Considering the Impact of 'Drug-like' Properties on the Chance of Success

Iskander Yusof and Matthew D. Segall

Optibrium Ltd., 7221 Cambridge Research Park, Beach Drive, Cambridge, CB25 9TL

Abstract

Many definitions of 'drug-like' compound properties have been published, based on analysis of simple molecular properties of successful drugs. These are typically presented as rules that indicate when a compound's properties differ significantly from those of the majority of drugs, which may indicate a higher risk of poor outcomes for *in vivo* pharmacokinetics or safety. We review the strengths and weaknesses of these rules and note, in particular, that overly rigid application of hard cut-offs can introduce artificial distinctions between similar compounds and runs the risk of missing valuable opportunities. Alternatively, compounds can be ranked according to their similarity to marketed drugs using a continuous measure of 'drug-likeness'. However, being 'similar' to known drugs does not necessarily mean that a compound is more likely to become a drug and we demonstrate how a new approach, utilising Bayesian methods, can be used to compare a set of successful drugs with a set of non-drug compounds in order to identify those properties whose values give the greatest distinction between the two sets, and hence the greatest increase in the likelihood of a compound becoming a successful drug. This analysis further illustrates that guidelines for 'drug-likeness' may not be generally applicable across all compound and target classes or therapeutic indications. Therefore, it may be more appropriate to consider specific guidelines for 'drug-likeness' dependent on the objectives of a project.

Definitions of 'Drug-likeness' relate simple molecular properties, such as molecular weight, physicochemical properties, number of rotatable bonds or number of aromatic rings to success against a drug discovery objective, usually achieving appropriate pharmacokinetics and safety. For example, Lipinski's, now ubiquitous, Rule of Five (RoF) [1] [2] defines four simple rules for the majority of compounds with good oral absorption:

- Logarithm of the octanol:water partition coefficient (logP) < 5
- Molecular weight (MW) < 500
- Number of Hydrogen bond donors (HBD) < 5
- Number of Hydrogen bond acceptors (HBA) < 10

Many alternative rules have been proposed relating compound properties such as polar surface area (PSA) and number of rotatable bonds (ROTB) to oral bioavailability. For example, Veber *et al.* investigated a dataset of 1,100 compounds with rat oral bioavailability data and found that those with ROTB of less than 10 and PSA less than 100 $Å^2$ had a higher probability of achieving oral bioavailability greater than 20% [3]. However, Lu *et al.* showed that the values of these cut-offs depended on the method used for the calculation of ROTB and PSA [4].

Rules have also been developed relating 'drug-like' properties to other outcomes. For example, Lovering *et al.* [5] identified the 'flatness' of a compound, as measured by the fraction of sp3 hybridized carbons, as an indicator of success in development. In a similar way Ritchie *et al.* [6] approached this question by relating the number of aromatic rings (AROM) to properties such as solubility, serum albumin binding and hERG inhibition. Furthermore, Hughes *et al.* explored the relationship of logP and PSA with observations of *in vivo* adverse toxicological events and found that compounds with logP >3 and PSA < 75 $Å^2$ have a significantly increased safety risk [7].

'Drug-like' properties and their associated rules have a number of strengths that have led to their popularity. In particular, they are very straightforward to understand and apply; the molecular properties on which they rely may be readily calculated and it is easy to identify a compound that fails on the criteria and how it may be modified to meet them. The rules for 'drug-like' properties provide useful guidelines to avoid common pitfalls that may be encountered downstream in discovery and development.

However, a number of weaknesses underlie this apparent simplicity. Most importantly, the rules for 'drug-like' properties apply only to the specific objective for which they were derived. For example, it is common to see the RoF used as a general definition of 'drug-like,' irrespective of the ultimate therapeutic goal of the project. However, the RoF relates only to oral absorption and the rules governing compounds for other routes of administration such as inhalation or topical application are quite different [8]; in these contexts, applying the RoF is likely to reject perfectly reasonable compounds.

In addition, the simple molecular properties which form the basis of the rules for 'drug-likeness' are only weakly predictive of a compound's ultimate biological properties. Given this, applying these rules as hard 'cut-offs,' or filters, runs the risk of missing valuable opportunities; for example, does a compound a MW of 501 have a significantly greater risk than a compound with MW of 499? It should also be noted that the RoF states that a compound has a higher risk of poor oral availability if it fails two or more of the criteria above, yet it is common to see all four criteria used independently as filters. Even when the RoF is applied as originally formulated, it is not highly predictive; table 1 shows the results of applying the RoF to 1191 marketed drugs labelled according to whether they have been approved for oral administration. One should not over interpret the results from a small, imbalanced set, however there are some notable observations that may be made: The RoF is not a guarantee of finding an orally available compound (we would not expect this as the RoF relates only to absorption and other factors such as first pass metabolism are important factors in determining oral bioavailability). Furthermore, many of the compounds that fail the ROF are orally administered and more non-orally administered compounds pass the RoF than fail, meaning that the specificity is poor.

		RoF result	
		Pass (≤ 1RoF failure)	**Fail (>1 RoF Failure)**
Route of administration	**Oral**	709	59
	Non-oral	333	90

From this we can see that passing the rules for 'drug-like' properties may bias the odds in favour of finding a successful compound, but applying these rules as hard filters runs the risk of rejecting valuable compounds. These rules are guidelines and should be given appropriate weight alongside other criteria for selecting a compound.

Many of the rules for 'drug-like' properties are derived from exploring the characteristics that successful drugs for the chosen objective have in common. Intuitively, it makes sense that compounds which differ significantly from the majority of known oral drugs will have a higher risk of failure due to inappropriate physicochemical or biological properties; there is little precedence for the success of such a compound and if successful it would be exceptional. However, being 'similar' to known drugs does not necessarily mean that a compound will have a better chance of becoming a drug than any other compound synthesised in the course of a drug discovery project. Some properties provide more information than others to distinguish successful compounds from unsuccessful. For example, if the distribution of a property is the same for oral drugs as that for all synthesised compounds then this property will tell us nothing about whether a compound will be more, or less, likely to become an oral drug. Rejecting a compound on the basis of such a characteristic would not be appropriate and would run the risk of missing valuable opportunities.

Quantitative Estimate of Drug-likeness

One approach to overcoming the problem of hard cut-offs and replace this with a continuous scale, by which chemistries can be ranked according to their drug-likeness, was recently published by Bickerton *et al.* [9]. The Quantitative Estimate of Drug-likeness (QED) relates the similarity of a compound's properties to those of oral drugs based on eight commonly used molecular properties: MW, logP, HBD, HBA, PSA, ROTB, AROM and count of alerts for undesirable substructures (ALERT).

The QED is based on a method for multi-parameter optimisation known as 'desirability functions' [10]. A desirability function relates the value of a compound characteristic to the 'desirability' of that outcome. The desirability is a number between zero and one, where a value of one indicates that the outcome is ideal and a value of zero indicates that the outcome is completely unacceptable.

To derive the QED metric, desirability functions were fitted to the distributions of the eight properties listed above for 771 marketed oral drugs. Using this method, a higher desirability score is assigned to a compound for a given property if the probability of observing that compound's property value is high for the marketed oral drugs. An example is shown in Figure 1 for molecular weight. The desirabilities of all the individual properties are combined into a single score, the QED, by taking their geometric mean, thus bringing together many of the factors considered by the drug-like property rules summarised above.

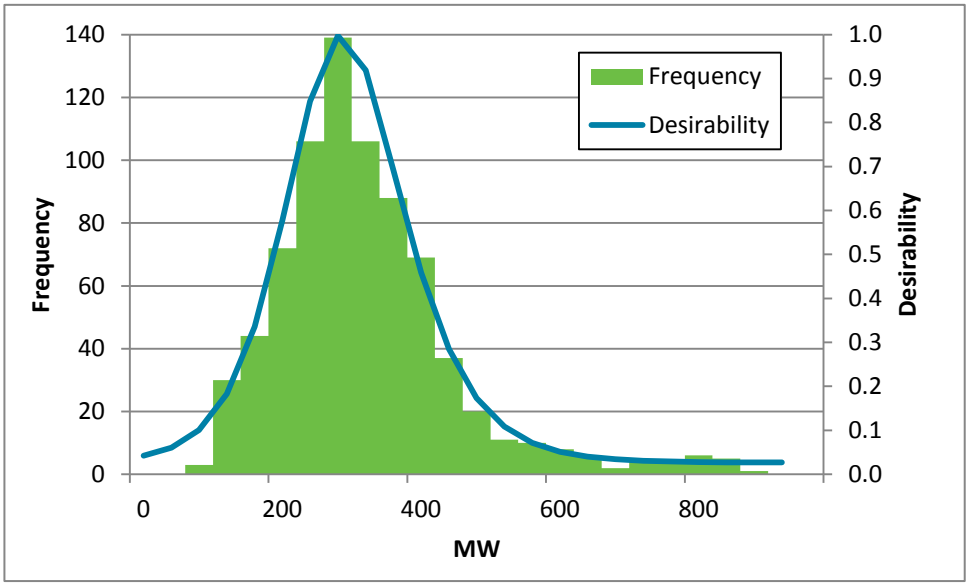

Figure 1 This graph shows the distribution of MW for a set of 771 orally absorbed small molecule drugs and a desirability function (blue), as used in QED, fitted to this distribution. The most desirable property values correspond to those most frequently observed in the set of drugs.

Bickerton *et al.* showed that the QED performed well in identifying a set of 771 marketed oral drugs, taken from DrugBank [11], from a set of 10,250 small-molecule ligands from the Protein Data Bank (PDB) ligand dictionary [12] (note that this was a different set of 771 oral drugs from that used to fit the desirability functions, although there was some overlap). Furthermore, the authors showed that the QED values agreed with medicinal chemists' subjective views of the attractiveness of compounds as hits on which to undertake further chemistry.

To aid interpretation, the desirabilities of the individual properties of a compound may be examined to identify those that differ most significantly from the majority of drugs which may, in turn, indicate strategies to improve the similarity.

Relative Drug Likelihood

The QED method overcomes one of the problems of 'drug-like' property rules by replacing hard cut-offs for individual properties with a single, continuous scale that combines many properties and allows chemistries to be ranked according to their similarity with oral drugs. However, as discussed above, a compound with a similar value of a property to known drugs does not necessarily have a higher chance of being a drug; some properties have more importance in distinguishing drugs from non-drugs and it is necessary to compare the properties of drugs and non-drugs to determine this. In essence, we are more interested in the properties that make drugs *different* from other compounds that might be considered in the course of a drug discovery project than the properties that drugs have in common.

In order to make this comparison, we may consider that a desirable value of a property is one that increases the probability of identifying a drug. Bayesian probability theory allows us to approach this question in a quantitative manner by comparing the distribution of properties of drugs with those of non-drugs. The underlying concept is that property values that are more likely to be observed in drugs than non-drugs are more desirable, because they imply an higher likelihood of a compound being a drug relative to a randomly selected compound (the mathematics are described in Box 1). This relative likelihood can be calculated as a function of a property value. A relative likelihood above one is good because it indicates that the property value increases the chance of a compound being successful while, conversely, a relative likelihood below one is bad. In a similar way to QED, the relative likelihoods of the individual properties of a compound can then be combined into a single score, the Relative Drug Likelihood (RDL), allowing chemistries to be prioritised.

One drawback of this approach is that it is necessary to specify a 'negative' set to represent the population of typical non-drug compounds from which we would like to identify potential drugs and an advantage of a similarity-based approach, such as QED, is that it may be applied in the absence of information on negative compounds. The suitable population of non-drugs is not one of totally random compounds, because the compounds chosen for synthesis in drug discovery would normally have been preselected in some way. A chemist's experienced eye cast over potential compounds is, itself, a form of pre-selection.

The choice of appropriate 'positive' (drug) and 'negative' (non-drug) sets for this analysis depends on the objective that we would like to explore. For example, we may wish to understand the properties that, in general, distinguish oral drugs from other compounds explored in small molecule drug discovery. To address this question, an appropriate negative set, representing a large, diverse set of compounds that have been considered in the course of medicinal chemistry projects, is provided by the ChEMBL database (https://www.ebi.ac.uk/chembldb/). To allow a direct comparison with the QED method, we have used the eight properties considered in QED for 1,000 randomly selected compounds from the ChEMBL database considered as a negative set and compared these with the 771 marketed oral drugs published in Bickerton *et al.* [9]. The property distributions and resulting relative likelihood functions are shown in Figure 2.

The performance of the QED and RDL methods can be compared for selection of a different set of 771 marketed oral drugs, taken from DrugBank, from the 656,737 compounds from the ChEMBL dataset remaining after the 1000 compounds used to construct the RDL have been removed (note that, for consistency, the second set of 771 drugs is the same as that used to benchmark the performance of QED in Bickerton *et al.* [9]). The receiver operating characteristic (ROC) plot for selection of the oral drugs from this larger set is shown in Figure 3. From this, it can be seen that the performance of the RDL is significantly higher than the QED metrics, indicating that the additional information derived from the comparison of positive and negative sets increases the ability to accurately distinguish successful compounds from the 'background' of compounds explored in drug discovery projects.

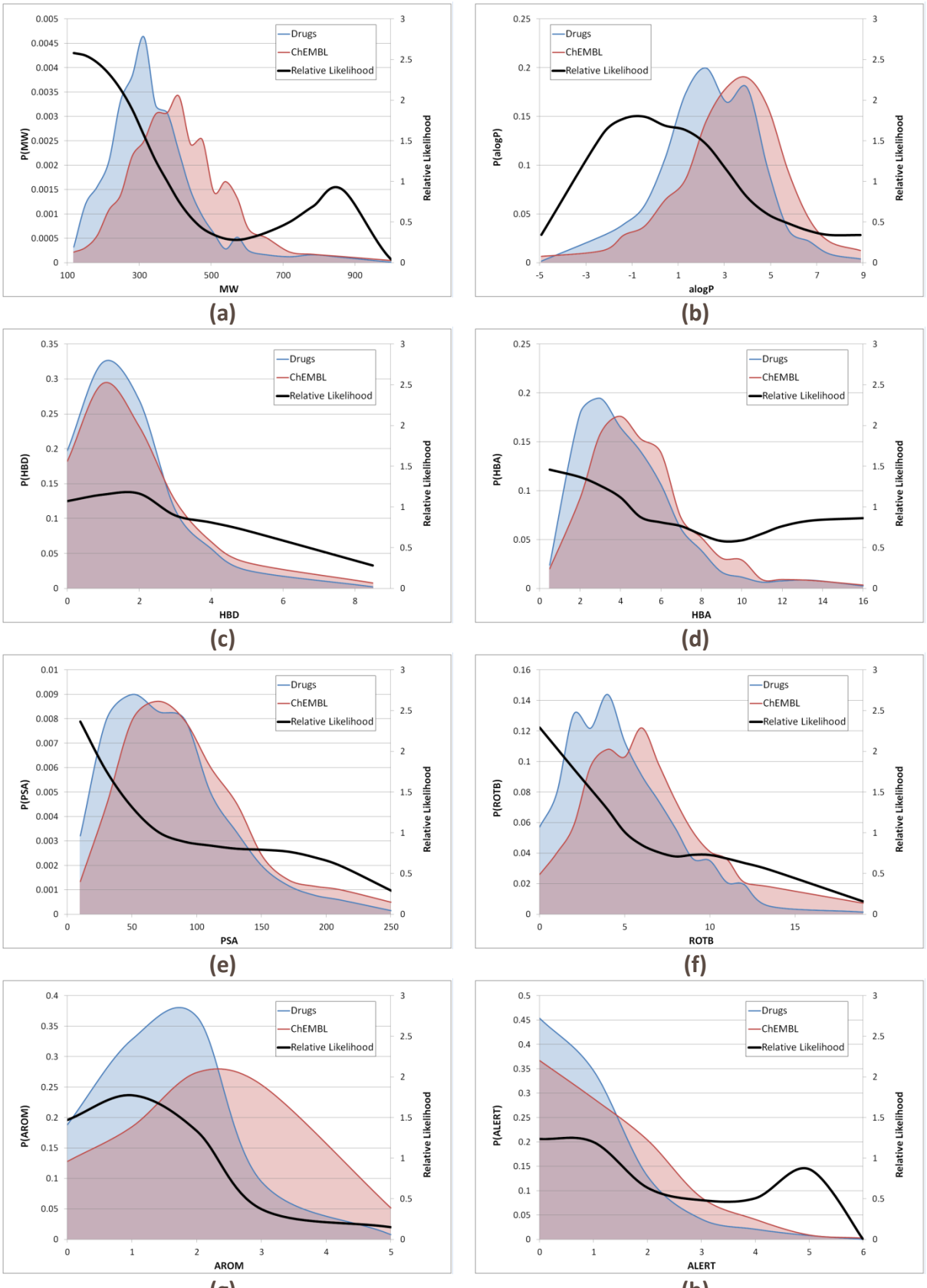

Figure 2 Graphs showing distributions of eight molecular properties studied for a set of 771 orally absorbed small molecule drugs (blue) and compounds in the ChEMBL database (red). MW (a), lipophilicity estimated by atom-based prediction of ALOGP (b), number of HBDs (c), number of HBAs (d), PSA (e), number of ROTBs (f), number of AROMs (g) and number of ALERTS (h).The resulting function indicating the relative likelihood of a compound being an oral drug (right vertical axis) is shown as a black line.

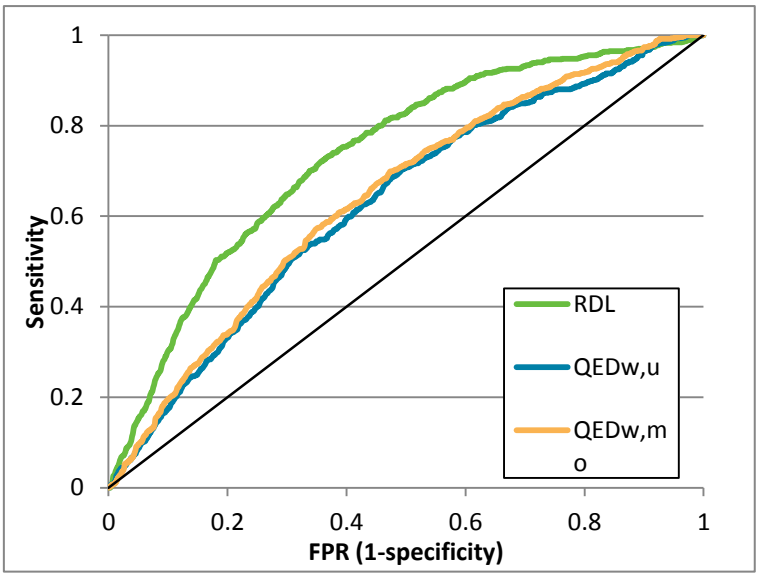

Figure 3 Receiver operating characteristic (ROC) plot of the true positive rate (sensitivity) against the false positive rate (FPR (1 - specificity)) for the classification of compounds as orally absorbed drugs or otherwise using RDL and the unweighted ($QED_{w,u}$) and mean optimal entropy ($QED_{w,mo}$) QED schemes. In this case, a set of 771 orally administered small molecule drugs were identified from a negative set of >650k compounds from ChEMBL.

A perfect classifier would be represented by the point in the top left and a performance below the identity line indicates worse performance than a random classification. A greater area under the curve (AUC) for a classifier indicates higher performance; the AUC for RDL is 0.73, $QED_{w,mo}$ is 0.64 and $QED_{w,u}$ is 0.63.

This observation is particularly notable because the compounds in the negative data set from ChEMBL are already biased by medicinal chemists experience and the application of drug-like property rules over several years. Therefore, identification of compounds from this set will be harder than from a truly random selection of compounds.

If we examine the graphs shown in Figure 2, we can make some interesting additional observations regarding the relative likelihood functions of the individual properties. The properties that provide the greatest ability to distinguish between the drugs and non-drugs are MW, PSA and ROTB, as indicated by the high values of the desirability functions for some values of these properties; in some cases indicating an enhancement in the relative likelihood of greater than two. Similarly, very low relative likelihoods indicate property values that correspond to a high risk of failing to achieve acceptable bioavailability after oral administration. Conversely, it can be seen that HBA provides relatively low power of discrimination, because the relative likelihood remains within a small range close to 1.

It is also notable that the relative likelihoods for some properties, e.g. MW, HBD, ROTB and ALERTS, exhibit a counter-intuitive increase for high values (although the value remains below 1.0, indicating that these property values are not good). There are, of course, a number of oral drugs with high values for MW, for example macrolides such as Sirolimus, and Erythromycin. It is hypothesised that these large, flexible molecules do not 'obey the rules' governing typical small molecule drugs; they may undergo hydrophobic collapse and, in practice, have a lower lipophilicity and volume than a more rigid compound of the same molecular weight and calculated logP. Similarly, there are a number of drugs such as Cimetidine, Cefpodoxime and Dantrolene which match several of the structural alerts that were applied. This may indicate that some of these alerts are not sufficiently specific or, alternatively, that there is an oversensitivity toward the exploration of chemistry containing substructures that have previously been associated with toxicity. Therefore, the increase in the relative likelihood may indicate that these regions of 'chemical space' have been underexplored in drug discovery projects relative to the number of oral drugs that have resulted from this chemistry. Alternatively, it is possible that these effects are artefacts of the data in ChEMBL, which have been abstracted from journals that may have greater focus on 'traditional' small molecule medicinal chemistry.

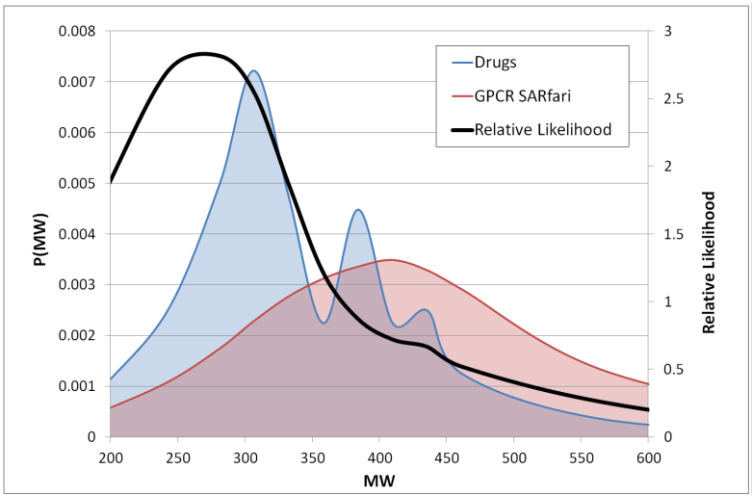

Figure 4 Graph showing the distributions of MW for oral administered drugs for GPCR targets (blue) and compounds screened against GPCR targets in the GPCR SARfari database (red). The resulting function indicating the relative likelihood of a compound being an oral drug for a target in this class (right vertical axis) is shown as a black line. This can be compared with the equivalent graph in Figure 2(a) comparing a general set of orally administered drugs and drug discovery compounds. This illustrates that the property values that distinguish successful from unsuccessful compounds can depend strongly on the specific objective.

These unexpected trends further highlight the challenge of deriving general rules for the properties that make a compound drug-like or otherwise. Many observations of drug like properties do not account for the different contexts from which the drugs analysed are derived and the requirements will depend heavily on the objective of the project, such as target class and therapeutic indication; thus, for example, antibiotic drugs do not provide a good guide to appropriate characteristics for a compound intended as a CNS drug.

An advantage of the RDL approach is that it can be easily adapted to different objectives. Example code to generate an RDL metric for any objective is provided in the supplementary information for this article. As a further illustrative example, we generated a target-class focussed RDL desirability function for MW using a positive set of 156 orally available drugs for G-Protein Coupled Receptor (GPCR) targets and a corresponding negative set of compounds screened against GPCR targets, derived from the GPCR SARfari database (https://www.ebi.ac.uk/chembl/sarfari/gpcrsarfari). The resulting distributions and relative likelihood function for MW are shown in Figure 4, from which we can see that the counterintuitive asymptotic behaviour seen in the relative likelihood function derived from general drugs and medicinal chemistry compounds is not present. This is because exploration of chemistry with high molecular weight has not yielded oral drugs for this target class.

As discussed above, the simple 'dug-like' molecular properties correlate only weakly with the *in vivo* disposition of a compound. However, it may still be useful to consider the relative likelihood of achieving objectives relating to parameters such as pharmacokinetic endpoints in order to identify trends with which to guide optimisation. As a brief example, we can consider the objective of identifying a compound with volume of distribution at steady state (VD_{ss}) of less than 1 L/kg which may be desirable to avoid high tissue binding. Using a database of pharmacokinetic parameters in human, published by Obach *et al.* [13], we identified a 'positive' set of 345 compounds and a 'negative' set of 322 compounds (see supplementary information). The resulting distributions and relative likelihood function for logP are shown in Figure 5, from which it can be seen that a negative logP, indicating high water solubility, significantly increases the likelihood of achieving a low VD_{ss}. Furthermore, the likelihood of achieving low VD_{ss} falls below average if the logP of a compound is greater than approximately 1. We can understand this because compounds with higher water solubility are likely to remain in solution in the plasma or extracellular fluid and will not readily bind to lipid tissues.

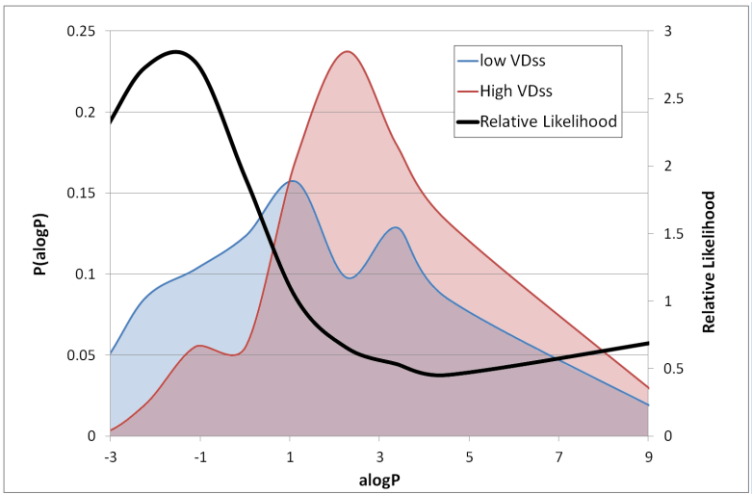

Figure 5 Graph showing the distributions of logP for compounds with low VD_{ss} (blue) and high VD_{ss} (red) based on a cut-off of 1 L/kg. The resulting relative likelihood function for identifying a compound with low VD_{ss} is shown as a black line (right vertical axis). This illustrates how the relative likelihood approach can identify property trends for individual parameters such as PK endpoints.

Conclusion

Consideration of drug-like properties can provide useful guidance when considering chemistries to pursue in the search for a novel drug; where possible, it is preferable to work in areas of chemistry that have lower risk of issues related to poor pharmacokinetics or safety. However, rules and scores for drug-like properties should be given appropriate weight in decision-making and should be balanced against other requirements for a drug, not least potency and, as we have discussed in this article, hard cut-offs should be avoided. It is also important to remember that these rules and trends are generated for specific objectives, most commonly oral absorption or bioavailability, and may not be applicable in other scenarios. If possible, the guidelines should be tailored to specific therapeutic or target classes and routes of administration.

Drug-like property guidelines are most commonly applied early in a drug discovery project, for example in design of screening libraries or hit prioritisation, where experimental data on many properties are limited or unavailable. When experimental data are available for compound potency and other properties, these clearly provide much more information about their desirability for further investigation. In this scenario, the search for drug-like properties has a much more limited role to play, perhaps providing guidelines regarding possible directions to improve poor experimental outcomes. For example, in the cases of QED and RDL, one can clearly identify the specific properties that contribute most to a reduction in the desirability or likelihood of success of a compound.

The guidelines for drug-like properties are, by definition, based on historical precedence. It is possible that exploration of new chemical space will yield new approaches to achieve appropriate pharmacokinetics and safety. Indeed, some therapeutic strategies, such as inhibition of protein-protein interactions or epigenetic targets, may demand the discovery of compounds that 'break the rules'. However, unprecedented approaches carry additional risks.

Finally, it is important to remember that while having good drug-like properties or a high QED or RDL indicates a higher chance of a compound becoming a drug, these are far from guarantees of success; the absolute probability is still small. There are many hurdles that a compound must overcome to become a successful drug and far more drug-like compounds fail than succeed. An important source of risk, that cannot be addressed by the properties of a candidate small molecule drug, is the biological risk associated with the therapeutic target. Lack of efficacy or toxicity associated with the biological mechanism of the target remain major causes of failure in later stages of clinical development. Introducing specific data on risk factors through good predictive or experimental methods and assessing the overall balance of properties using a multi-parameter optimisation

[14] approach, as a compound progresses, will allow these additional sources of risk to be mitigated to achieve greater confidence in its success.

Acknowledgements

The authors would like to thank Edmund Champness, Nick Foster, Chris Leeding, James Chisholm and Steve Smith for useful discussions and comments on the draft manuscript.

Supplementary Information

A full derivation of the RDL method along with example code to calculate the RDL for any positive and negative sets of compounds are provided in the supplementary information. Furthermore, the positive and negative data sets used in the examples presented herein are also provided.

References

1 Lipinski, C.A., Lombardo, F., Dominy, B.W., and Feeney, P.J. (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.*, 23, 3-25.

2 Leeson, P.D. and Springthorpe, B. (2007) The influence of drug-like concepts on decision-making in medicinal chemistry. *Nat. Rev. Drug Discov.*, 6, 881-890.

3 Veber, D.F., Johnson, S.R., Cheng, H.Y., Smith, B.R., Ward, K.W., and Kopple, K.D. (2002) Molecular properties that influence the oral bioavailability of drug candidates. *J. Med. Chem*, 45, 2615-2623.

4 Lu, J.J., Crimin, K., Goodwin, J.T. et al. (2004) Influence of Molecular Flexibility and Polar Surface Area Metrics on Oral Bioavailability in the Rat. *J. Med. Chem.*, 47, 6104-6107.

5 Lovering, F, Bikker, J., and Humblet, C. (2009) Escape from flatland: increasing saturation as an approach to improving clinical success. *J. Med. Chem.*, 52, 6752-6756.

6 Ritchie, T.J. and Macdonald, S.J.F. (2009) The impact of aromatic ring count on compound developability – are too many aromatic rings a liability in drug design? *Drug Discov. Today*, 14, 1011-1020.

7 Hughes, J.D., Blagg, J., Price, D.A. et al. (2008) Physicochemical drug properties associated with in vivo toxicological outcomes. *Bioorg. Med. Chem. Lett.*, 18, 4872-4875.

8 Choy, Y.B. and Prausnitz, M.R. (2011) The rule of five for non-oral routes of drug delivery: ophthalmic, inhalation and transdermal. *Pharm. Res.*, 28, 943-948.

9 Bickerton, G.R., Paolini, G.V., Besnard, J., Muresan, S., and Hopkins, A.L. (2012) Quantifying the chemical beauty of drugs. *Nature Chemistry*, 4, 90-98.

10 Harrington, E.C. (1965) The desirability function. *Ind. Qual. Control*, 21, 494-498.

11 Wishart, DS, Knox, C, Guo, AC et al. (2008) DrugBank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Res.*, 36, D901-6.

12 Dimitropoulos, D., Ionides, J., and Henrick, K. In Baxevanis, A. D. et al., eds., *Current Protocols in Bioinformatics*. Wiley, 2006.

13 Obach, R.S., Lombardo, F., and Waters, N.J. (2008) Trend analysis of a database of intravenous pharmacokinetic parameters in humans for 670 drug compounds. *Drug Metab. Dispos.*, 36, 1385-1405.

14 Segall, M.D. (2012) Multi-Parameter Optimization: Identifying high quality compounds with a balance of properties. *Curr. Pharm. Des.*, 18, 1292-1310.

Box 1: Bayesian Probability and Relative Likelihood

Bayesian probability theory allows us to infer the probability of an outcome for a new compound based on data observed for previous compounds with known outcomes. In this context, Bayes' theorem states:

$$P(D|X) = \frac{P(X|D)P(D)}{P(X)}.$$

Here, P(D|X) is the probability of a compound being a drug given the value of a property X, which this is what we would like to know; in Bayesian terms this is known as the posterior. P(X|D) is the probability of the property X given that a compound is a drug, known as the likelihood. P(X) is the probability distribution for the property X for all compounds, whether drugs or not, and is known as the evidence. Finally, P(D) is the probability of a compound being a successful drug, given no further information, which is the 'prior probability' of a compound being a drug (a very small number, based on historical evidence!).

In a similar way, we can compute the posterior probability of a compound not being a drug, given the value of a property X:

$$P(D'|X) = \frac{P(X|D')P(D')}{P(X)}.$$

We can then determine if a compound is *more likely* to be a drug based on its property values, by taking the ratio between the posterior probability of a compound being a drug and not being a drug to arrive at the equation:

$$\frac{P(D|X)}{P(D'|X)} = \frac{P(X|D)}{P(X|D')}\frac{P(D)}{P(D')}.$$

A desirable value for a property is one for which this ratio is relatively high, i.e. the probability of a compound being a drug is increased relative to the probability of it not being a drug. P(D) and P(D') are unknown, but they are constants and hence this ratio is directly proportional to the ratio of the likelihoods of property X for drug and non-drugs. Therefore, our measure of desirability is the *relative likelihood*

$$d(x) = \frac{P(X=x|D)}{P(X=x|D')}.$$

The relative likelihood functions above for different properties can be combined into a single desirability index, the Relative Drug Likelihood (RDL), by taking the geometric mean of the desirabilities of the individual properties:

$$\mathrm{RDL} = \exp\left(\frac{1}{n}\sum_{i=1}^{n} \ln\left(d_i(x_i)\right)\right).$$

The mathematical details of the derivation of the RDL metric are provided in the supplementary information.