*The traditional physical-property-based argument for drug attrition is developed and extended to one that incorporates drug–transporter interactions. A new algorithm is proposed that facilitates the evaluation of this hybrid property space.*



# A new paradigm for navigating compound property related drug attrition

**Patrick Barton[1] and Robert J. Riley[2]**

[1] School of life Sciences, University of Nottingham, Nottingham, UK
[2] Evotec UK, Abingdon, Oxfordshire, UK

**Improving the efficiency of drug discovery remains a major focus for the pharmaceutical industry. Toxicity accounts for 90% of withdrawals and major early-stage terminations relate to suboptimal efficacy and safety. Traditional oral drug space is well defined with respect to physicochemical properties and ADMET risks but increased focus on ligand-lipophilicity efficiency, maximizing enthalpy contributions and new target classes challenge this paradigm. A hybrid space has been identified that combines physical properties and key interactions attributable to drug transporters. A novel algorithm is proposed that incorporates drug–transporter interactions and its utility evaluated against popular ligand efficiency indices. Simply reducing the bulk properties of compounds can exchange one problem for another and creates high-risk areas that challenge the successful delivery from a balanced portfolio.**

## Introduction

Recent statistics indicate that the well-documented challenges facing the pharmaceutical industry continue with rising costs and attrition together with increased pressure from regulators and payors being major contributors [1–4]. For over a decade now, companies have rightly focused on compound properties to stem ADMET (Absorption, Distribution, Metabolism, Excretion, Toxicity)-related attrition and increase the likelihood of candidates surviving early toxicity assessment, progressing into humans and ideally achieving appropriate exposure. The latter should not only be viewed from the systemic circulation but also at the target tissue for efficacy without paying the penalty at tissues associated with (often) off-target safety concerns.

Clearly, intrinsic drug affinity at desired or nondesired (potentially toxic) targets is governed by a combination of bulk physicochemical properties [5] and specific structural features for example basic p$K_a$ for hERG (human ether-a-go-go-related gene) [6,7], phospholipidosis (PLD) inducers [8], structural motifs for reactive metabolites [9,10] and transporter interactions such as with the bile salt excretory pump (BSEP) [11]. Nevertheless, sufficient exposure at the target tissue will still be required in addition to inherent affinity. Optimizing the exposure of potent compounds at the desired site of action and in tissues associated with toxicity is fundamental to addressing attrition

*Corresponding author:* Barton, P. (patrick.barton@nottingham.ac.uk)

**Patrick Barton** is Associate Professor of Drug Discovery in the School of Life Sciences at the University of Nottingham, UK. His main research interests are in drug metabolism and pharmacokinetics/pharmacodynamics with a particular interest in the role of drug-transporters and their influence on ADMET properties. Prior to joining the University of Nottingham Dr Barton had 20 years' experience in the pharmaceutical industry working in drug discovery for AstraZeneca. During this period he held the positions of Associate Director of Discovery Drug Metabolism, Group Leader of Computational and Physical Chemistry and Global Leader of Predictive Drug Metabolism. He supported multiple projects across the respiratory and inflammation, cardiovascular and oncology portfolio resulting in multiple clinical candidates.

**Robert J. Riley** is currently Vice President of Drug Discovery at Evotec. His current responsibilities include managing drug discovery alliances with several partners and leadership of the DMPK and Separations and Analytical Services groups. He was previously head of Chemistry and DMPK within Respiratory and Inflammation at AstraZeneca where he supported the delivery of multiple candidate drugs, several of which are still in late-stage development. He was also directly involved in the discovery of Brilinta®. He retains a keen interest in the relationship between compound properties, drug design and 'survivability'.



1359-6446/
http://dx.doi.org/10.1016/j.drudis.2015.09.010

via efficacy and safety [12,13]. In turn, key determinants of drug exposure are the molecular properties of the drug candidate [14–16]. The overarching goal therefore remains to achieve good pharmacokinetics and optimized exposure from a modest dose often preferring the oral route. Therefore, an enhanced understanding of the interplay between molecular structure and exposure remains paramount to successful drug discovery and we move to a paradigm based upon 'design better, develop faster and succeed more often' from the rather uninspiring but often quoted mantra of 'fail fast and fail cheap'.

## Drug-like metrics: rules and indices

Emphasis has therefore been placed on defining a series of rules and indices for compounds at various stages of preclinical development. It is important to distinguish between efficiency metrics that incorporate the affinity of compounds for their targets from more-generic drug-like properties. The term drug-likeness is applied in drug discovery to identify virtual or real molecules that occupy what is considered to be drug-like chemical space, based on physicochemical properties [17–19]. Often this entails examining the calculated physicochemical properties of molecules and favoring those found in marketed drug molecules or clinical candidates. It is now widely accepted that drug-like compounds tend to demonstrate certain favorable ADMET properties such as aqueous solubility and cell permeability. One approach is to use property cutoff filters above which compounds fail, for example the now famous Rule of Five (Ro5) [20].

Clearly, the value of these filters depends on the method used for their calculation and their associated errors [21]. Nevertheless, this concept and resulting rules have gained popularity as a result of their simplicity; they are easy to interpret and the molecular properties on which they rely can be readily calculated and it is easy to identify compounds that meet or fail the criteria and to develop optimization strategies. The underlying rationale is to increase the probability of designing, prioritizing and thereby developing compounds with an acceptable ADMET profile and to minimize property space that is sparsely populated with marketed drugs (exception space). Precedence for the success of the latter group is rare and would be exceptional [22].

Recognition of the property inflation or molecular obesity associated with compound progression has led to more rigorous rules or filters, for example Rule of Four for leads [23] and the Rule of Three for fragment-based lead generation [24]. However, their specificity can be poor and concern has also been expressed that these rules or indices should not be seen as hard cutoffs but rather guidelines given the errors in their measurement [25,26], and that application of multiparametric optimization methods would add considerable insight and benefit [27,28]. Moreover, several valuable therapeutic discoveries lie in chemical space beyond typical oral drug space – so-termed exception or beyond the Ro5 (bRo5) space [22].

It can therefore be concluded that adopting the rules for drug-like properties biases the odds in favor of finding a successful compound, but applying these rules as rigid filters runs the risk of rejecting valuable compounds. Consequently, several groups have cautioned that such assessment should go beyond such simple rules. The introduction of the term ligand efficiency (LE) was the first of several expressions that have recognized the importance of combining physicochemical properties and potency at the intended target [29,30]. A comprehensive review of the development of these metrics or indices is beyond the scope of this work. The interested reader is directed to Fig. 1, which provides the chronology of major developments in this area.

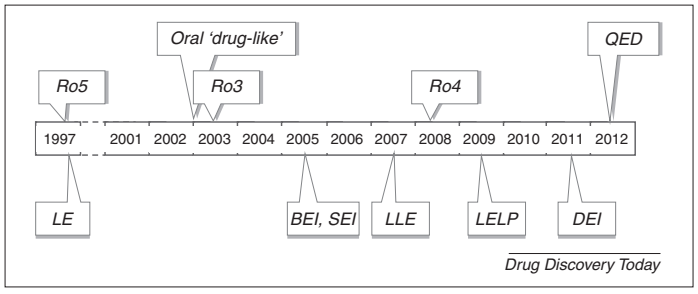


**FIGURE 1**

Schematic timeline illustrating major developments of generic drug-like properties (above timeline) and efficiency metrics (below timeline). Rule of Five (Ro5) [20], ligand efficiency (LE) [59], oral drug-like [15,49], Rule of Three (Ro3) [60], binding efficiency indices (BEI) [49], surface efficiency indices (SEI) [49], ligand-lipophilicity efficiency (LLE) [29,30], Rule of Four (Ro4) [23], lipophilicity-corrected ligand efficiency (LELP) [61], drug efficiency index (DEI) [62], quantitative estimate of drug likeness (QED) [25].

Analysis of historical oral marketed drug data might not be reflective of future trends in preclinical drug discovery because increasingly different approaches are adopted particularly for target engagement. These include optimization of receptor kinetics through consideration of enthalpy binding [31], novel target classes [32,33], allosteric modulators [34] and targeting novel signaling pathways such as β-arrestin [35]. These relatively recent approaches could yield compounds with a physical property distribution distinct from that previously observed and characterized. If this is coupled to classical strategies for the optimization of ADMET properties using traditional methodologies with established compound collections, vigilance and the early identification of chemical space which could deviate from our previous experience is paramount. Maintaining compound properties that are based on historical drugs while targeting novel mechanisms for efficacy might not serve the industry well. Therefore we need to be able to respond rapidly to an ever-changing concept of what constitutes drug-like space.

## Compound properties and their relation to attrition

A recent study by Pfizer [36] reported an analysis of compound attrition resulting from preclinical toxicity in rat demonstrated an increased likelihood of failure if compounds were basic and had $\log P > 3$, PSA $< 75$ Å$^2$. One interpretation of these findings would be that lipophilic, relatively promiscuous compounds with little polar functionality tend to have an increased incidence of toxicity. A similar finding was also reported around this time for a range of veterinary drugs [37] and reports have indicated the importance of lipophilicity in hepatotoxicity with a $\log P > 3$ being suggested as a threshold, together with considerations of total daily dose. A similar analysis by AstraZeneca [38] on their compound failures revealed a different profile, with the majority of attrition occurring with PSA $> 75$ Å$^2$ and $\log P < 3$. Although attrition in the high-log *P*–low-PSA space can readily be rationalized via consideration of promiscuity and interactions across a range of systems [30], the

apparent failure of compounds in the low-clog *P*–high-PSA region is arguably less intuitive.

This apparent dichotomy can be rationalized by several factors such as: chemist bias in design strategies and synthesis; targets within the companies' portfolios at any given time. Such factors have been theorized by others: a recent analysis [2] has suggested that companies like Pfizer and Vertex synthesize compounds with low molecular weight (MW), low log *P* and few H-bond donors (HBDs) and/or H-bond acceptors (HBAs); whereas compounds profiled in companies such as AstraZeneca and Roche appear to achieve low log *P* while increasing HBDs and HBAs and MW.

Interestingly, an analysis of AstraZeneca and Bayer compounds has also revealed marked differences in their compound collections reemphasizing the 'organizational factor' [39]. An update from Pfizer has demonstrated a shift in their attrition data with more compounds appearing in the PSA > 75 Å$^2$, log *P* < 3 space again indicating the importance of the time element and probably targets within a portfolio and their associated chemical strategies. The fundamental role played by exposure was also re-emphasized in this publication [40].

These reports are intriguing because they have been very powerful, abundant and yielded valuable and influential datasets. However, although clearly philatelic, little mechanistic or functional rationale has been provided. The recent paper utilizing the quantitative estimate of drug-likeness (QED) algorithm started to address this link with respect to ADMET properties but requires numerous physicochemical descriptors. Given the obvious link between drug disposition and exposure in target tissues for efficacy and toxicity, we endeavored to analyze further the 3/75 rule with respect to the Biopharmaceutical Drug Disposition Classification System (BDDCS) (Fig. 2).

## BDDCS and transporter interaction

The BDDCS [41] looks to separate compounds into four classes based on their permeability and solubility properties. The system has been insightful in predicting drug–transporter interactions, particularly gastrointestinal and hepatic efflux [multi-drug-resistance protein (MDR)1] and hepatic uptake [e.g. organic anion transporting polypeptide (OATP)1B1 and OATP1B3] [42]. This model has highlighted that compounds falling into BDDCS class III tend to be soluble, metabolically stable, poorly permeable and demonstrate disposition related to transporter interactions. Figure 2 also shows that, for this dataset of >900 compounds [38], class III compounds separate significantly from other BDDCS groups based upon a simple bivariate plot of PSA and log *P*. The observation that compounds with PSA > 75 Å$^2$, *c* log *P* < 3 tend to be class III is important and suggested that this rather simple, visual relationship should be examined further for statistical significance.

The hypothesis that compounds with PSA > 75 Å$^2$, log *P* < 3 belong to BDDCS class III whereas compounds for which this condition is not true belong to class I, II or class IV is a typical 2 class situation and is normally summarized in a confusion or error matrix that cross-tabulates the observed and predicted patterns as shown in Tables 1 and 2. The analysis indicates a correct classification rate of 74.3% which supports the hypothesis that BDDCS class III compounds tend to have PSA > 75 Å$^2$, log *P* < 3. However a more rigorous predictive measure of the model is the kappa (κ) value, which considers chance correlations within the data. In this case a κ value of 0.41 indicates a strong model with moderate-to-high predictive power [43].

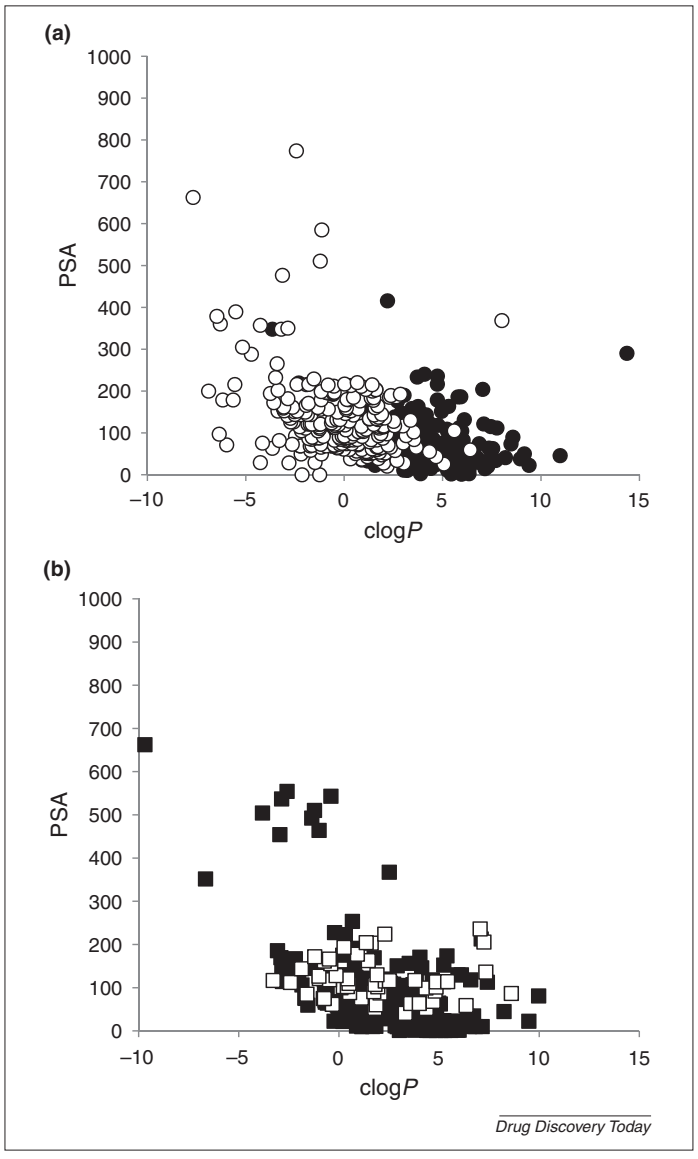


**FIGURE 2**

**(a)** Mapping of >900 compounds on a bivariate plot of PSA versus *c* log *P* – Biopharmaceutical Drug Disposition Classification System (BDDCS) class II compounds (dark circle) compared with class III compounds (open circle). **(b)** Comparison of BDDCS class I (closed squares) and class IV (open squares).

It is important to specify that this simple, binary hypothesis is central to our thinking because the authors are suggesting that significant attrition can be associated with class III compounds, because they tend to be transporter substrates or inhibitors. It also

**TABLE 1**

**Confusion matrix for the classification of Biopharmaceutical Drug Disposition Classification System (BDDCS) class III compounds from other BDDCS classes using *c* log *P* and PSA**

| | Observed class III | Observed other |
|---|---|---|
| **Predicted class III** | 170 | 77 |
| **Predicted other** | 157 | 506 |

**TABLE 2**

**Statistical analysis for classification of Biopharmaceutical Drug Disposition Classification System (BDDCS) class III compounds from other BDDCS classes using *c* log *P* and PSA**

| Measure | Model prediction |
|---|---|
| **Correct classification rate** | 74.3% |
| **Sensitivity** | 51.9% |
| **False-positive rate** | 13.2% |
| **False-negative rate** | 48.0% |
| **Positive predictive power** | 68.8% |
| **Negative predictive power** | 76.3% |
| **Kappa (κ)** | 0.41 |

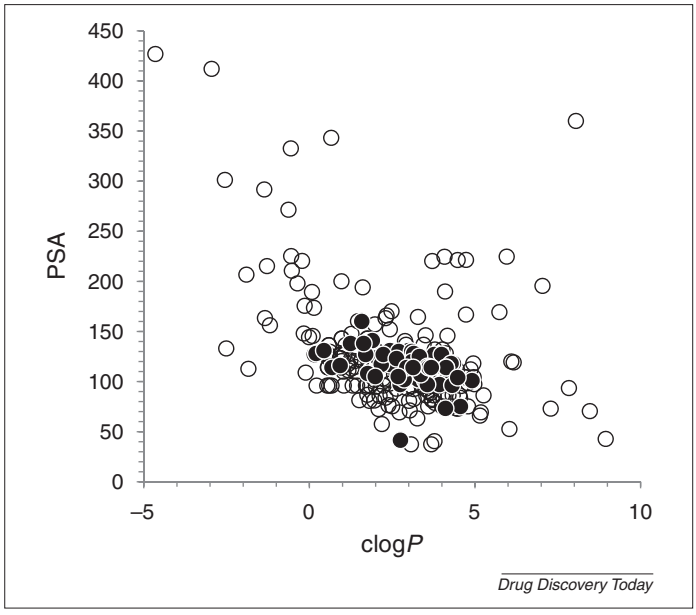


**FIGURE 3**

Mapping of *c* log *P* versus PSA demonstrating the overlap between rat (closed circles) and human (open circles) organic anion transport proteins ligands.

supports a more mechanistic interpretation regarding what has previously been a simple empirical observation.

This analysis suggests that the compounds reported by AstraZeneca [38], and more recently by Pfizer, in profiling their attrition [36] are probably class III – poorly permeable, metabolically stable substrates for uptake transporters. Such compounds can have concentrations in certain tissues in excess of the circulating plasma that could well be species-dependent and confound interpretation and translation of efficacy and, perhaps more importantly, toxicology studies. Profiling compounds recently reported to be substrates for rat [44] and human [44–47] OATPs supported the hypothesis that they do conform to the paradigm of PSA > 75 Å$^2$, *c* log *P* < 3, although the influence of log *P* is less obvious in one of the more limited human datasets (Fig. 3). Building on this exercise, a further challenge became defining an algorithm that could quantify the relative risks in this PSA/log *P* transporter space. Several groups have proposed and/or reviewed a range of ligand efficiency metrics, which continue to gain popularity [48].

## ADMET efficiency index (AEI): a hybrid index encompassing ligand-lipophilicity efficiency (LLE) and BDDCS space

As discussed earlier, consideration of PSA [as in surface efficiency index (SEI)] [49] in such metrics has perhaps not been as prominent as it could be, although it has been suggested by several groups as potentially offering differentiation beyond lipophilicity alone. Our analysis of the PSA/log *P* grid led us to consider an algorithm that encompassed the key benefits of LLE (pActivity – log *P*) yet also factored in consideration of PSA, typically normalized to a value documented to be compatible with extensive human absorption [49,50]. This new indirect efficiency metric that aims to encompass the LLE but encodes for an ADMET score is described in Eqn 1:

$$\mathrm{AEI} = \frac{(\mathrm{pActivity} - |\log P|)}{\mathrm{PSA}} \times 100 \qquad (1)$$

This algorithm benefits from LLE which has been demonstrated to be effective in reducing attrition through a secondary pharmacology argument generally considered to be nonspecific in nature [30]. The PSA term is used to score the data with respect to potential transporter interactions for which moderate lipophilicity coupled with a propensity for H-bond interaction contribute to the generic pharmacophore [51]. This is also supported by the observation that BDDCS class III compounds have PSA > 75 Å$^2$ and *c* log *P* < 3 and are known to have an increased prevalence of transporter interactions. The authors also considered the need for log *P* and PSA terms in this index because both can be considered ways in which the general polarity of compound can be modified. However, the need for both terms is rationalized by that fact that the octanol:water partition coefficient, log *P*, is a poor descriptor for H-bonding ability of a compound owing to the relatively high concentration of water in saturated octanol [52]. Hence, to accommodate general lipophilicity and H-bonding potential into this index, log *P* and PSA are required. The form of the log *P* term in the AEI is the modulus and can be rationalized. Firstly, to moderate the influence of the LLE term (pActivity – log *P*) from overpowering the influence of PSA in AEI the modulus of log *P* has to be used. It is clear that the term in LLE becomes positive when log *P* is negative and, although this is useful in terms of rationalizing ligand efficiency, this becomes detrimental in terms of rationalizing ADMET properties such as permeability and hence absorption which should become less favorable as log *P* becomes negative. By contrast, large positive values of log *P* are detrimental to LLE and AEI.

During the preparation of this manuscript, an assessment of the utility of the QED algorithm in its ability to predict the pharmaceutical and pharmacokinetic profiles of oral drug molecules was published. In this analysis a dataset of 771 oral drug molecules were separated into the top and bottom 25 compounds based upon the QED score [53]. These top and bottom 25 compounds were then analyzed for differences in their physical properties and human pharmacokinetic data. With the exception of fraction absorbed the QED algorithm was unable to separate the top and bottom 25 compounds in terms of pharmacokinetic parameters. It is seen from Fig. 4 that a simple bivariate plot of log *P* and PSA clearly demonstrates that the compounds described as 'ugly' are separated by PSA with all ugly compounds having PSA > 73 Å$^2$, which fits well with the cutoff proposed by Pfizer and all non-ugly

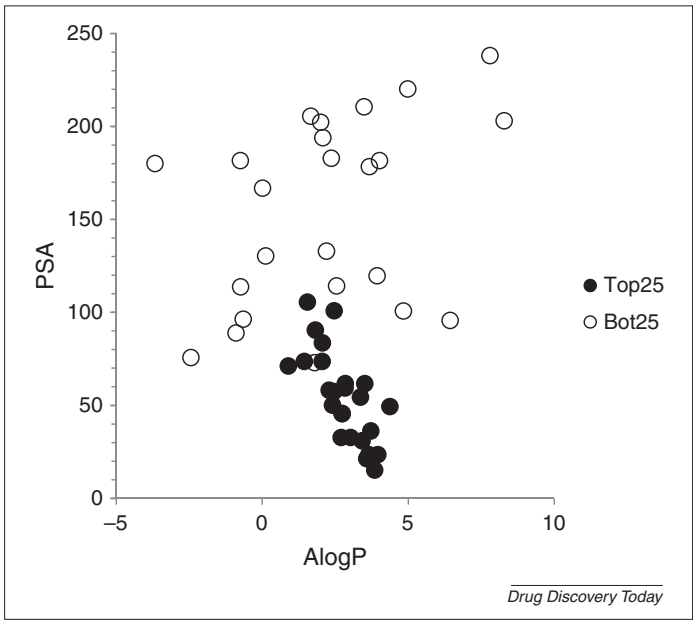


**FIGURE 4**

Plot of TPSA versus Atom based log *P* (Alog *P*) for the top 25 scoring (closed circles) and bottom 25 scoring (open circles) drugs based on quantitative estimate of drug likeness (QED) index.

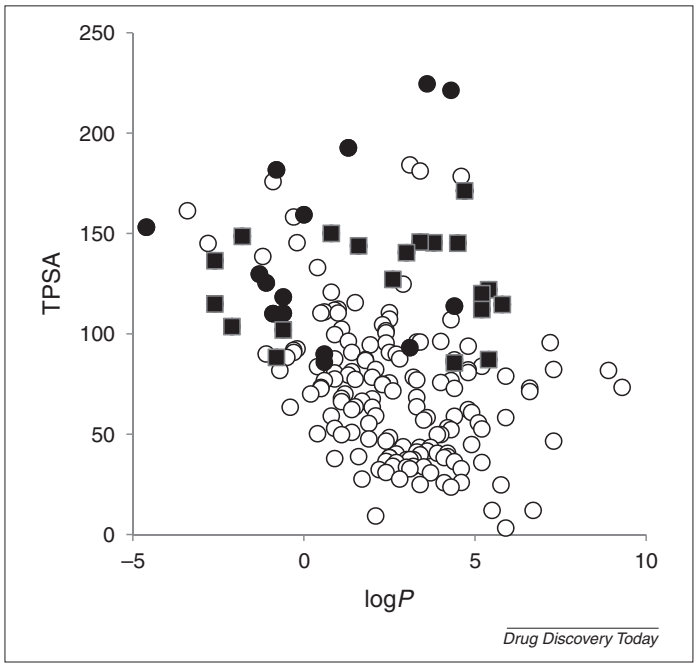


**FIGURE 5**

Plot of TPSA versus log *P* for top US 200 prescribed drugs. All compounds (open circles), injectable route of administration (closed circles) and sartans, statins and HIV protease inhibitors (closed squares).

drugs having PSA < 100 Å$^2$. Although it is not possible for us to calculate AEI for all of these compounds owing to the lack of readily available potency data, it is clear that PSA is a significant factor in this separation, justifying further its inclusion in AEI. This conclusion is also supported by an analysis of the top 200 US marketed drugs [54] (Fig. 5), which reveals that class III compounds are relatively limited in this group and dominated by certain therapeutic classes – antimicrobial (bacterial and viral), angiotensin II antagonists and statins, the disposition of which is known to be influenced by drug transporters [55]. To validate this approach further, the resultant ADMET score was applied to a range of datasets.

## HIV protease and HMG-CoA reductase inhibitors

These compound classes are often studied because of the interesting balance between enthalpic and entropic contributions to their efficacy [56]. Table 3 shows that, for the HIV protease inhibitors analyzed, darunavir was ranked the highest by LLE and AEI and represents the most recent HIV protease to reach the market and hence can be expected to have enhanced properties over historical antivirals. Amprenavir and atazanavir start to highlight and differentiate AEI from LLE and other indices. Amprenavir and atazanavir show good potency but possess greater PSA or lower affinity than darunavir and hence the indices start to diverge. The properties of atazanavir limit its absorption from acidic

**TABLE 3**

**Comparison of properties, ligand efficiency (LE), ligand-lipophilic efficiency (LLE) and lipophilicity-corrected ligand efficiency (LELP) together with the new ADMET efficiency index (AEI) for two well-studied compound classes**

| Compound | $pIC_{50}$ | log *P* | $N_h$ | PSA | LE | LLE | LELP | AEI |
|---|---|---|---|---|---|---|---|---|
| **HIV protease inhibitors** | | | | | | | | |
| Darunavir | 10.9 | 2.8 | 38 | 149 | 0.40 | 8.1 | 7.1 | 5.4 |
| Indinavir | 9 | 2.8 | 45 | 118 | 0.28 | 6.2 | 10.2 | 5.3 |
| Lopinavir | 11 | 4.7 | 46 | 120 | 0.33 | 6.3 | 14.3 | 5.3 |
| Amprenavir | 9.6 | 2.4 | 35 | 140 | 0.38 | 7.2 | 6.4 | 5.1 |
| Saquinavir | 9.5 | 3.2 | 49 | 167 | 0.27 | 6.3 | 11.9 | 3.8 |
| Tipranavir | 12.1 | 7.8 | 42 | 114 | 0.40 | 4.3 | 19.8 | 3.8 |
| Nelfinavir | 9.3 | 4.7 | 40 | 127 | 0.33 | 4.6 | 14.7 | 3.6 |
| Atazanavir | 10.4 | 4.5 | 51 | 171 | 0.29 | 5.9 | 16.2 | 3.5 |
| Ritonavir | 10 | 5.2 | 50 | 202 | 0.28 | 4.8 | 19.1 | 2.4 |
| **HMG-CoA reductase inhibitors** | | | | | | | | |
| Cerivastatin | 8.3 | 2.6 | 33 | 99.9 | 0.35 | 5.7 | 7.5 | 5.7 |
| Rosuvastatin | 9 | 1.9 | 33 | 149.3 | 0.38 | 7.1 | 5.2 | 4.8 |
| Pravastatin | 7 | 1.6 | 30 | 124.3 | 0.33 | 5.4 | 5.1 | 4.3 |
| Fluvastatin | 6.5 | 3.8 | 30 | 82.7 | 0.30 | 2.7 | 12.8 | 3.3 |
| Atorvastatin | 8 | 5.4 | 41 | 111.8 | 0.27 | 2.6 | 20.3 | 2.3 |

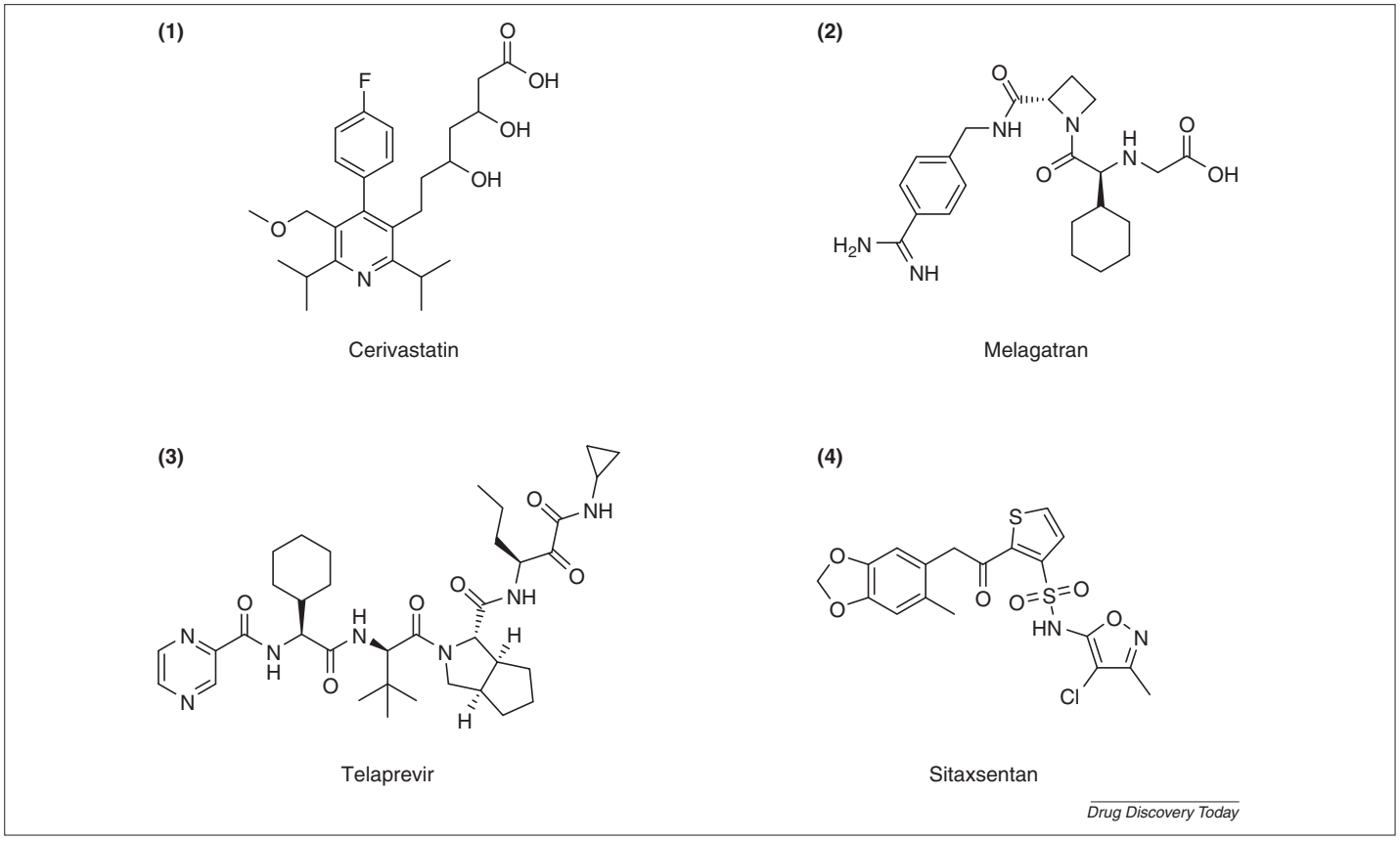


**FIGURE 6**

Compounds highlighted in this article. **1** Cerivastatin an HMG-Co reductase inhibitor withdrawn from the market because of reports of rhabdomyolysis. **2** Melagatran a thrombin inhibitor and the active form of ximelagatran withdrawn from market owing to liver injury. **3** Telaprevir a protease inhibitor for the treatment of hepatitis and discontinued as a result of increased competition. **4** Sitaxsentan is a endothelin-A antagonist used in the treatment of pulmonary arterial hypertension and withdrawn from market because of liver toxicity.

environments only, whereas amprenavir was subsequently replaced by the prodrug fosamprenavir which exhibits enhanced drug-like properties and was designed to reduce pill burden, and improved dissolution properties showed a clear improvement in bioavailbility relative to amprenavir. These small but significant changes in the balance of potency, lipophilicity and H-bonding potential start to highlight the benefit of AEI.

For HMG-CoA reductase inhibitors, a fairly good agreement was observed across all indices aside from a switch between cerivastatin (compound **1** in Fig. 6) and rosuvastatin. Although the magnitude of the AEI value for all HMG-CoA reductase inhibitors would indicate nonoptimal properties, this serves as a useful reminder that some transporter interactions can be beneficial and indeed targeted in drug design. In the case of HMG-CoA reductase inhibitors hepatic uptake is integral to their mechanism-of-action and safety profile. The hepatic uptake for rosuvastatin and atorvastatin is much higher than cerivastatin and this is reflected in their relatively lower AEI score. Of course this does not diminish the fact that these compounds would be flagged as having the propensity for transporter interactions or account for the fact that cerivastatin was withdrawn from the market for concerns around rhabdomylosis as a consequence of several well-documented drug–drug interactions (DDIs).

## Compounds optimized by applying LLE

A comparison was also made between LLE and AEI for a recently published dataset of 59 compounds (47 targets) where LLE was explicitly applied in their optimization. It is important to note that these compounds should theoretically have been optimized for oral delivery and contain no fatal flaws (e.g. metabolic instability) [2]. Although broad agreement between LLE and AEI was observed for this dataset (see Supplementary material online) [63], there were notable exceptions. For example lorcaserin scores very highly on AEI and moderately on LLE and can be rationalized in terms of its relatively low affinity (0.1 μM) and moderate lipophilicity ($c\log P = 3.2$) giving rise to moderate LLE. However, because the H-bonding potential (as reflected in PSA) is very low (PSA = 12 Å$^2$) this compound scores well in terms of AEI. Melagatran (compound **2**; Fig. 6) is an example of a compound scoring higher in LLE than in AEI and serves to highlight the difference between a ligand-based index and a more drug-like index. Melagatran is a potent (pActivity = 8.8) hydrophilic compound ($c\log P = -0.75$) with high H-bonding potential (TPSA = 148 Å$^2$). All of these properties would give rise to this compound scoring well in the LLE index. It is particularly interesting to note that melagatran has a negative $c\log P$ which inflates the LLE score [LLE = 8.8 − (−0.75)]. This is in contrast to AEI where the negative value of $c\log P$ has little influence but the high H-bonding potential reduces the score significantly because this can give rise to an increased probability of unfavorable affects. It is worth noting at this stage that ximelagatran, the prodrug of melagatran, has been associated with a number of unfavorable liver toxicity issues that culminated in its withdrawal from the market and further development activities in

**TABLE 4**

**Categorization of 57 recent ligand-lipophilic efficiency (LLE) optimized compounds**

| Class | *N* | LLE | | AEI | | Daily dose (mg) | | % class with daily dose (mg) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Mean | SE | Mean | SE | Mean | Median | <10 | <50 | ≤300 | >300 | >1000 |
| **A** | 15 | 6.62 | 0.29 | 9.91 | 1.58 | 71.37 | 25 | 33 | 80 | 93 | 7 | 0 |
| **B** | 17 | 5.87 | 0.45 | 5.42 | 0.18 | 181.32 | 100 | 18 | 29 | 76 | 24 | 0 |
| **C** | 16 | 3.02 | 0.33 | 2.78 | 0.22 | 926.56 | 600 | 0 | 0 | 44 | 56 | 31 |

Class A: ADMET efficiency index (AEI) > 7. Class B: 4 < AEI < 7. Class C: AEI < 4 compounds together with the maximal prescribed daily dose. Also shown is the percentage of compounds within a banded range of total daily dose.

2006. Interestingly, the lowest scoring drug with the new AEI is telaprevir (compound **3**; Fig. 6), an antiviral agent for the treatment of hepatitis C, which was recently withdrawn (for commercial reasons) by Vertex after only 2 years of marketing.

To broaden the utility of AEI, an attempt was made to relate it to the maximum daily dose of compounds using this same dataset. The maximum daily dose is a single numerical value that is a distillation of pharmacokinetics, pharmacodynamics and toxicity associated with any drug. High daily doses have been related to adverse events [57] and a low daily dose has been described as a way to mitigate deficiencies in potential therapeutic agents [58]. Hence targeting low daily dose represents a measure of drug efficiency and represents a multiparametric, integrated goal in drug discovery. From our analyses (Table 4) we propose three categories for the AEIs: Class A: AEI > 7, Class B: 4< AEI < 7, Class C: AEI < 4, the rationale for this being: clear distinction for best compounds; reference to total daily dose; approximately even distribution across this dataset. Figure 7 shows the relationship between AEI and maximum daily dose and demonstrates values of AEI > 7 corresponds to a low maximum daily dose. In this way the authors suggest that AEI values >7 identify compounds with superior intrinsic properties and minimal transporter interaction.

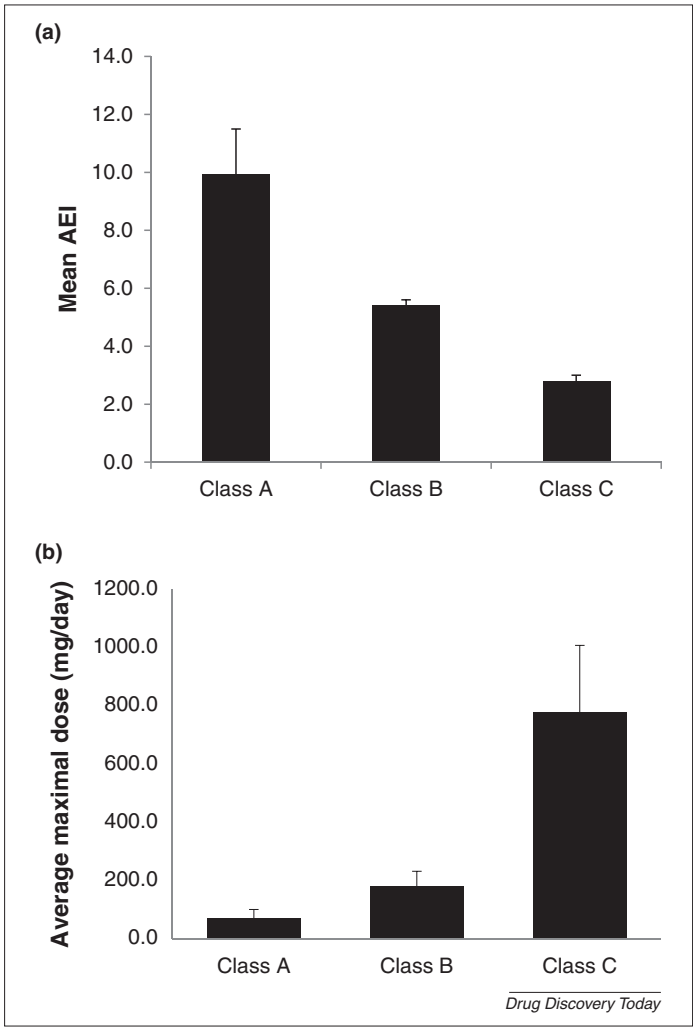


**FIGURE 7**

Catogerization of 59 ligand-lipophilicity efficiency (LLE) optimized compounds by **(a)** ADMET efficieny index (AEI) and **(b)** total daily dose. Class A: AEI > 7. Class B: 4 < AEI < 7. Class C: AEI < 4.

Compounds falling into class A have AEI > 7 with an AEI ranging from 7 to 31 and a mean of 9.9. For this class the median daily dose is 25 mg with 93% of this class having a total daily dose of <300 mg. Compounds within the range 4 < AEI < 7 (class B) tend to show higher maximum daily dose and potential suboptimal ADMET properties. The mean dose for this class is 100 mg with 24% having a total daily dose of >300 mg. Class C compounds where AEI < 4 are the most concerning representing a relatively poor LLE and ADMET profile. The median dose for these compounds is 600 mg and 44% have doses >300 mg daily.

This analysis also afforded a valuable comparison within a target class. This is exemplified in Table 5 for endothelin antagonists. As with statins, the AEI score for several of these compounds would trigger some further thinking. For example, as first-to-market in pulmonary arterial hypertension (PAH) bosentan still earns ~US$1 billion per year. However, whereas hepatic uptake is key to statin efficacy and benefits their safety profile overall, it is believed to be linked to hepatic toxicity of compounds such as bosentan via elevation of intrahepatic concentrations and inhibition of transporters such as BSEP, as illustrated in this dataset. By contrast, the AEI score for ambrisentan is superior as a result of improved potency and physicochemical properties and this culminates in a much reduced, once-daily dose. These properties also address the propensity to be actively taken up by the liver and inhibit BSEP.

The much improved AEI scores for sitaxsentan and macitentan are also consistent with a lower, once-daily dose compared with bosentan. However, sitaxsentan was withdrawn from the market suggesting these improvements were not sufficient to address hepatotoxicity, which could also have been linked to reactive metabolite formation, probably via the methylene dioxy motif. Interestingly, Table 5 suggests that macitentan has similar intrinsic OATP and BSEP liabilities to bosentan but these might be offset by the much lower dose and associated exposure in the liver.

**TABLE 5**

**Application of ADMET efficiency index (AEI) thinking to a series of endothelin antagonists indicating how potency and physical properties relate to AEI and ligand-lipophilic efficiency (LLE) together with the transporter affinity for organic anion transporting polypeptide (OATP)1B1 and bile salt excretory pump (BSEP)**

| | Bosentan | Sitaxsentan | Macitentan | Ambrisentan |
|---|---|---|---|---|
| **Daily dose (mg)** | 250 | 100 | 10 | 10 |
| **MW** | 552 | 455 | 588 | 378 |
| **$pIC_{50}$** | 7.9 | 8.9 | 8.9 | 8.6 |
| ***c* log *P*** | 4.2 | 2.8 | 3.4 | 3.8 |
| **PSA** | 146 | 144 | 128 | 82 |
| **LLE** | 3.8 | 6.1 | 5.5 | 4.8 |
| **AEI** | 2.6 | 4.2 | 4.3 | 5.9 |
| **OATP1B1 $IC_{50}$ (μM)** | 5 | ND | 2 | 47 |
| **Cell: media** | 20 | 40 | 100 | 2 |
| **BSEP $IC_{50}$ (μM)** | 42 | 25 | 12 | >100 |
| **'Effective' BSEP $IC_{50}$ (μM)** | 2.1 | 0.6 | 0.1 | >50 |

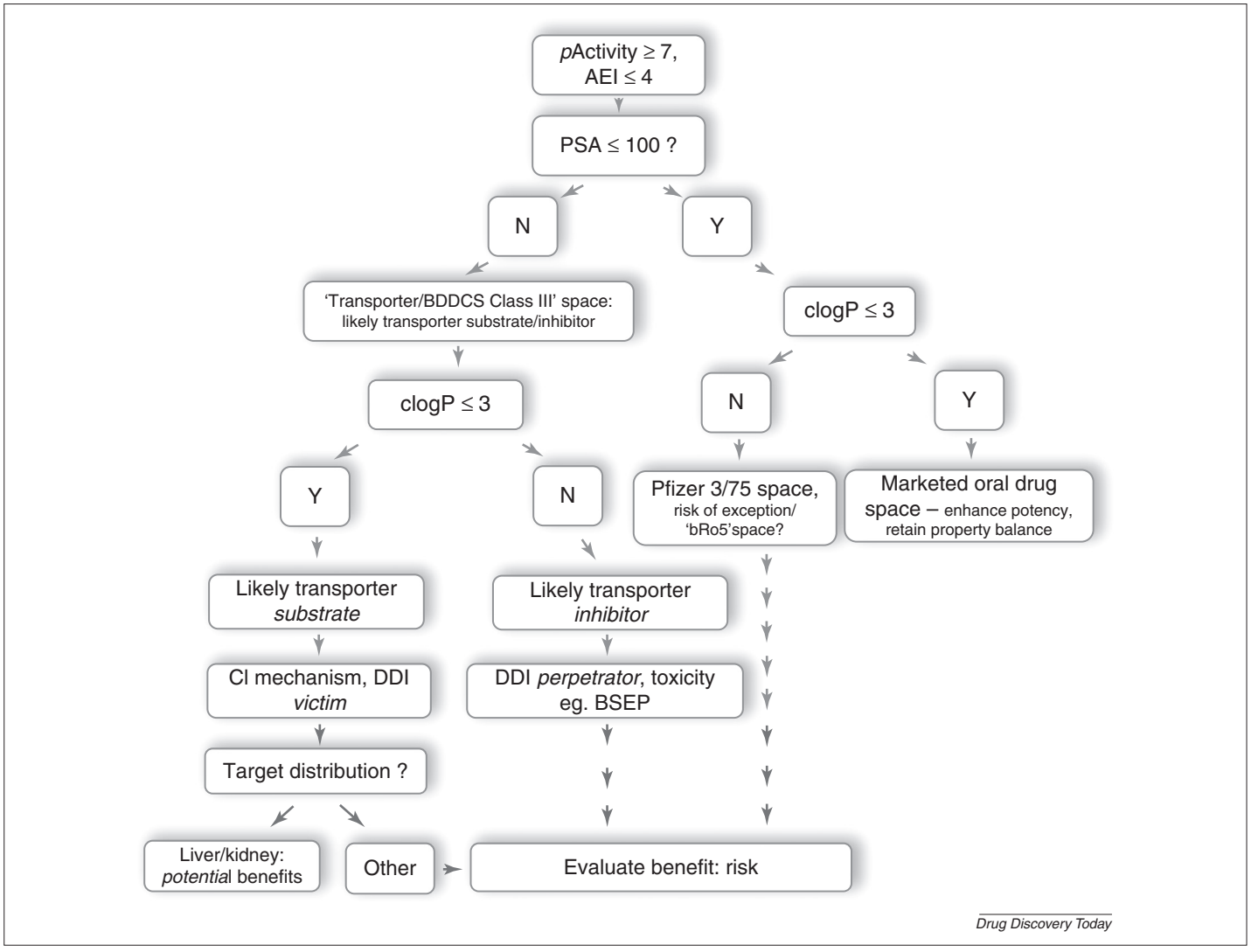


**FIGURE 8**

Decision tree proposed as a useful framework by which it is possible to gauge and anticipate potential attrition scenarios using ADMET efficiency index (AEI) in a drug discovery setting.

## Individual company strategies to compound optimization

Individual companies have clearly committed significant resource and effort to understand the underlying science influencing compound attrition. These company-specific approaches have identified different working practices within organizations and analysis reveals different companies tend to occupy different areas of chemical space [2,36]. Some organizations [38] have challenged the attempts to find simple guidelines that describe safety-related drug attrition and conclude the current models do not adequately describe their attrition data. They further conclude that if these models had been applied then a number of compounds would have failed to reach Phase II and hence would have been prevented from having potential clinical utility. The work of Leeson and colleagues [2] examined a number of physical properties of compounds acting at specific drug targets across a number of pharmaceutical companies. Careful examination of these data helps rationalize this juxtaposition between Pfizer attrition data and that of AstraZeneca: AstraZeneca manipulate *c* log *P* into a positive area of property space by increasing PSA whereas Pfizer appear to manage *c* log *P* without increasing polarity and presumably by maintaining a lower MW. If this is put in the context of AEI then it is plausible to argue that the attrition observed across different organizations differs because they work in different areas of property space.

Although, in general, organizations strive to maintain good overall lipophilicity, they differ significantly with respect to their HBA and HBD profiles. By implication, we argue that attrition can indeed relate to general physicochemical-property-related attrition but for some companies could be the result of the relatively high PSA values, which enhances the probability of the compounds falling into BDDCS class III and increased transporter-related attrition.

## Concluding remarks

Arguments have been presented for a new index that seeks to balance the physicochemical properties, which are known to give rise to nonspecific toxic events, with the increasing implication of drug–transporter interactions in drug toxicity. The AEI score looks to bridge the well-established lipophilic efficiency and known ADMET properties to maximize target engagement by ensuring suitable exposure at the site of action while minimizing drug–transporter involvement, which can often lead to DDIs or toxicity. By implication AEI presents an alternative paradigm to the traditional physical-property-related attrition which distinguishes between physical-property-related toxicity and transporter-related toxicity. Compounds for which PSA $> 75$ Å$^2$ and *c* log *P* $< 3$ demonstrate an increased propensity for drug–transporter interactions, typified by BDDCS class III compounds. PSA, an often understated ADMET-related property, has been demonstrated to be a pharmacophoric component of drug–transporter interactions in addition to specific structural motifs.

A key feature of AEI is its potential link to clinical dose. In all scenarios it is possible to mitigate potential defects in compound properties by targeting a low clinical dose. For example, in the case of the endothelin antagonists ambrisentan and bosentan, both compounds have *c* log *P* $> 3$ and PSA $> 75$ Å$^2$ and drug–transporter liability but, because of its superior potency and pharmacokinetics and low dose, ambrisentan is better tolerated. To facilitate the use of AEI in discovery programs we have demonstrated the utility of a simple *in silico* model that utilized the fact that many BDDCS class III compounds interact with drug transporters. The model shows that compounds having *c* log *P* $< 3$ and PSA $> 75$ Å$^2$ have a 73% probability of being a BDDCS class III compound. This statistically significant model facilitates early identification of compounds falling into this class and hence suitable mitigation can be put in place if required.

Here, to facilitate the use of AEI further in drug discovery projects the authors have put forward a decision tree (Fig. 8) to facilitate decision making and identify areas that might require further investigation. The tree starts by looking at the intrinsic potency of the compounds that should ideally be less than 0.1 μM and asks if AEI $< 4$. If, in general, potency is less than 1 μM but PSA $< 100$ Å$^2$ and the compound is not too lipophilic our analysis would suggest that the series needs to be further optimized only in terms of potency while retaining the balance of properties already present in the compounds. Ideally this would be conducted as part of a multiparametric optimization program incorporating desired ranges of properties together with associated errors, which negates the requirements for hard cutoffs [27]. In this way, a more probabilistic approach would be adopted. If the properties are as described above but the compounds are becoming rather lipophilic, *c* log *P* $> 3$, then general physical properties need to be modified. The potential fate of such compounds would be typified by the Pfizer 3/75 rule. For compounds with AEI score $<4$ and PSA $> 100$ Å$^2$ consideration needs to be given to the compounds being BDDCS class III and hence potential transporter substrates or inhibitors. The project teams could then consider further investigation of the benefit:risk of occupying this area of property space. The properties of such compounds can be inconsistent with oral delivery from simple formulations and hence other approaches could be considered: parenteral administration, enhanced formulations or designing a prodrug to overcome absorption issues, depending on a number of considerations including the intended therapeutic use. For larger molecules, creative synthetic solutions might also be considered including intramolecular H-bonding and macrocyclization.

The aim of the AEI is not to restrict the properties of new compounds to regions considered as marketed drug space but rather to act as an indicator of the potential deficiencies or challenges in a compound series. It is recognized, for example, that natural products represent one source of drug products often with properties that are bRo5. As drug discovery continues to target innovative approaches to more effective medicines, such as the disruption of protein–protein interactions, AEI serves to highlight concerns that small-molecule drug discovery needs to address early on rather than leaving late-stage attrition to continue to take its toll on the industry.

## Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at http://dx.doi.org/10.1016/j.drudis.2015.09.010.

## References


1 Kola, I. and Landis, J. (2004) Can the pharmaceutical industry reduce attrition rates? *Nat. Rev. Drug Discov.* 3, 711–716
2 Leeson, P.D. and St-Gallay, S.A. (2011) The influence of the 'organizational factor' on compound quality in drug discovery. *Nat. Rev. Drug Discov.* 10, 749–765
3 Paul, S.M. *et al.* (2010) How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nat. Rev. Drug Discov.* 9, 203–214
4 Swinney, D.C. and Anthony, J. (2011) How were new medicines discovered? *Nat. Rev. Drug Discov.* 10, 507–519
5 Hann, M.M. and Keserü, G.M. (2012) Finding the sweet spot: the role of nature and nurture in medicinal chemistry. *Nat. Rev. Drug Discov.* 11, 355–365
6 Springer, C. and Sokolnicki, K.L. (2013) A fingerprint pair analysis of hERG inhibition data. *Chem. Cent. J.* 7, 167
7 Waring, M.J. and Johnstone, C. (2007) A quantitative assessment of hERG liability as a function of lipophilicity. *Bioorg. Med. Chem. Lett.* 17, 1759–1764
8 Hanumegowda, U.M. *et al.* (2010) Phospholipidosis as a function of basicity, lipophilicity, and volume of distribution of compounds. *Chem. Res. Toxicol.* 23, 749–755
9 Riley, R.J. *et al.* (2007) Time-dependent CYP inhibition. *Expert Opin. Drug Metab. Toxicol.* 3, 51–66
10 Srivastava, A. *et al.* (2014) Identification and mitigation of a reactive metabolite liability associated with aminoimidazoles. *Chem. Res. Toxicol.* 27, 1586–1597
11 Warner, D.J. *et al.* (2012) Mitigating the inhibition of human bile salt export pump by drugs: opportunities provided by physicochemical property modulation, *in silico* modeling, and structural modification. *Drug Metab. Dispos.* 40, 2332–2341
12 Harnisch, L. *et al.* (2013) Modeling and simulation as a tool to bridge efficacy and safety data in special populations. *CPT Pharmacometrics Syst. Pharmacol.* 2, e28
13 Morgan, P. *et al.* (2012) Can the flow of medicines be improved? Fundamental pharmacokinetic and pharmacological principles toward improving Phase II survival. *Drug Discov. Today* 17, 419–424
14 Grime, K.H. *et al.* (2013) Application of *in silico*, *in vitro* and preclinical pharmacokinetic data for the effective and efficient prediction of human pharmacokinetics. *Mol. Pharm.* 10, 1191–1206
15 Veber, D.F. *et al.* (2002) Molecular properties that influence the oral bioavailability of drug candidates. *J. Med. Chem.* 45, 2615–2623
16 Waring, M.J. (2009) Defining optimum lipophilicity and molecular weight ranges for drug candidates – molecular weight dependent lower log*D* limits based on permeability. *Bioorg. Med. Chem. Lett.* 19, 2844–2851
17 Lipinski, C.A. (2004) Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discov. Today Technol.* 1, 337–341
18 Leeson, P.D. and Davis, A.M. (2004) Time-related differences in the physical property profiles of oral drugs. *J. Med. Chem.* 47, 6338–6348
19 Wenlock, M.C. *et al.* (2003) A comparison of physiochemical property profiles of development and marketed oral drugs. *J. Med. Chem.* 46, 1250–1256
20 Lipinski, C.A. *et al.* (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* 23, 3–25
21 Petit, J. *et al.* (2012) Softening the Rule of Five – where to draw the line? *Bioorg. Med. Chem.* 20, 5343–5351
22 Doak, B.C. *et al.* (2014) Oral druggable space beyond the rule of 5: insights from drugs and clinical candidates. *Chem. Biol.* 21, 1115–1142
23 Gleeson, M.P. (2008) Generation of a set of simple, interpretable ADMET rules of thumb. *J. Med. Chem.* 51, 817–834
24 Jhoti, H. *et al.* (2013) The 'Rule of Three' for fragment-based drug discovery: where are we now? *Nat. Rev. Drug Discov.* 12, 644
25 Bickerton, G.R. *et al.* (2012) Quantifying the chemical beauty of drugs. *Nat. Chem.* 4, 90–98
26 Wenlock, M.C. and Carlsson, L.A. (2015) How experimental errors influence drug metabolism and pharmacokinetic QSAR/QSPR models. *J. Chem. Inf. Model.* 55, 125–134
27 Segall, M.D. (2012) Multi-parameter optimization: identifying high quality compounds with a balance of properties. *Curr. Pharm. Des.* 18, 1292–1310
28 Segall, M. *et al.* (2011) Applying medicinal chemistry transformations and multiparameter optimization to guide the search for high-quality leads and candidates. *J. Chem. Inf. Model.* 51, 2967–2976
29 Ryckmans, T. *et al.* (2009) Rapid assessment of a novel series of selective CB(2) agonists using parallel synthesis protocols: a lipophilic efficiency (LipE) analysis. *Bioorg. Med. Chem. Lett.* 19, 4406–4409
30 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
31 Holdgate, G.A. and Gill, A.L. (2011) Kinetic efficiency: the missing metric for enhancing compound quality? *Drug Discov. Today* 16, 910–913
32 Morphy, R. (2006) The influence of target family and functional activity on the physicochemical properties of pre-clinical compounds. *J. Med. Chem.* 49, 2969–2978
33 Beaumont, K. *et al.* (2005) Oral delivery of G protein-coupled receptor modulators: an explanation for the observed class difference. *Bioorg. Med. Chem. Lett.* 15, 3658–3664
34 Conn, P.J. *et al.* (2009) Allosteric modulators of GPCRs: a novel approach for the treatment of CNS disorders. *Nat. Rev. Drug Discov.* 8, 41–54
35 DeWire, S.M. *et al.* (2007) Beta-arrestins and cell signaling. *Annu. Rev. Physiol.* 69, 483–510
36 Hughes, J.D. *et al.* (2008) Physiochemical drug properties associated with *in vivo* toxicological outcomes. *Bioorg. Med. Chem. Lett.* 18, 4872–4875
37 Grabowski, T. *et al.* (2010) Correlations between no observed effect level and selected parameters of the chemical structure for veterinary drugs. *Toxicol. In Vitro* 24, 953–959
38 Muthas, D. *et al.* (2013) A critical assessment of modeling safety-related drug attrition. *MedChemComm* 4, 1058–1065
39 Kogej, T. *et al.* (2013) Big pharma screening collections: more of the same or unique libraries? The AstraZeneca–Bayer Pharma AG case. *Drug Discov. Today* 18, 1014–1024
40 Wager, T.T. *et al.* (2013) Improving the odds of success in drug discovery: choosing the best compounds for *in vivo* toxicology studies. *J. Med. Chem.* 56, 9771–9779
41 Wu, C.Y. and Benet, L.Z. (2005) Predicting drug disposition via application of BCS: transport/absorption/elimination interplay and development of a biopharmaceutics drug disposition classification system. *Pharm. Res.* 22, 11–23
42 Benet, L.Z. *et al.* (2011) BDDCS applied to over 900 drugs. *AAPS J.* 13, 519–547
43 Landis, J.R. and Koch, G.G. (1977) The measurement of observer agreement for categorical data. *Biometrics* 33, 159–174
44 Varma, M.V. *et al.* (2012) Physicochemical property space of hepatobiliary transport and computational models for predicting rat biliary excretion. *Drug Metab. Dispos.* 40, 1527–1537
45 Tu, M. *et al.* (2013) Medicinal chemistry design principles for liver targeting through OATP transporters. *Curr. Top. Med. Chem.* 13, 857–866
46 Shitara, Y. *et al.* (2013) Clinical significance of organic anion transporting polypeptides (OATPs) in drug disposition: their roles in hepatic clearance and intestinal absorption. *Biopharm. Drug Dispos.* 34, 45–78
47 Karlgren, M. *et al.* (2012) Classification of inhibitors of hepatic organic anion transporting polypeptides (OATPs): influence of protein expression on drug–drug interactions. *J. Med. Chem.* 55, 4740–4763
48 Murray, C.W. *et al.* (2014) Validity of ligand efficiency metrics. *ACS Med. Chem. Lett.* 5, 616–618
49 Abad-Zapatero, C. and Metz, J.T. (2005) Ligand efficiency indices as guideposts for drug discovery. *Drug Discov. Today* 10, 464–469
50 Clark, D.E. (1999) Rapid calculation of polar molecular surface area and its application to the prediction of transport phenomena. 2. Prediction of blood–brain barrier penetration. *J. Pharm. Sci.* 88, 815–821
51 Chang, C. *et al.* (2006) Pharmacophore-based discovery of ligands for drug transporters. *Adv. Drug Deliv. Rev.* 58, 1431–1450
52 Toulmin, A. *et al.* (2008) Toward prediction of alkane/water partition coefficients. *J. Med. Chem.* 51, 3720–3730
53 Ritchie, T.J. and Macdonald, S.J. (2014) How drug-like are 'ugly' drugs: do drug-likeness metrics predict ADME behaviour in humans? *Drug Discov. Today* 19, 489–495
54 Zhong, H.A. *et al.* (2013) Understanding the molecular properties and metabolism of top prescribed drugs. *Curr. Top. Med. Chem.* 13, 1290–1307
55 International Transporter Consortium (2010) Membrane transporters in drug development. *Nat. Rev. Drug Discov.* 9, 215–236
56 Meanwell, N.A. (2011) Improving drug candidates by design: a focus on physicochemical properties as a means of improving compound disposition and safety. *Chem. Res. Toxicol.* 24, 1420–1456
57 Chen, M. *et al.* (2013) High lipophilicity and high daily dose of oral medications are associated with significant risk for drug-induced liver injury. *Hepatology* 58, 388–396
58 Nakayama, S. *et al.* (2009) A zone classification system for risk assessment of idiosyncratic drug toxicity using daily dose and covalent binding. *Drug Metab. Dispos.* 37, 1970–1977
59 Kuntz, I.D. *et al.* (1999) The maximal affinity of ligands. *Proc. Natl. Acad. Sci. U. S. A.* 96, 9997–10002
60 Congreve, M. *et al.* (2003) A 'Rule of Three' for fragment-based lead discovery? *Drug Discov. Today* 8, 876–877
61 Keseru, G.M. and Makara, G.M. (2009) The influence of lead discovery strategies on the properties of drug candidates. *Nat. Rev. Drug Discov.* 8, 203–212
62 Braggio, S. *et al.* (2010) Drug efficiency: a new concept to guide lead optimization programs towards the selection of better clinical candidates. *Expert Opin. Drug Discov.* 5, 609–618
63 Hopkins, A.L. , Keserü, G.M. , Leeson, P.D. , Rees, D.C. and Reynolds, C.H. (2014) The role of ligand efficiency metrics in drug discovery. *Nat. Rev. Drug Discov.* 13, 105–121