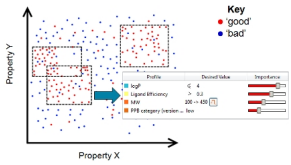
We recently submitted this article, co-authored with Nigel Greene and Falgun Shah of Pfizer's compound safety prediction group. In the article, we discuss new 'rule induction' methods that explore complex data to find interpretable, multi-parameter rules, tailored to any drug discovery objective that can be used to identify compounds with a higher chance of success. This is illustrated with applications to simple ‘drug like’ properties for oral drugs and exploration of experimental target inhibition data to find rules for selecting compounds with a low risk of cardio- and hepatotoxicity.



Abstract

Drug discovery is widely recognised as a process of multi-parameter optimisation, with the objective of finding compounds that meet a profile of many property criteria. These will depend on the ultimate therapeutic goal of the project and are typically chosen based on the subjective opinion of the project team. However, analysis of historical data can help to guide the determination of the most appropriate profile. In this article we describe computational approaches, described as rule induction, that enable an objective analysis of complex data to objectively identify multi-parameter rules that distinguish compounds likely to be successful for a project’s goal. The resulting rules are interpretable and modifiable, allowing experts to understand and adjust them based on their knowledge of the underlying biology and chemistry. Furthermore, the importance of each property criterion can be identified, allowing experimental resources to be focused on generating the most critical data necessary to make effective compound prioritisation decisions. We illustrate this method with two applications: determining rules for simple, calculated properties for orally administered drugs, which are compared with previously published measures of ‘drug-likeness’; and exploring experimental target inhibition

data to find rules to reduce the risk of hepatotoxicity and cardiotoxicity.

You can download a copy of this article as a PDF and the supplementary information .