Title Page

In Silico Predictions of Blood-Brain Barrier Penetration: Considerations to ‘Keep in Mind’

JAY T. GOODWIN and DAVID E. CLARK

ADMETRx, Inc., 4717 Campus Drive, Suite 600, Kalamazoo, Michigan, USA (J.T.G.)
and Argenta Discovery Ltd., 8/9 Spire Green Centre, Flex Meadow, Harlow, Essex,
CM19 5TR, UK (D.E.C.)
**Running Title Page**

**Running title:** Predicting Blood-Brain Barrier Penetration

Corresponding author:

Jay T. Goodwin

ADMETRx, Inc.

4717 Campus Drive, Suite 600

Kalamazoo, MI 49008

Phone: (269) 372-3272

Fax: (269) 585-6016

jtgoodwin@admetrx.com

**Document statistics:**

- Number of text pages: 23
- Number of tables: 1
- Number of figures: 1
- Number of references: 40
- Number of words in Abstract: 246

**Non-standard abbreviations:**

- ADMET, absorption, distribution, metabolism, excretion, toxicity
- ARD, automatic relevance determination
- AUC, area under the curve
- BB, blood(plasma)-brain partitioning
- BBB, blood-brain barrier
- BCSFB, brain-cerebrospinal fluid barrier
- CNS, central nervous system
- CNS^-, not pharmacologically active in CNS
- CNS^+, pharmacologically active in CNS
- PS, permeability-surface area
- (T)PSA, (topological) polar surface area
Abstract

Within drug discovery it is desirable to determine whether a compound will penetrate and distribute within the CNS with the requisite pharmacokinetic and pharmacodynamic (PK/PD) performance required for a CNS target, or if it will be excluded from the CNS wherein potential toxicities would mitigate its applicability. A variety of in vivo and in vitro methods for assessing CNS penetration have therefore been developed and applied to advancing drug candidates with the desired properties. In silico methods to predict CNS penetration from chemical structure have been developed to address virtual screening and prospective design. In silico predictive methods are impacted by the quality, quantity, sources, and generation of the measured data available for model development. Key considerations for predictions of CNS penetration include comparison of local (in chemistry-space) versus global (more structurally diverse) models, and where in the drug discovery process such models may be best deployed. Preference should also be given to in vitro and in vivo measurements of greater mechanistic clarity that better support development of structure-property relationships. Although there are numerous statistical methods which have been brought to bear on the prediction of CNS penetration, of greater concern is that such models are appropriate for the quality of measured data available, and that such models are statistically validated. In addition, assessment of prediction uncertainty and relevance of predictive models to structures of interest are critical. This article will address these key considerations for development and application of in silico methods in drug discovery.
Introduction

The main interfaces between the central nervous system (CNS) and the peripheral circulation are the blood-brain barrier (BBB) and the blood-cerebrospinal fluid barrier (BCSFB). The surface area of the former (about 20 m²) is several thousand times larger than the latter; thus the BBB represents the most important barrier between the CNS and the systemic circulation (Pardridge, 2002; Graff and Pollack, 2004). There are at least two aspects of the BBB that make it a formidable barrier to candidate CNS drugs. First, in terms of its morphology, the endothelial cells of which the BBB is comprised are connected by complex tight junctions, which severely limit any paracellular transport. Furthermore, a minimal level of pinocytosis and lack of significant membrane fenestrae affords an additional hindrance to the transport of hydrophilic molecules. Second, the BBB includes an array of metabolic enzyme systems and efflux transporters that constitutes a biochemical barrier to the majority of xenobiotics (de Boer, 2003). This combination of physical and biochemical barriers establishes the BBB endothelium as quite distinct from other endothelia, and it has been estimated to prevent the brain uptake of more than 98% of potential neurotherapeutics (Pardridge, 2002).

Determinants of BBB penetration

The generally accepted biophysical/physicochemical models of BBB permeability have as their primary determinants for passive transport the solute’s lipophilicity, hydrogen-bond desolvation potential, pKₐ/charge and molecular size (Abraham, 2004; Atkinson et al., 2002; Chikhale et al., 1994; Jezequel, 1992; Young et al., 1988; Levin, 1980). Such models must however also acknowledge the potential for active mechanisms, which are generally dependent upon some specificity of molecular recognition and solute concentration at the transporter. BBB permeability is therefore impacted by other factors that determine solute concentration at the brain capillary surface, including plasma protein binding, blood flow through and partitioning into capillary membranes, and distribution into brain parenchyma (Kalvass, 2002). These, and other pharmacokinetic parameters such as absorption, first-pass metabolism, distribution into other tissues, and elimination pathways are in turn complex functions of physicochemical, biochemical and physiological determinants, not all of which are
directly related to the chemical structure of the drug (Burton, 2002). The question then becomes if, and how well, these determinants are captured by current in silico methods and which determinants can be related directly to solute chemical structure. Qualitative, expert-based rules such as those proposed by Lipinski (Raub, 2004; Raub, et al., 2004) or simple molecular polar surface metrics are considered grossly reflective of the major determinants of passive cellular membrane permeability (Clark, 1999). These simplified rule-based approaches have utility because they approximate our understanding of the underlying mechanisms contributing to CNS penetration. However, they are unlikely to reflect accurately the complexity of the interactions and combination of these physicochemical and biochemical determinants.

Now that predictions of CNS penetration are playing a greater role in discovery decision-making, an examination of the primary endpoints that are measured and modelled is warranted.

**How is BBB Permeation Measured?**

Log blood/brain (BB) is a measure of the partitioning between blood (plasma) and brain tissue, quantified by the ratio of the solute concentrations in brain and plasma. Log BB is generally obtained by methods such as the Brain Uptake Index (Oldendorf, 1981), which measures first-pass extraction from a single intravenous injection to yield log BB. Log BB can also be determined from the bolus carotid intravenous method (Ohno, 1978) with AUC (Area Under the Curve) plasma quantitation and single-point brain concentration determination. In this case, a permeability-surface area coefficient is determined, and log BB can be calculated from log PS with certain assumptions regarding the endothelial surface area. It is an apparent value as the perfusate contains serum protein and other blood components to which the compound of interest may bind. Note that BBB permeability is implicitly captured in this measure, and that it does not distinguish free and plasma-bound solute (Pardridge, 2004). Nor does it address intra-cerebral distribution, differentiating between intracellular and extracellular concentrations. As a consequence, log BB is but a crude assessment of the likely concentrations at the targeted site of CNS activity, whether extracellular or cytosolic. Results from log BB measurements are also dependent upon experimental conditions, particularly dosing regimen (single bolus dose vs. multiple doses vs. continuous infusion) as well as sampling time from dose.
Ideally, brain concentrations are measured at the plasma t\textsubscript{max} for the solute. However, differences in brain and systemic clearance can lead to variations in the measured log BB for a given compound depending upon sampling time. It is appropriate to consider the various contributions to experimental uncertainty when employing such data in model development, particularly when attempting to infer improved performance of a given method or set of computational descriptors relative to others.

When considering CNS penetration, perhaps the most appropriate \textit{in vivo} measure is the log of the permeability-surface area coefficient (log PS) (Martin, 2004; Pardridge, 2004), which represents the \textit{permeability} of a given solute across the brain capillary endothelium (the anatomical representation of the BBB). This measure reflects the free (unbound) extracellular solute concentrations, and is most often performed following the perfusion method established by Takasato (Takasato et al., 1984). This method eliminates serum binding and provides a direct measure of trans-BBB apparent permeability. However, it is a resource-intensive measure, requiring microsurgical expertise, and therefore of relatively low throughput. Solubility of the solute of interest in the perfusate can also be a limiting factor. As with log BB, this measure does not address specific intra-cerebral solute tissue distribution, and represents the potential combination of passive and active transport mechanisms (deconvolution requires additional experiments, perhaps in combination with \textit{in vitro} methods (as addressed in Current Issues & Future Directions)). Considering the greater mechanistic clarity of log PS compared with log BB, the former property is likely to be more informative as a measure of CNS penetration for use in lead optimisation. Its value will be greatly enhanced in combination with measures of other determinants of CNS penetration including plasma protein binding and susceptibility to active transport (Raub et al., 2004; Mahar-Doan et al., 2002).

There have been numerous efforts recently to classify chemical structures as CNS\textsuperscript{+} or CNS\textsuperscript{-}, where these predictions are derived from available data for marketed compounds (for examples, see: Ajay et al., 1999). These data represent pharmacological activity, which is the \textit{activity} of the compound at the receptor/targeted site of action, and are indirect and implicit measures of BBB
permeability and intra-cerebral concentration and distribution. It should be understood that if a given compound does not show CNS pharmacological activity, the compound may yet access the CNS but without the targeted or modelled pharmacological potency – “the absence of evidence is not the evidence of absence”. Such predictions may be best suited for combinatorial library design, in which a general bias for chemistries more likely to manifest CNS potency is desired, to select from among a much larger virtual or existing compound library, and to reduce the burden on compound acquisition, library synthesis, or high-throughput screening resources. Another critical consideration is the utility of such classification methods (e.g., CNS+ vs. CNS−) in prioritization and selection of potential candidates. What is the error in such classification, particularly at the boundaries separating the classes? What is the associate risk of advancing poor candidates or excluding viable candidates as a result of such classification error? Practitioners should also be cognizant of the risks in assuming that marketed drugs and existing compounds for which such pharmacological data exist represent the breadth and diversity of all chemistries likely to demonstrate favourable CNS activities.

**Why Predict BBB Permeation?**

What are the decisions that are made in drug discovery with respect to CNS penetration? Certainly, in the simplest terms, it is desirable to determine whether a compound will penetrate and distribute within the CNS with the requisite pharmacokinetic and pharmacodynamic (PK/PD) performance required for a CNS target, or if it will be excluded from the CNS wherein potential toxicities would mitigate its applicability. As a result, a variety of *in vivo* and *in vitro* methods for assessing CNS penetration have been developed and applied to advancing drug candidates with the desired properties. However, such methods are inevitably resource-intensive in terms of skilled personnel, animals, cell culturing, bioanalytical support, and of course they are retrospective in their application, requiring the existence of synthesized compound. *In silico* prediction methods address this limitation, supporting the prospective design and selection of candidate structures prior to synthesis.
A Bit of History

The prediction of BBB permeation from computed or experimental parameters has been of interest for at least thirty years. Some of the earliest attempts sought to discover relationships between brain penetration (usually determined as log BB) and partition coefficients measured in various solvent systems. For instance, it was suggested in the early 1970s that the optimum log P oct value for diffusion into the CNS was around 2 (Hansch and Glave, 1972; Levin, 1980). In something of a seminal paper, Young and co-workers at Smith Kline & French discovered a linear relationship between log BB and ∆ log P (= log P oct – log P cyc) for 20 histamine H₂ receptor antagonists (Young et al., 1988). This seems to have been one of the first indications of the crucial role of hydrogen-bonding capacity, as well as lipophilicity¹, in passive diffusion across the BBB. The 20 compounds in the Young paper are still part of most log BB training sets today, having been added to over the years, notably by Abraham and co-workers (Platts et al., 2001). Since the mid-nineties, when interest in predictive ADMET began to burgeon, many groups have developed predictive models using log BB data or surrogate measures such as CNS (in)activity. These have been thoroughly reviewed at regular intervals in recent years (Norinder and Haeberlein, 2002; Clark, 2003; Ecker and Noe, 2004) and so will not be covered here.

The State of the Art

The state of the art in log BB prediction can be illustrated by a recent paper (Winkler and Burden, 2004). A Bayesian neural network was used to model the relationship between an 85-compound log BB data set and a set of computed molecular property descriptors including counts of hydrogen-bond acceptors and donors, hydrophobes, rotatable bonds, log P, molecular weight and polar surface area (PSA), a popular measure of hydrogen-bonding capacity. The model with the best statistics (r² = 0.81, s = 0.37) was obtained using four nodes in the hidden layer. Applying this model to a 21-compound test set gave rise to a q² value of 0.65 (standard error of prediction = 0.54). The statistics for the training and test sets are typical of what is observed in high-quality log BB models (Clark, 2003). Given that the experimental error in log
BB data is estimated to be up to 0.3 log units, it is clear that the current generation of in silico models appears to be approaching the limits of predictive accuracy that might be expected. Applying automatic relevance determination (ARD) to the whole data set showed that the most important descriptor in the model was log P, closely followed by the count of rotatable bonds and PSA. The combination of lipophilicity (log P) and hydrogen-bonding (PSA) descriptors is a feature of many modern log BB models (Clark, 2003). The count of hydrogen-bond donors was more important than that of acceptors, mirroring findings by other workers in the field of absorption prediction (Clark, 2005).

It would seem then that, given the current data sets available, the prediction of log BB is about as good as can be expected. There is, however, little room for complacency, especially given some recent high-profile criticism levelled at log BB as an index of brain penetration (Martin, 2004; Pardridge, 2004). In particular, the fact that log BB reflects the total drug concentration in the brain, rather than the free drug concentration, which crucially determines the receptor saturation by the drug, has been highlighted (Atkinson et al., 2002). It is Pardridge’s contention that the best index of BBB permeability is the BBB permeability-surface area (PS) product, because this can predict the level of free drug in the brain (Pardridge, 2004).

It is thus timely that two papers on log PS prediction have been published recently. In the first of these, Abraham has collated a set of log PS values for 30 neutral compounds and correlated these data ($r^2 = 0.87$, $s = 0.52$) with his five solute descriptors previously used for log BB modelling (Abraham, 2004). It is interesting to compare two equations for log BB and log PS:

\[
\log \text{BB} = 0.021 + 0.463E - 0.864S - 0.564A - 0.731B + 0.933V - 0.567I \\
N = 148, r^2 = 0.75, s = 0.34, F = 69
\]

(Equation 6 in Platts et al., 2001)

\[
\log \text{PS} = -0.639 + 0.312E - 1.009S - 1.895A - 1.636B + 1.709V \\
N = 30, r^2 = 0.87, s = 0.52, F = 32.2
\]
where $E$ is an excess molar refraction, $S$ is dipolarity/polarisability, $A$ is the hydrogen-bond acidity of the compound, $B$ is the hydrogen-bond basicity of the compound, and $V$ is its characteristic (McGowan) volume. With the exception of the additional indicator variable ($I$) for carboxylic acids, the two equations are reassuringly similar in terms of the descriptors that they comprise and the signs of the coefficients. This indicates that the factors that determine log BB are also those that determine log PS, suggesting that the knowledge about the molecular determinants of brain permeation gained over recent years from log BB modelling is still useful, at least in a qualitative sense.

The second paper (Liu et al., 2004) describes the measurement of log PS values for 28 drug-like compounds. Of these, 23 compounds were assumed to cross the BBB by passive diffusion and were selected for modelling. Using stepwise multivariate linear regression and a variety of computed descriptors, the following equation was derived:

$$\text{logPS} = -2.19 + 0.262 \times \log D + 0.0683 \text{ vsa\_base} - 0.009 \text{ TPSA}$$

$n = 23, r^2 = 0.74, s = 0.5, F=18.2$

(Equation 3 in Liu et al., 2004)

where log $D$ (at pH 7.4) was calculated by ACD/LogD software (Advanced Chemistry Development, Inc., Toronto, ON, Canada), vsa\_base is the van der Waals’ surface area due to basic atoms and TPSA is the topological polar surface area – the latter descriptors being computed by MOE (Chemical Computing Group, Montreal, QB, Canada). Once again, the combination of polar surface area and lipophilicity descriptors previously observed in many log BB models is noteworthy, and reassuring. This model was found to predict two small literature log PS data sets quite well.

**Some Success Stories**
Although it is hard to find instances in the medicinal chemistry literature in which BBB permeability calculations have been used successfully, some examples are beginning to emerge in which it is clear that, at least in a qualitative sense, predictive modelling has had influence, either directly or indirectly. For example, scientists at AstraZeneca, citing a log BB modelling paper (Clark, 1999), have described their efforts to produce a non-CNS penetrating kappa opioid receptor agonist by adding polar substituents to increase the PSA and reduce lipophilicity (Semple, 2003). Conversely, Neurogen’s scientists have reported the results of lipophilicity and PSA calculations as evidence that a series of corticotrophin releasing factor-1 receptor antagonists was likely to permeate the BBB (Hodgetts et al., 2003), again citing a review of predictive modelling (Atkinson, 2002) as their justification. Finally, the Volsurf BBB permeation model (Cruciani, 2000) has been employed recently to predict the BBB permeability of some prodrugs of non-steroidal anti-inflammatory agents (Perioli et al., 2004). It seems likely that, given the lag time inherent in most medicinal chemistry publications, more such applications will be reported in the future reflecting the uptake of in silico BBB permeation predictions in the last few years.

Current Issues and Future Directions

The increased ease, speed, and throughput of in silico methods do not come without a cost, viz., the increased risk that such predictions may be somewhat or totally inaccurate for the structures of interest. Such inaccuracies would derive from two primary sources: the nature of the chemistry-space represented by the compounds used in training these models (does it cover, and how well, the chemistry of interest?), and the mechanistic clarity and relevance represented in the data as they are measured.

A common refrain at the present time in the field of predictive ADMET is the need for better and larger data sets from which to build the next generation of models. The area of BBB permeation is no exception to this. Even after more than a decade of intense interest in predicting BBB permeation, the number of log BB measurements available in the public domain is still probably fewer than 200. If, as has been
suggested above, the focus should now shift to log PS, the need for additional data is even more acute, since the available log PS data number only a few tens. The composition, as well as size, of data sets also needs attention. To date, the publicly available data sets comprise a motley assortment of compounds from various sources with little guarantee of their being derived by a common experimental protocol. If the science of BBB permeation prediction is to advance as we would wish, then what is required is bespoke data generation for the purposes of model building, creating data sets that span appropriate ranges in the biological endpoint, molecular diversity and physicochemical properties.

Again, considering the possible structural diversity of chemistry-space (even subsetting for ‘small’ molecular therapeutics of MW $\leq 500$), and the potential liabilities associated with ‘global’ models – sparseness of coverage in chemical-space regions of interest, and lack of motivation to generate such a large and comprehensive set of measurements due to the substantial financial burden for any one organization, it may be more appropriate to focus resources for quantitative model development within lead optimization series. In such settings there will be sufficient amounts of compound generated for measurements (less required for in vitro measures, more required for in vivo confirmatory experiments), and greater structural similarity to support local model development. The importance of local chemistry-space-relevant model development should not be underestimated – through working in more homologous chemistry-spaces, there is a greater likelihood of reducing complexities due to multiple mechanisms and the resulting need to generate nonlinear statistical/computational models using methods such as neural networks. Such nonlinear models in turn make structure-based mechanistic interpretation of predicted trends from computational descriptors difficult at best and are thus best avoided during lead optimization when the key question asked by the medicinal chemist is “Which compound should I make next?”

In vivo methods, while offering a more direct measure of performance and greater acceptance in discovery decision-making as a result, often cannot meet the time constraints placed on such decision-making due to their low throughput. In addition,
in vivo methods are not as readily amenable to the mechanistic deconvolution of the various determinants of CNS penetration, including passive and active transport. Therefore in vitro methods that can supplant more resource-intensive in vivo methods – such as the use of wild-type and transformed/transfected cell lines (such as MDCK transfected with human MDR1, or other transformed cell lines (Terasaki et al, 2003)) - have emerged and have found utility in combination with in vivo methods for assessing CNS penetration with greater throughput and increased mechanistic clarity and relevance (Raub et al., 2004). Such mechanistic clarity, along with more rapid turnaround times for the generation of data, are critical for development and application of predictive models of CNS penetration to be applied in discovery decision-making.

No method is without its drawbacks, and there are several concerns associated with predicting CNS penetration. These can perhaps best be enumerated by indicating where in the discovery process such methods are currently applied. The earliest stages in drug discovery where knowledge of likely CNS penetration will assist in decision-making are in library design and virtual screening, as well as in the hit-to-lead/lead selection stages (Figure 1). In silico methods offer the promise of rapid, relatively inexpensive (beyond the initial investment in software models and infrastructure) characterization, and can be used with the objective of populating libraries either exclusively or a with preponderance of structures anticipated to be either CNS penetrants or non-penetrants (for library design) or biasing lead selection towards desired CNS penetration (or lack thereof). Existing models such as those of Cruciani et al. (Cruciani, 2000) or simpler metric/rule-based approaches (Raub, 2004; Norinder, 2002: Clark, 1999 & 2003) (simple indicator variables, surface area calculations) are readily applied in library design and virtual screening. However, these models are in large part based upon phenomenological data (e.g., pharmacologically-active CNS+ vs. non-active, CNS- compounds) or experimentally measured sets of pharmacokinetic property endpoints of limited size and structural diversity. The compounds that comprise these datasets are subsets of existing, synthesized compounds, which are in turn vastly small (and potentially sparsely sampled) subsets of possible chemistry-space. It is appropriate to consider the relevance of the training data structural composition relative to the chemical classes of
interest. Such models may be globally applied but perhaps not locally relevant. Therefore, there is an inherent risk in application of such models – that credible compounds will be excluded from further consideration before there is an opportunity to validate them through in vivo studies (false negatives). As such, these models or mnemonic devices should mirror current biophysical and biochemical understanding of pharmacokinetic processes involved in CNS penetration (e.g., plasma binding, passive and active transport mechanisms, metabolism) to imbue some sense of confidence in their applicability. The simpler expert-rule/metric-based approaches as referred to above may therefore be most appropriate at this stage. Their application may be best suited in combination with some structural similarity clustering to ascertain variability of the CNS metrics across structural classes. Such a combination of methods will help to diagnose liabilities intrinsic to the core template, or due to side-chain modifications.

Arguably the more appropriate setting for determination of CNS penetration is in the lead optimization stage in drug discovery. At this point in the discovery process, there is generally a sufficient investment of synthetic capacity and medicinal chemistry support for compound design and scale-up to support experimental measurements. Furthermore, the data resulting from in vitro and in vivo assessments of CNS penetration may be used for the development of more locally (in the sense of chemistry-space) relevant and quantitatively predictive models. In addition, there usually exist sufficient resources for confirmation of predicted CNS penetration, and in turn the application of such data to iterative model development for improved prospective design of additional analogs.

Regardless of the stage in the drug discovery process, it is the responsibility of the user to determine the appropriateness of any predictive model that they may choose to apply (Table 1), whether made available from the scientific literature or developed within their own organizations. Model ‘appropriateness’ will be dictated by the quality (mechanistic clarity and relevance to the process to be modelled, statistical robustness provided by replicate measurements, range and distribution of measured data) (Stouch, 2003), quantity, and whether the chemical structures in the training set
are representative of the chemical-space of interest (does the model extrapolate or interpolate predictions for the structures of interest?). What is the statistical confidence surrounding the predicted value, and how good does it need to be for the decisions to be made based upon such predictions? There have been numerous statistical methods developed and applied to modelling of CNS penetration, and aside from concerns as to the appropriateness of qualitative predictions (uncertainty increases as predictions approach category boundaries), utilization of one or more of these methods is basically a matter of user preference.

**Concluding Remarks**

The utility of methods for predicting CNS penetration is fundamentally dependent upon what is to be decided, and at what point in the drug discovery process. This will impact the type of data to be generated (*in silico, in vitro, or in vivo*) and will determine the requirements for mechanistic relevance and clarity of the measured endpoint and how they can be related to chemical structure. The discovery scientist will need to determine how much uncertainty associated with data (predicted or measured) can be accommodated in the decision process (risk assessment). They should consider the advantages of employing local as opposed to global models, and integration with *in vitro* and *in vivo* measurements to validate and iteratively refine models and extend their utility. These issues will continue to drive measurement of data of appropriate statistical and mechanistic clarity for development and improvement of *in silico* models. This will in turn lead to reduction of uncertainty and will improve efficiencies through focus on local models – into lead refinement where greater resources are made available, and where the opportunities exist for ongoing validation of *in silico* and *in vitro* methods against *in vivo* measurements.
References


Footnotes

[1] Note that lipophilicity, as measured by octanol/water partitioning, is in fact a function of both hydrophobicity (cavity formation in solvent) and hydrogen-bond donor potential, as indicated by solvatochromic analysis (El Tayar, et al., 1991). Hydrogen-bond potential, as represented by ΔlogP (the difference in octanol/water and heptane (or isooctane)/water partitioning) is primarily a function of both hydrogen-bond acceptor and donor energetics, with minimal contribution from solute volume (or cavity-formation).
Figure Legends

Figure 1. Application of *in silico*, *in vitro*, and *in vivo* CNS penetration assessment methods to the various stages of drug discovery.
Tables

Table 1

- Understand the mechanistic nature of the property to be modeled, and how it is experimentally measured; mechanistically-convoluted and statistically poor data will not lead to informative models;
- Use methods to assess BBB permeability/CNS penetration that are mechanistically more straightforward \(\rightarrow\) Log PS is more informative than Log BB; \textit{in vitro} passive and active transport experiments augment \textit{in vivo} measurements;
- \textit{In silico} models can be used to suggest compounds for more resource-intensive \textit{in vitro} and \textit{in vivo} measurements, and resulting data can be used to iteratively refine and improve predictive quality of \textit{in silico} models;
- Examine the range and distribution of measured data available. Available data should cover the range of values relevant to the given property and that required for informed decision-making. Data should not be underrepresented in the regions of importance for decision-making;
- Focus \textit{in vitro} and \textit{in vivo} measurements on the chemistry-space of interest – best applied in lead optimization;
- Employ structure-based computational descriptors that appeal to biophysical and medicinal chemical intuition. Clarity of interpretation of the model results may be of more use than moderate improvements in model accuracy for prospective chemical design;
- Ensure that enough measured data are available for creating training and validation data subsets, and that the range and distribution of measured data and computed descriptors are comparable in both subsets. If not enough data are available, use recommended cross-validation methods such as boosting or leave-one-out approaches;
- Do not use statistical and computational methods as “black boxes” – their strengths and limitations must be understood;
- Assess the appropriateness of any model to predict a value for a structure of interest; does the structure fall within the ‘model-space’ – is the prediction interpolative or extrapolative?
- Determine the confidence or error associated with a given predicted value

Table 1. Essential considerations for development and application of effective and validated predictive models. For several representative references, see: Stouch, et al., (2003); Golbraikh & Trophsa (2002); Hawkins (2004); also see \texttt{http://home.neo.rr.com/catbar/chemo/pitfalls.htm}. 
Figure 1

Assessing CNS Penetration

Rules-based filters

In Silico

Structure-based quantitative property relationships

In Vitro

In Vivo

Drug Discovery Stages

Library design & acquisition

HTS

Hit triage/hit to lead

Lead optimization

Preclinical development

Clinical stages

Marketed drug