

## Expectations of a chemist from a 'good' QSAR Model

**Jon Heal, Bill Hamilton, Joe Sheridan, Prosarix Ltd  
and Mukund S Chorghade, Chorghade Enterprises**

---

Constructing a predictive model for compound activity (QSAR), or some other property, (QSPR) gleaned from experimental data is now common practice in drug discovery. Judicious use of QSAR has provided increased efficiency in the hit finding and lead optimisation stages of a project: computational predictive models have been used by researchers in several drug discovery campaigns.

The degree of reliance on QSAR depends on the type of property being predicted, the stage of the project and the relative ease and cost of compound synthesis and subsequent testing. Many QSAR models provide useful predictions; a number do not, despite good statistics generated from internal data used in training.

Cheminformaticians understand that a useful QSAR model must comply with general characteristics proposed by Topliss & Costello<sup>1</sup> and Unger and Hansch<sup>2</sup>:

- The training set must have sufficient examples to cover the range of properties required to be predicted by the model. Usually this includes several log orders of magnitude of the end point being predicted.
- The model should be based on a number of non-correlated descriptors far less numerous than the number of compounds in the training set (at least 5 – 10 fold) and biophysically relating to the property being predicted.
- The simplest model should preferably be selected.
- The model should be characterised by a number of statistical parameters (including correlation, standard deviation, F values and confidence intervals) and cross-validated to test internal predictivity. Ideally, model performance should be measured against a separate test set (external predictivity).

Adhering to these general principles improves the likelihood of the QSAR model being predictive. Often however, there is no extensive external testing (using compounds unseen in training). In this familiar scenario, correlations for external predictions can be far worse than expected regardless of the encouraging statistics generated by the training data: the reasons for such poor model performance are not fully understood yet; notwithstanding the many theoretical studies into this area.

The 'Kubinyi Paradox', following systematic investigations into the relationships between internal and external predictivity, states that high internal predictivity may often result in low external predictivity and vice versa<sup>3</sup>. One explanation for this is that the overall error of the prediction is compounded when errors inherent in the model are coupled with experimental errors in the data from external compounds. This unsuspected conclusion dictates a more critical definition of a 'good' QSAR model as one, which has been validated with significant external testing. Unfortunately, this degree of validation is usually only gained as a project progresses by testing sets of synthesised analogues selected from model predictions. Many models, when first constructed, do not have significant external validation.

Even for a model exhibiting good external predictivity, a second problem relating to the chemical space of the training set and to the scope of the trained model may become apparent. A training set, by definition, is always limited and the model only learns about the properties displayed by this set *i.e.* most models are local. A model for predicting properties for compounds which are substantially dissimilar to the training set will exhibit significantly diminished predictive capability because the range

of descriptor values in the test set are outside of the range 'seen' in the training set. Understanding the scope of the model is critical to recognising the capability of the model to make predictions on diverse structures, although, depending on the descriptors used, this may not always be obvious! An awareness of this problem is only possible from an analysis of the training data. Usually the chemist does not have access to this and is less likely to understand this potential pitfall than the originator of the model. Poor predictions relating to model scope are very common: it is the responsibility of the chemoinformatician to explain where the model is likely to be applicable and flag its possible limitations. Conversely, it is also the responsibility of the chemist to see a QSAR model not as a black box but instead to understand the likely scope of the model based on the chemical space of the training set.

This issue must be considered in the context of where and how the model is being applied. QSAR models may be used either for hit finding (virtual screening) or more classically for lead optimisation (analogue prediction). QSPR models for ADMET predictions may be used as a filter in library profiling or in lead optimisation. The significance of caveats relating to model quality in both of these scenarios requires some clarification. A chemist is less likely to be troubled by inaccuracies during hit finding than lead optimisation.

To exemplify this, imagine a screening program against a given target wherein a limited library of compounds, from in an in-house database of several million, will be purchased for testing based on model predictions. ADMET predictive models and activity predictions are used to filter the compounds. Millions of predictions are possibly being made at this stage. In this scenario, the hit rate from the QSAR and ADMET models can be relatively low, maybe only 30%, but yield a number of interesting compounds that save considerable resource, compared to HTS testing of all of the compounds, and represent a significant enrichment. In hit finding, the accuracy of the model is competing with random selection. This contrasts to the lead optimisation scenario in which fewer predictions are made and the cost of inaccuracy is much greater as it leads to both unnecessary compound synthesis and longer lead optimisation timeframes.

A QSAR model in this phase must be at least as predictive as a chemist's intuition. Since many analogues around a hit are likely to show some activity, the objective value of the QSAR model may be hard to determine. The crunch point occurs when the model predicts that a set of compounds, which a chemist would otherwise favour are predicted to be inactive by the model. It is usual for a model without confirmed external predictive capability to be assessed by such compounds in order to gain validation.

ADMET QSPR models have utility in lead optimisation: the decision not to make compounds that are predicted to exhibit properties outside of a desired range is significant, especially when the cost of testing compounds for those properties is lower than synthesis costs.

The role of QSAR in lead optimisation may be greeted with healthy scepticism. A chemist should question if the model adheres to the rules of Topliss & Costello<sup>1</sup>, and Unger & Hansch<sup>2</sup>. Unfortunately due to a number of reasons (including the ease of generating such models) some do not! A model should not be totally relied upon until at least one round of external testing has provided validation.

Further meaningful predictions may be expected as long as chemical space is similar to compounds seen in training. Model originators should be aware of and inform the chemist when predictions are being made for structures, which may lie outside the scope of the model generation. Finally, the value of the prediction should be balanced against the cost of compound synthesis and testing.

QSAR remains a very powerful technique likely to play an increasingly important role in compound design and selection. It is a highly efficient method of maximising the value of experimental data; its place in the medicinal chemist's tool kit looks assured.

**References:**

1. Topliss JG, Costello RJ (1972) Chance correlations in structure-activity studies using multiple regression analysis, *J. Med. Chem.* **15**:1066-1068
2. Unger SH, Hansch C (1973) On model building in structure-activity relationships. A re-examination of adrenergic blocking activity of beta-halo-beta-arylalkylamines, *J. Med. Chem.* **16**:745-749
3. van Drie JH (2003) Pharmacophore discovery - lessons learned. *Curr. Pharm. Des.* **9**:1649-1664