To model or not to model?

Lilia Fisk, Stepane Werner, Thierry Hansen
Lhasa Limited, Granary Wharf House, 2 Canal Wharf, Leeds, LS11 5PS
E-mail: Lilia.Fisk@lhasalimited.org

Introduction

Drug development continues to be an expensive activity due to significant numbers of drug candidates failing for reasons of efficacy or toxicity [1]. To alleviate these issues and help with the prioritisation of compounds, pharmaceutical companies use in vitro testing and in silico modelling strategies in order to identify potential secondary pharmacology liabilities which may lead to adverse outcomes. There is also an increasing awareness that testing and modelling should attempt to rationalise the biological mechanisms at play in these scenarios. For example, a number of biological targets are linked with pathways that could result in hepatotoxicity in vivo. Many, in cases, in vitro assays can be considered as “representatives” of the molecular initiating events (MIEs) for hepatotoxicity [2]. In silico modelling techniques can take advantage of data generated in this way alongside knowledge of the chemical structure to understand and predict potential liabilities in a biological setting. Prior to beginning any in silico study, the suitability and appropriateness of the available data for the specific models should be assessed.

We present research that explored a number of descriptors based on different methods to fragment a chemical structure (including chemical reactivity and pharmapohore types). Whereas it is expected that the type of interaction with a target will narrow the choice of suitable descriptors, it is also important to assess if the data being analysed is itself appropriately modelable. Two published approaches to measure “how modelable is this dataset?” were implemented in-house and compared: the MODellability Index (MOIDI) [3], and the Bayes Imbalance Impact Index (BII3) [4].

The MDOI index is based on the detection of activity cliffs in the dataset. For classification models, an activity cliff is defined as the pair of molecules with the smallest distance between them and belonging to different class (figure 1A). The MOIDI index is computed as the ratio between the count of molecules not detected as activity cliffs and the total number of molecules in each class (figure 1B).

Method

Datasets

20 hepatotoxicity MIE datasets were collected from public sources. Structures were standardised to ensure that structural representation was consistent between datasets.

Descriptors

Structural fingerprints: ATOMPAIR_EXTENDEDDBS YL (APES) - topological atom pair using an extended version of the sibyl atom type as label; BF_DF_2 - breadth and depth first and fragmentation up to depth 2, ECFP6 - extended-connectivity fingerprints with maximum diameter 6 [9]. Numerical value descriptors (largest plane, LogP and molecular weight MW) were discretised into bins of different sizes.

MOIDI and BII3 calculation

For each dataset a feature space based on a descriptor was created, and indexes were calculated for combination of pairs datasets and descriptors.

Model generation and performance

In-house developed SOHN (Self-Organising Hypothesis Networks) [6] methodology used for creating models, and 5 times cross-validation (SCV) used to generate performance metrics - MCC (Matthews Correlation Coefficient) [7].

Results 1

Structural and physicochemical descriptors were used in the assessment whether modelability indexes are indicative of a model performance. Structural descriptors: APES, BF_DF_2, ECFP6, largest plane; physicochemical descriptors: LogP, molecular weight. Linear regression indicated that 3 out 6 descriptors have a good correlation for SOHN model predictivity (based on SCV) with MODI and BII3 indexes.

The BII3 index reflects the effect of the imbalance on the classification performance. The classification difficulty is caused by class imbalance (or the imbalance ratio (IR) - the ratio between the number samples in the majority and the minority classes). The higher the value of BII3, the more seriously the imbalance influences the classification of the dataset. The imbalance issue can be corrected using an appropriate learning algorithm.

MIE datasets

Hepatotoxicity MIE datasets were used to evaluate the correlation between performance of the model and modelability indexes.

The hepatotoxicity MIE datasets varied in size and number of positive compounds. The number of compounds per dataset and bias (imbalance ratio) within each dataset are shown in figure 3.

Results 2

Each specific descriptor was assessed for its usefulness in modelling SAR (Structure Activity Relationship) by calculating the weighted average across MIEs MCC value for a descriptor where the weighting factor are MODI and BII3 indexes respectively to reflect an actual contribution. A heat map of weighted MCC values gives a better summary of a difference in performance of MIE models per descriptor and an interaction type.

Conclusions

During this investigation it was identified that both modelability indexes correlated well with SOHN model performance based on the MCC for SCV. While some descriptors significantly correlated with good in their ability to “describe” a chemical - target interaction, there was a clear indication that certain types of interactions were easier to model than others. Liver toxicity MIE models for nuclear receptor interactions built on pharmacophore type descriptors showed a better predictivity in comparison to models for transporter interactions.

References