Addressing Toxicity Risk when Designing and Selecting Compounds in Early Drug Discovery

Matthew Segall*, Chris Barber†, Nicholas Foster*

*Optibrium Limited, Cambridge, UK †Lhasa Limited, Leeds, UK

Abstract

It has been estimated that toxicity accounts for approximately 30% of expensive, late stage failures in clinical development. Therefore, identifying and prioritising chemistries with a lower risk of toxicity, as early as possible in the drug discovery process, would help to address the high attrition rate in pharmaceutical R&D. We will describe how expert knowledge-based prediction of toxicity can alert chemists if their proposed compounds are likely to have an increased risk of causing toxicity, based on precedence for similar compounds where experimental data are available. However, an alert for potential toxicity should be given appropriate weight in the selection of compounds. It is important to balance potential opportunities against the risk of late stage failures cause by toxicity; an alert may not be sufficient reason to 'kill' a compound or chemical series. If a series achieves good outcomes for other requirements, it may be appropriate to progress selected compounds and generate experimental data to confirm or refute a prediction of potential toxicity. We will discuss how multi-parameter optimisation approaches can be used to balance the potential for toxicity with other properties required in a high quality candidate drug, such as potency and appropriate absorption, distribution, metabolism and elimination (ADME). Furthermore, it may be possible to modify a compound to reduce its likelihood of toxicity and we will describe how information on the region of a compound that triggers a toxicity alert can be interactively visualised to guide this redesign.

Knowledge-based Toxicity Prediction

Knowledge-based predictive systems for small molecule toxicity, such as Derek Nexus™ [1], emulate the decision-making process of an expert by applying a form of artificial intelligence. A knowledge base is used to make a prediction by inferring relationships between facts through a process known as reasoning. This allows for the introduction of associated data such as reactivity or knowledge of the mechanism of action, and can cope with uncertainty and conflicting data. Expert systems are particularly well suited to making predictions for toxicities derived through multiple mechanisms for which only incomplete datasets are available and can often provide more interpretable results.

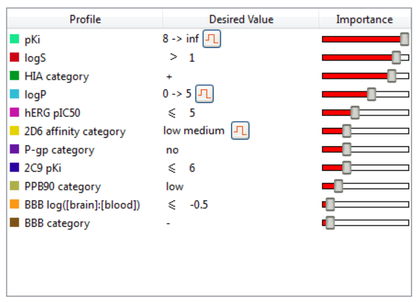
Derek Nexus provides a prediction and, if positive, an associated likelihood qualifying the prediction; some of these are shown in the table below. In practise, it has been demonstrated that likelihood can be taken as a level of confidence since it correlates well with the accuracy of a prediction[2].

Result	Interpretation
No Report	Nothing to report
Equivocal	There is an equal weight of evidence for and against the proposition
Plausible	The weight of evidence supports the proposition
Probable	There is at least one strong argument for the proposition and none against it

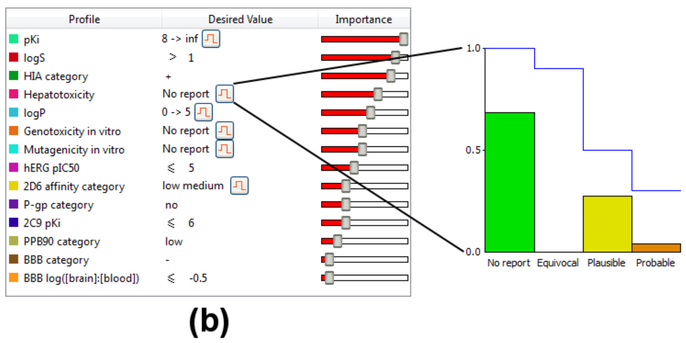
Expert systems are frequently applied in the later stages of drug development, for risk assessment and design of experiments to support a regulatory submission. In such cases, features including mechanistic interpretation, expert commentary, documentation, validation statistics and supporting data are particularly valuable. However, these methods are less commonly used early in drug discovery, where the numbers of compounds considered are much larger and the scientists using the predictions are less likely to be expert toxicologists, making detailed examination of each prediction impractical. Below, we will discuss how expert knowledge based prediction of toxicity has been integrated within the StarDrop™ [3] platform to guide the design and selection of compounds in early drug discovery.

Multi-Parameter Optimisation

Predictions of toxicity risk should be balanced against other properties, such as target potency and ADME, and given appropriate weight in the selection and design of compounds. Multi-parameter optimisation (MPO) methods [4], such as Probabilistic Scoring, allow a project team to define a profile of ideal property criteria and their importance to a project's objective, as illustrated below. Predicted and experimental compound property data are then assessed against the profile to prioritise compounds with the best overall chance of downstream success. The uncertainty in the property data, due to experimental variability or statistical errors in predictions, can be explicitly taken into account, to identify when compounds can be confidently distinguished and avoid inappropriate rejection of compounds, leading to missed opportunities.



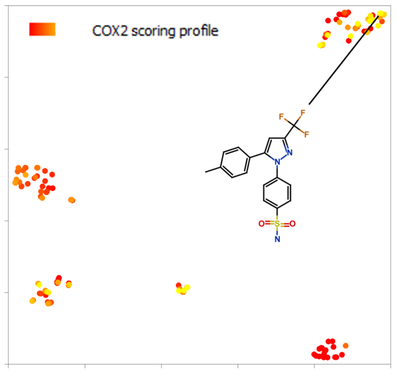
(a)

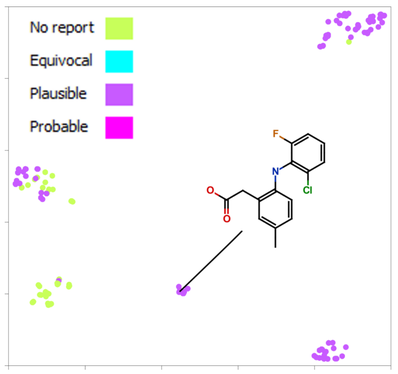


(b)

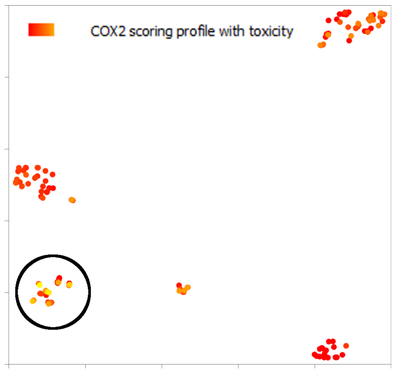
Series Selection in Early Drug Discovery

In early 'hit-to-lead' it is common to consider a library of compounds, representing multiple chemical series, with the objective to efficiently identify one or more high quality lead series for progression. The 'chemical space' on the right illustrates a compound library, with experimental screening data for potency against the COX2 enzyme and predicted ADME properties. The points are coloured by the score against the multi-parameter profile (a) shown below, left, from low (red) to high (yellow). This indicates three clusters of compounds with a good balance of potency and ADME properties, one of which includes the 'gold standard' drug in this class, Celecoxib (highlighted).





The same chemical space is shown here, coloured by predictions of hepatotoxicity using Derek Nexus, indicating clusters with evidence of hepatotoxicity. One such cluster includes the drug Lumiracoxib (highlighted), which was withdrawn from the market in several countries, mostly due to hepatotoxicity concerns.

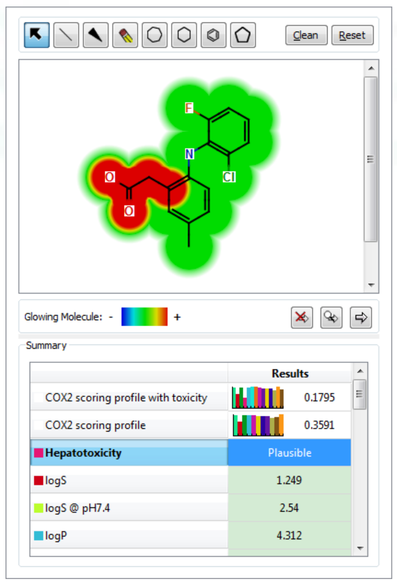


The toxicity predictions can be combined with *in vitro* and *in silico* data for other properties in an overall scoring profile (b) shown below, left. The points to the right are coloured by the resulting scores, from low (red) to high (yellow). One cluster clearly stands out (circled), with the highest likelihood of yielding a potent lead series with good ADME properties and reduced toxicity risk.

It is notable that Celecoxib is also predicted to have plausible evidence of toxicity, illustrating the potential for false positive predictions. However, the score for Celecoxib (0.15±0.08) is not statistically significantly different from the top-scoring compound (0.45±0.30) and therefore this compound would not be rejected outright. An MPO approach balances toxicity predictions against other compound attributes and to use them as triggers for early assessment of the potential risk, rather than immediately eliminating any compound that fired an alert.

Guiding Compound Design

An advantage of a knowledge-based approach to toxicity prediction is that the structural features of a compound associated with an increased toxicity risk are identified. In the case of Lumiracoxib (right), a single functionality is highlighted as the cause of the structural alert for hepatotoxicity, in common with other members of the series. Approaches for reducing this risk, while retaining potency and other desirable properties, can be investigated at an early stage before deciding if this series should be rejected. An interactive design environment, coupling predictive models within an MPO environment enables strategies for design of compounds to be explored with instant feedback.



Conclusions

Knowledge based predictions of toxicity can provide a useful guide to the design and selection of compounds with reduced toxicity risk in early drug discovery. Toxicity predictions must be balanced against other important property requirements and the uncertainty in predictions should be taken into account.

References

[1] Derek Nexus, Lhasa Limited. http://www.lhasalimited.org/

[2] Judson, P.J. *et al.* (2013) Toxicology Research, **2**(1), pp. 70–79

[3] StarDrop, Optibrium Ltd. http://www.optibrium.com/

[4] Segall, M.D. (2012) Curr. Pharm. Des. **18**(9) pp. 1292-1310