prerequisite for the first-in-man trial. In many cases, the findings and the severity of effects in one species does not concur with the observation in the other species. The toxicologist thus faces the question which species is more predictive for the human. It is not always the most sensitive species, which best predicts the human effects, as has already been discussed above for certain thyroid effects in rats, which are known to be of low relevance for man. Toxicology text books describe numerous examples of species-specific toxicities, but for rare findings the toxicologist lacks a solid database for an assessment.

Toxicologists at Sanofi have recently presented an eTOX data mining approach for the assess-

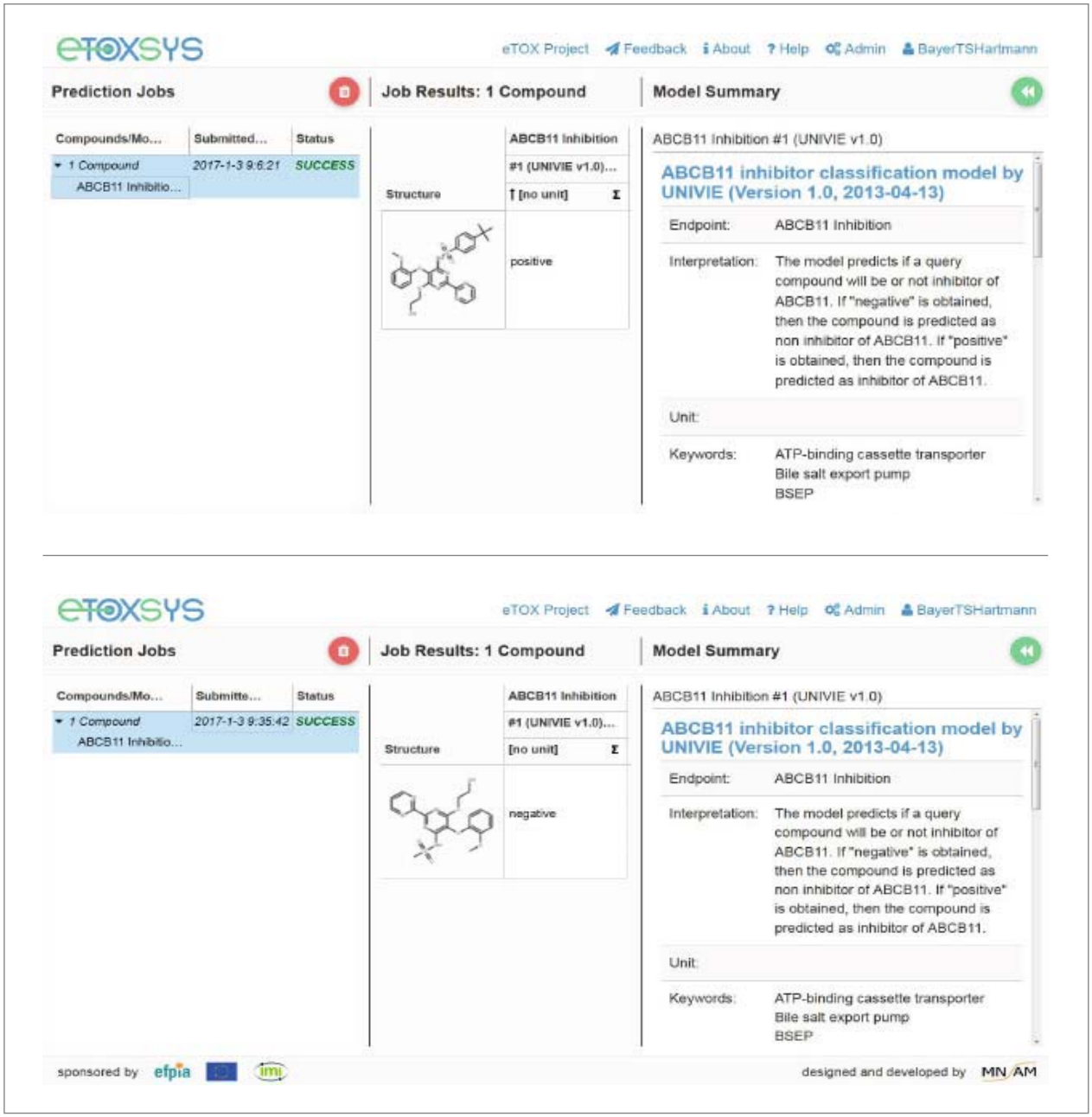

Figure 10: Illustrative example of the theoretical application of the in silico tools: The bile salt export pump (BSEP/ABCB11) is involved in bile acid-dependent and -independent bile secretion, respectively. It has been reported that bosentan, a marketed endothelin receptor antagonist, inhibits BSEP, which may lead to cholestatic liver injury due to the intracellular accumulation of bile salts. The effect is predicted by an in silico tool in eTOXsys developed by the University of Vienna (upper screenshot). The tool could be used to modify the structure (e.g., by eliminating the benzylic side chain) in order to obtain a structure which results in a negative prediction (lower screenshot). Such tools could help the medicinal chemist to develop hypotheses on possible structural modifications of candidate compounds.

ment of species-specificity of a kidney finding, namely "renal papillary necrosis" (Drewe et al. 2016). Their analysis of the eTOX data showed that this finding occurred more frequently in rat compared to dog studies. The toxicologists furthermore found an association between the occurrence of this finding and compounds acting as inhibitors of the targets EGF or VEGF receptor. These results show that the rat is evidently more sensitive than the dog, but it does not immediately answer the question, which of the two species is the more predictive for humans. Today, the eTOX database cannot yet provide responses because the corresponding human data lacks for comparison. On the other hand, the results of such preclinical sensitivity analysis can lead to further hypotheses, e.g., is the higher sensitivity of the rat due to the higher glomerular filtration rate (GFR) compared to dog? A higher GFR may lead to higher local drug levels in the kidney of these animals. Based on this assumption, mechanistic studies can be designed. In case these studies confirm the hypothesis, it can be concluded that for this specific finding the dog is more relevant to humans than the rat, since GFR in humans is closer to dogs than to rats.

7. Access to the eTOX database

The eTOX database is currently fully accessible only to the partners of the eTOX consortium. However, since all IMI projects are public-private partnerships it was planned from the beginning of eTOX to provide access to external parties. During the course of the project a data access strategy was developed (Cases et al. 2014) which allowed other European research collaborations to retrieve and re-use data sets for their purpose.

In addition, a smaller subset of the data base was made available as a demo database to illustrate the capabilities of the system and the level of detail to a broader audience. This demo database is accessible via the link https://etoxsys.eu.

In order to maintain the database after the end of the project a sustainability plan was developed. Key component of this plan is a business broker, who will assure that the modular aspects of the project, i.e., the database, the developed ontologies, the user interface and the different models will be managed as a single commercial entity. The business broker will continue to act as an honest broker for the data but no data dump into the public domain is intended.

The business broker assisted by a User Board will determine the fees for future customers based on fair and reasonable grounds, i.e., the access and maintenance fees will consider aspects such as new data contribution (in-kind contribution), the number of users per institution, the role of the institution (commercial vs. academia/regulatory) and, if commercial, the turn-over of the company. The commercialization of the database is intended immediately after the end of the project, i.e., for the first quarter 2017.

8. Outlook

The above described example for the analysis of animal to human translation shows, that an important part is still missing in the big data puzzle, namely a well-structured collection of clinical data. To be able to answer such translational questions, i.e., the transferability of the preclinical findings to humans, access to clinical data is key. The challenge is similar to the one Harry Olson and collaborators encountered, except that large preclinical data sets are today available through eTOX. For the clinical data the heterogeneity is, however, by far larger. A correlation analysis should not only include the clinical data acquired during drug development (phases I-III), but also post-marketing studies as well as reports in the frame of pharmacovigilance. In contrast to the eTOX data collection, the privacy of patient data would prevent the sharing of some clinical data through a single big consolidated database, i.e., alternative approaches are required.

Based on the success of eTOX and the remaining open questions, another project under Innovative Medicines Initiative 2 will try to resolve these issues with the help of intelligent interfaces to existing clinical databases of different origins. The interfaces shall protect patient data privacy but allow for statistical translational analyses. In a few years from now, it should thus be possible to determine the predictivity of a preclinical finding in a specific species for the human situation within a few minutes time.

The new IMI project called "eTransafe" is intended to start in the second quarter of 2017.

9. Acknowledgements

The research leading to these results has received support from the Innovative Medicines Initiative (IMI) 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. |

References

Cases M., Briggs K., Steger-Hartmann T., Pognan F., Marc P., Kleinöder T., Schwab C.H., Pastor M., Wichard J., Sanz, F.: The eTOX Data-Sharing Project to Advance in Silico Drug-Induced Toxicity Prediction. International Journal of Molecular Sciences 15, 21136–21154 (2014).

Drewe W.C., Briggs K.A., Schwab C.H., Kleinöder T., Yang C., Steger-Hartmann T., Pawletta M., Stolte M., Amberg A.: eTOXsys - An Integrated Platform for Profiling And Data Mining Across

Multiple Databases and Prediction Models. EUROTOX poster P05–017, Sevilla (2016).

Olson H., Betton G., Robinson D., Thomas K., Monro A. et al.: Concordance of the Toxicity of Pharmaceuticals in Humans and in Animals. Regulatory Toxicology and Pharmacology 32, 56–67 (2000).

Ravagli C., Pognan F., Marc P. OntoBrowser: a collaborative tool for curation of ontologies by subject matter experts. Bioinformatics 1–2 (2016).

Rouquié D., Tinwell H., Blanck O., Schorsch F., Geter D., Wason S., Bars R.: Thyroid tumor formation in the male mouse induced by fluopyram is mediated by activation of hepatic CAR/PXR nuclear receptors. Regulatory Toxicology and Pharmacology 70, 673–80 (2014).

Shanks N., Greek R., and Greek J.: Are animal models predictive for humans? Philosophy, Ethics, and Humanities in Medicine 4:2 (2009).

AUTHORS

Dr. Thomas Steger-Hartmann is Head of Investigational Toxicology at Bayer AG, Germany. He has more than 20 years of experience in the field of pre-clinical and eco-toxicological assessment of pharmaceuticals and early drug development, particularly in the field of contrast agents, hormonal compounds and oncology indications. He is a trained biologist and holds a M.Sc. degree from the University of Michigan and a PhD from the University of Freiburg i. Br., Germany. Dr. Steger-Hartmann is involved in several international research collaborations in the field of toxicology and environmental safety. He is currently deputy leader of the European Innovative Medicines Initiative (IMI) consortium 'eTOX – Expert Systems for in Silico Drug Toxicity Prediction'.

Contact:
thomas.steger-hartmann@bayer.com

For the past 26 years, **Dr. François Pognan's** entire focus has been the understanding and prediction of drug-induced toxicity mechanisms using a variety of tools, ranging from in vitro approaches to in vivo investigative studies. He holds a PhD in Molecular Oncogenesis from University Pierre and Marie Curie in Paris, and has been occupying various mechanistic toxicology positions in Pfizer-France, AstraZeneca UK and USA, and now in Novartis-Pharma in Switzerland as Executive Director in Pre-Clinical Safety. Currently, he is also leading the eTOX consortium.