Multi-Parameter Optimisation in Drug Discovery

By Matthew Segall at Optibrium Ltd

Drug discovery activities are producing ever-larger volumes of complex data that carry significant levels of uncertainty; multi-parameter optimisation methods enable this data to be better utilised to quickly target compounds with a good balance of properties – but they all have their strengths and weaknesses.

One of the major challenges of drug discovery is to identify a compound with a good balance of the many physicochemical and biological properties necessary to become a successful, efficacious and safe drug. Having identified poor pharmacokinetics (PK) as a major cause of clinical failure in the later 1990s, projects now routinely attempt to mitigate this risk by assessing the absorption, distribution, metabolism and elimination (ADME) properties of compounds in early drug discovery. However, more recently, safety has become a greater cause of failure in the clinic leading to the assessment of toxicity earlier in the drug discovery process. The result has been an increase in cost and a reduction in the productivity of drug discovery, but unfortunately a corresponding increase in the success rate in development has yet to be observed (1).

The increase in the volume and complexity of the data now available in early drug discovery, through experimental and computational approaches, has posed a new challenge: how can the full value of this data be realised to quickly identify those compounds with the best chance of downstream success, while confidently eliminating those chemistries which will not have an appropriate balance of properties? This challenge is made greater by the fact that the data generated in drug discovery have significant levels of uncertainty, both in terms of experimental variability and statistical error in predictions, and the need to assess the relevance of the data generated using *in vitro* or *in vivo* animal models to the ultimate, human patient (the 'translation' problem). Furthermore, different compound properties will have varying degrees of importance to the overall objective of a drug discovery project that should be taken into account; it may be necessary and appropriate to make a compromise on less important properties in order to achieve critical requirements.

Keywords

Drug discovery
Data visualisation
Multi-parameter optimisation (MPO)
Ligand efficiency (LE)
Ligand lipophilic efficiency (LLE)
Probabilistic scoring

The challenge of simultaneously optimising multiple factors has been previously addressed in many fields, from engineering disciplines such as automotive and aeronautical design, to economics. The methods used to achieve this are described as multi-parameter optimisation (MPO), multi-dimensional optimisation (MDO) or multi-objective optimisation (MOOP). For simplicity, we will use the term MPO to refer to all of these methods. We can learn from these approaches to better use the data generated in drug discovery in order to quickly design and select compounds with a good balance of properties. However, we must make allowance for the greater uncertainty in drug discovery data in comparison with that available in disciplines such as engineering, where simulations and measurements typically have much greater precision.

One common question is: "Can't this be easily solved by visualisation of the data?" While visualisation is necessary to understand and communicate results, it is rarely sufficient to allow conclusions to be easily drawn, given the complexity of the data at hand. Figure 1 shows a typical three-dimensional plot with additional information shown by colour and size of points. Even with only five-dimensional data, it is difficult to confidently draw a conclusion from these visualisations, even before we consider the relative importance of each property or the uncertainty in the data. An MPO method helps a project team to define a set of criteria and use this pro-actively to guide their decisions to quickly target high quality compounds.

In this article we will briefly outline some of the MPO methods being applied in drug discovery and discuss their strengths and weaknesses. A more

comprehensive review, including example applications may be found in reference 2.

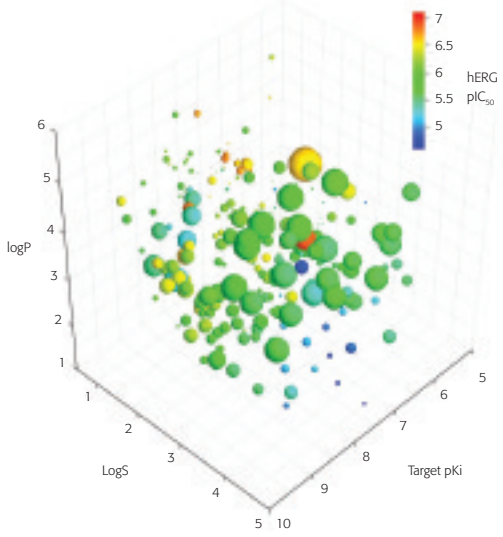
Rules of Thumb

A very common approach to steering chemistry toward appropriate physicochemical properties is to apply a 'rule of thumb'. The best known of these is probably Lipinski's Rule of Five (RoF) which relates four simple characteristics of a compound (the number of hydrogen bond donors and acceptors, lipophilicity and molecular weight) to its likelihood of achieving good oral absorption (3). However, many other characteristics have since been proposed as guides to achieving good PK, developability and safety, including polar surface area, proportion of sp^3 carbons and number of aromatic rings.

The simplicity of these rules of thumb makes them very attractive; they are easy to interpret and apply, helping chemists to avoid chemistries with a low chance of success while optimising potency against their therapeutic target. However, their simplicity also gives rise to a significant weakness; the simple characteristics of a compound on which they are based are only weakly predictive of the ultimate outcome. This weakness is compounded by the temptation to apply these rules rigidly, as hard filters, to eliminate compounds. Is a compound with a calculated logP of 4.9 significantly better than one with logP of 5.1? This leads to a risk of missing valuable opportunities by incorrectly eliminating good compounds.

It is also important to recognise the context in which it is appropriate to apply a rule of thumb; each rule relates to a specific drug discovery objective. For example, the RoF was developed as a guideline to improve the chance of finding an orally absorbed compound. However, it is frequently applied as a definition of 'drug likeness', which is inappropriate if the project objective is to design a compound intended for another route of administration such as topical, intravenous or inhaled (4).

Figure 1: A three-dimensional plot of compound data. Target affinity (pKi), solubility (logS) and lipophilicity (logP) are plotted on the x-, y- and z-axis respectively. Point colour indicates the inhibition of the hERG ion channel (pIC_{50}) and size indicates the degree of blood-brain barrier penetration.



Calculated Metrics

A recent trend has been to combine potency with other parameters into a single metric which may be monitored during optimisation. One commonly used example is the ligand efficiency (LE), which is defined as:

$$LE = \frac{1.4 \times pIC_{50}}{N_H}$$

where $pIC_{50} = -log(IC_{50})$, the IC_{50} is expressed in molar concentration and N_H is the number of heavy atoms (5). This penalises large compounds over small compounds with similar potency because larger compounds tend to have poorer physicochemical and ADME properties. Another popular metric is the ligand lipophilic efficiency (LLE): $LLE = pIC_{50} - logP$, where logP is the octanol:water partition coefficient. This penalises compounds that achieve increased potency through increased lipophilicity, because high logP is associated with poor solubility, absorption and metabolic stability along with an increased risk of non-specific interactions and toxicity (6).

These calculated metrics provide a good approach to guide the selection and optimisation of compounds because it is only necessary to monitor a single parameter. Moreover, they are generally easy to interpret; for example, increasing potency by adding a large lipophilic group is unlikely to be beneficial in the long term compared with introducing a new specific interaction with the target via a hydrogen bonding functionality.

However, similar caveats apply to these calculated metrics as to rules of thumb regarding the context and rigidity with which they are applied. Furthermore, it is important to bear in mind that property values such as potency and logP will have significant uncertainty, so the uncertainties of metrics such as LLE may be high and their ability to confidently distinguish compounds limited.

Pareto Optimisation

First proposed by the Italian economist Vilfredo Pareto early in the 20th century, the idea behind Pareto optimisation is that there is not a single, optimal option, but a family of optima that represent different balances

Figure 2: Plot illustrating the concept of Pareto optimality in which the ideal outcome corresponds to the top right corner of the plot. The points show data for activity against the therapeutic target and stability in human liver microsomes for a set of 75 compounds. The red points are Pareto optimal or 'non-dominated' points; for example, in the case of point A, there are no points with a higher value for both parameters. However, blue points are not Pareto optimal; for example, point B is 'dominated' by point C

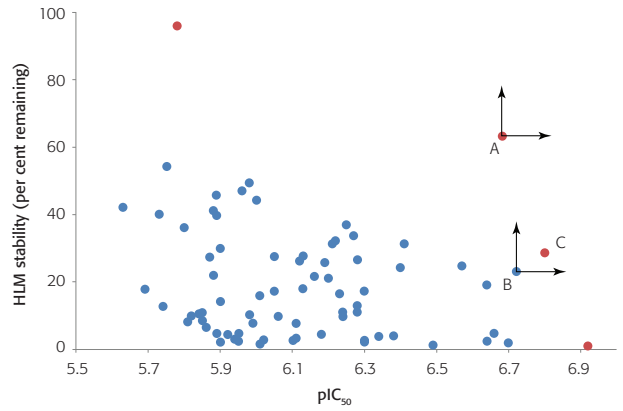


Figure 3: Two examples of desirability functions: (a) defines an ideal property range of 4 to 6; a compound with a property below this range is less desirable, but still potentially acceptable, but above this range a compound would be absolutely rejected; in the case of (b), a compound with property value greater than 8 would be ideal; a compound with property value below 2 would be rejected; and the desirability increases linearly with property value between 2 and 8

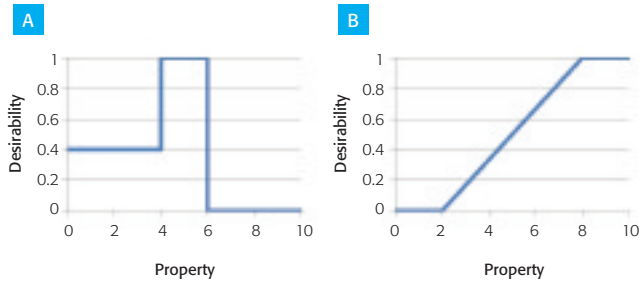
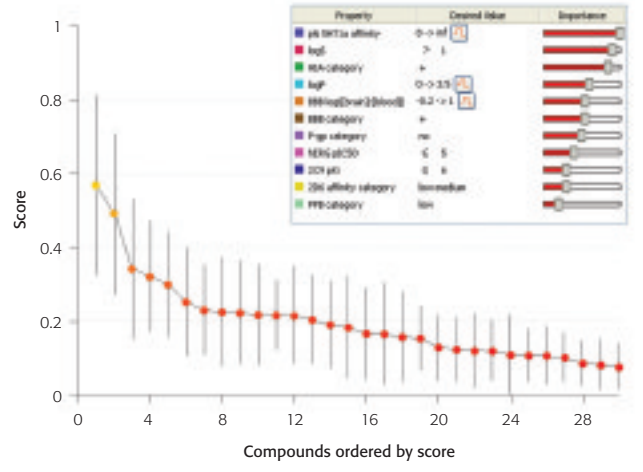


Figure 4: An example output from probabilistic scoring for 30 compounds. The compounds are ordered from left to right along the x-axis in order of their score, and overall score for each compound is plotted on the y-axis. The overall uncertainty in each score (one standard deviation), due to the uncertainty in the underlying data, is shown by error bars around the corresponding point. From this it can be seen that approximately the bottom 50 per cent of compounds may be confidently rejected, as their error bars do not overlap with that of the top-scoring compound. The scores have been calculated against the inset scoring profile, showing the property criteria and importance of each criterion to the overall project objective. Underlying each criterion is a desirability function



of the properties being optimised (7). These solutions are described as 'non-dominated' in that there are no alternatives that are better in all properties being optimised (see Figure 2).

Pareto optimisation is ideal when the appropriate balance of properties is not known *a priori*, because it helps to explore a range of different options that are all optimal combinations of the properties. However, for large numbers of parameters (for example, >5), the number of Pareto optimal solutions can become overwhelming, making it impossible to evaluate them all and choose between them. It is also difficult to take into account differences in the weights assigned to different properties and the uncertainties in the underlying data; it may not be possible to say with confidence which compounds are Pareto optimal.

Desirability Functions

Desirability functions were first developed as an approach to combining multiple factors and calculate a single value used to monitor quality control (8). A desirability function maps a property value onto a scale between zero and one that represents the 'desirability score' of an outcome with that property value; an ideal outcome will achieve a desirability score of one, while a completely unacceptable outcome will receive a desirability score of zero. Two simple, linear examples are shown in Figure 3 – but more complex, non-linear functions may also be used. The overall quality is then calculated by combining the desirability scores for the individual properties by taking their sum or arithmetic or geometric means.

Desirability functions give great flexibility in defining property criteria, their relative weights and the acceptable trade-offs. The results may also be easily interpreted, because the impact of each individual property to the overall desirability index can be calculated to guide strategies to improve the overall quality. A disadvantage of this approach over Pareto optimisation is that it assumes that the profile of properties that an ideal compound would posses is known *a priori*, based on analysis of historical data or the expertise of the project team.

Probabilistic Scoring

The probabilistic scoring method builds on the flexibility and interpretability of desirability functions by explicitly taking into account the uncertainty in the underlying data (9). Thus, not only is a score calculated for each compound that represents the likelihood of achieving the ideal profile of properties, but an uncertainty is also estimated for each score (illustrated in Figure 4). This makes it clear when the available data

confidently distinguishes between compounds and avoids too much weight being given to data with high uncertainty; it also mitigates the risk of incorrectly rejecting valuable compounds based on uncertain measurements or predictions.

In common with the desirability function approach, probabilistic scoring requires the ideal property profile and the weights of the individual property criteria to be specified by the project team. In addition, the uncertainties in the input data must be provided, ideally based on validation of the assays or predictive models used to generate the data or the variability between multiple measurements of a single compound. However, where rigorous experimental errors are not available for an assay, it is still valuable to include an estimate of the confidence in the data, based on the experience of the experimentalist to prevent over-interpretation of the data when choosing between compounds.

Where this information is available, probabilistic scoring brings together the requirements for MPO in drug discovery: flexibility in specifying the property criteria and weights for a project's therapeutic objectives, interpretability of the results to guide decisions on the optimisation and selection of compounds, and a rigorous treatment of uncertainties to give confidence in those decisions.

Conclusion

MPO methods offer the ability to guide design and selection to quickly identify compounds that are likely to achieve a successful outcome in the clinic and occupy a strong market position. The earlier these can be applied in the drug discovery process, the more likely it is that quick progress can be made to identify high quality lead and candidate compounds. Too much focus on a single optimisation goal early in the process – typically target potency – can make it very difficult to correct deficiencies later, leading to multiple, long iterations in lead optimisation or expensive, late stage project failures.

The complexity of the design goals and data in drug discovery makes it very difficult to achieve MPO without computational support; psychologists have shown that people find it challenging to make good, objective decisions based on complex, uncertain data when the stakes are high, falling prey to common 'cognitive' biases (10). However, any software to support MPO must be easily accessible to the key decision-makers in drug discovery – chemists and biologists as well as computational experts – to allow exploration of potential strategies with intuitive and immediate feedback. This can achieve the best combination of the expertise of drug discovery scientists with computational support to guide confident, objective decisions and quickly target high quality chemistry.

References

1. Paul SM *et al*, How to improve R&D productivity: the pharmaceutical industry's grand challenge, *Nat Rev Drug Discov* 9, pp203-14, 2010
2. Segall MD, Multi-Parameter Optimization: Identifying high quality compounds with a balance of properties, *Curr Pharm Des*, Vol (in press), 2011, Preprint may be downloaded from www.optibrium.com/community
3. Lipinski CA, Lombardo F, Dominy BW and Feeney PJ, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, *Adv Drug Deliv Rev* 23, pp3-25, 1997
4. Choy YB and Prausnitz MR, The rule of five for non-oral routes of drug delivery: ophthalmic, inhalation and transdermal, *Pharm Res* 28(5), pp943-948, 2011
5. Hopkins AL, Groom CR and Alexander A, Ligand Efficiency: a useful metric for lead selection, *Drug Discov Today* 9(10), pp430-431, 2004
6. Leeson PD and Springthorpe B, The influence of drug-like concepts on decision-making in medicinal chemistry, *Nat Rev Drug Discov* 6(11), pp881-890, 2007
7. Jaffe W, Pareto translated: A review article, *J Economic Literature* 10(4), pp1,190-1,201, 1972
8. Harrington EC, The desirability function, *Ind Qual Control* 21(10), pp494-498, 1965
9. Segall MD, Champness E, Obrezanova O and Leeding C, Beyond profiling: Using ADMET models to guide decisions, *Chemistry & Biodiversity* 6(11), pp2,144-2,151, 2009
10. Chadwick AT and Segall MD, Overcoming psychological barriers to good discovery decisions, *Drug Discov Today* 15(13/14), pp561-569, 2010

Matthew Segall is Director and Chief Executive Officer at Optibrium Ltd (Cambridge, UK). He has an MSc in computation from the University of Oxford and a PhD in physics from the University of Cambridge. As Associate Director at Camitro, ArQule and then Inpharmatica, he led a team developing predictive ADME and decision-support tools for drug discovery. In 2006, he became responsible for management of Inpharmatica's ADME division, providing experimental services and the StarDrop software platform. Following the acquisition of Inpharmatica, Matt became Senior Director responsible for BioFocus's ADMET division, and in 2009 led a management buy-out of the StarDrop business to found Optibrium, a company that develops software to guide decision-making through complex and uncertain data. Email: matt@optibrium.com