We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

BOOK CITATION INDEX
CLARIVATE ANALYTICS
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Use of Analytical Methods and *In-silico* Techniques in Public Health Research

Peter Papoh Ndibewu and Prince Ngobeni

Additional information is available at the end of the chapter

1. Introduction

In this chapter, we briefly summarize the concept of analytical methodologies used for detecting, measuring, and/or monitoring in public health research. Additionally, this chapter describes *in silico* ADME, ADMET and ADME/Tox approaches or models relevant to analytical methodologies for public health researchers or practitioners.

Recently, adopted technologies to cope with this type of scientific demand in terms of drug development and testing are the applications of *in silico* techniques used in pharmaceutical companies in the process of drug discovery (Ndibewu et al., 2012; Vallero, 2012; Lipinski et al., 1997; Leeson & Springthorpe, 2007). Two fast growing examples involve the use of cheminformatics (also known as chemoinformtaics) and chemical informatics which is the use of computer and informational techniques, applied to a range of problems in the field of chemistry (Langdon et al., 2011; Brown, 2011). These *in silico* techniques are employed in pharmaceutical companies in the process of drug discovery (Lipinski, 2004). These methods can also be used in chemical and allied industries in various other forms (Pradeep, 2009). Combined with the accuracy of data obtained with validated analytical methods, this encompasses the mixing of those information resources necessary to transform data into information and information into knowledge for the intended purpose of making better decisions faster in the area of drug lead identification and optimization. The outcome is clearly the efficiency in public health research and the beneficiary is mankind. In the section that follows, we briefly summarize *in silico* pharmacokinetics, widely designated as ADME approaches or models relevant to analytical methodologies for public health researchers or practitioners (Yamashita & Hashida, 2004). Pharmacokinetics is the study of the time course of a drug within the body and incorporates the processes of absorption, distribution, metabolism and

excretion (ADME) (Van de Waterbeemd & Gifford, 2003). The extent of distribution will depend on the structural and physicochemical properties of the compound (Fig. 1).
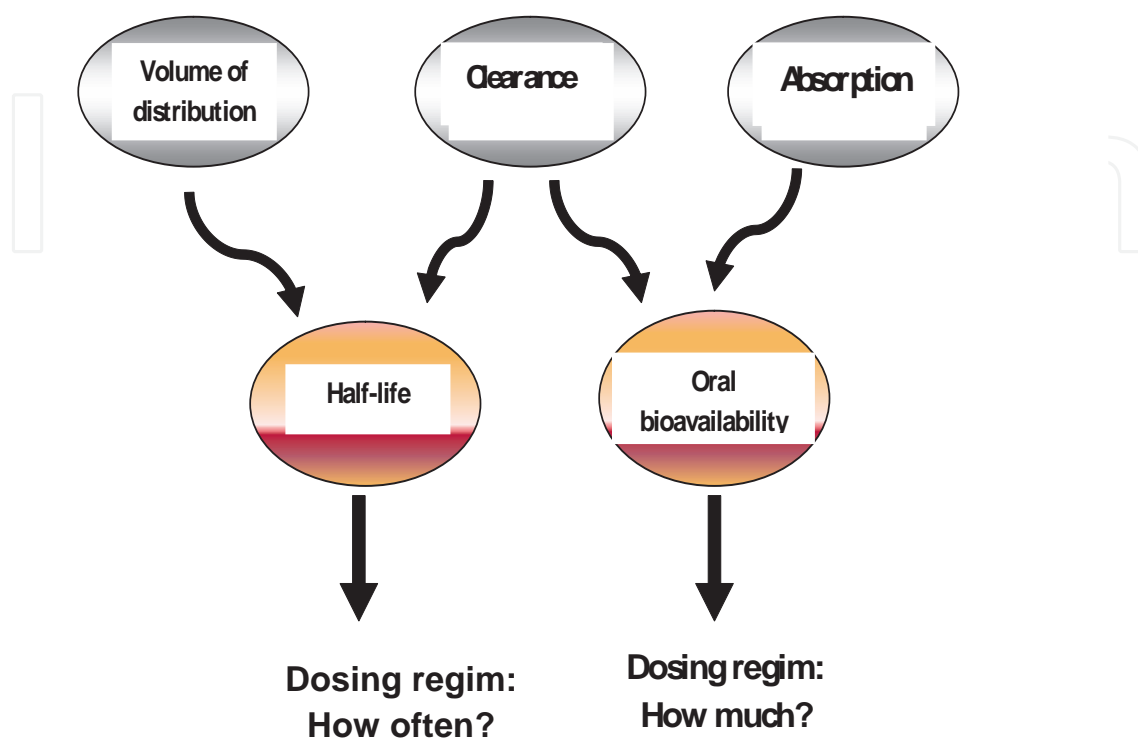

Figure 1. ADME flow chart showing volume of distribution (V_d): is a theoretical concept connecting administered dose with actual initial concentration (C_o) present in circulation, $V_d = Dose/C_o$; Unbound volume, $V_{du} = Vd/f_u$, where f_u is the fraction of unbound; Clearance (Cl) of the drug from the body mainly takes place via the liver and the kidney; Bioavailability is given by area under the curve (AUC) = F x Dose/Cl; Half-life ($t_{1/2}$) is the time taken for a drug concentration in plasma to reduce by 50% (is function of the clearance and volume of distribution, and determines how often a drug needs to be administered ($t_{1/2} = 0.693V_d/Cl$) (Van de Waterbeemd & Gifford, 2003).

Figure 2 shows a simple ADME decision-making flow sheet which will incorporate predictors for volume of distribution, oral bioavailability, half-life ($t_{1/2}$), distribution\protein binding module including percentage plasma protein binding values (*%PPB*) and drug affinity constant to human serum albumin represented by log K^a_{HSA} constants. The adaptability of this simple form has continuously refine existing models by building on larger and higher-quality data sets crucial to the success of the *in silico* approaches as grouped by Van de Waterbeem and Gifford (2003) in the form of problem areas for which predictive models could be helpful (Fig. 2). And during these past few years, the range of models have further expanded to include, for example, models for various transporters, metabolism by non-P450 enzymes, plasma protein binding, and so (Lu et al., 2003; Ekins et al., 2001).

Figures 2a and 2b outline the parameters in the prediction of a safe drug given in acceptable dose, which it is ultimately hoped will be reliably obtainable from molecular structure and appropriate descriptors using a suite of predictive models. This expression had earlier been made clear by Japertas and coworkers (2011).
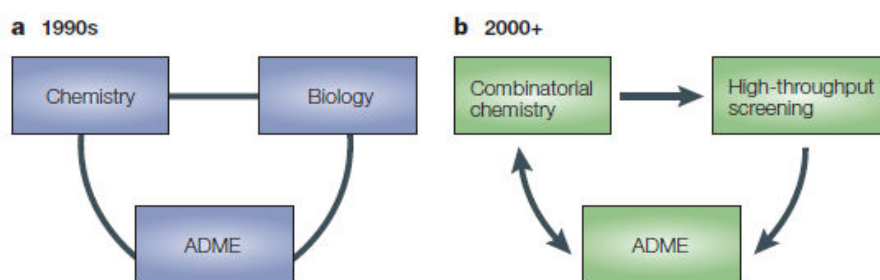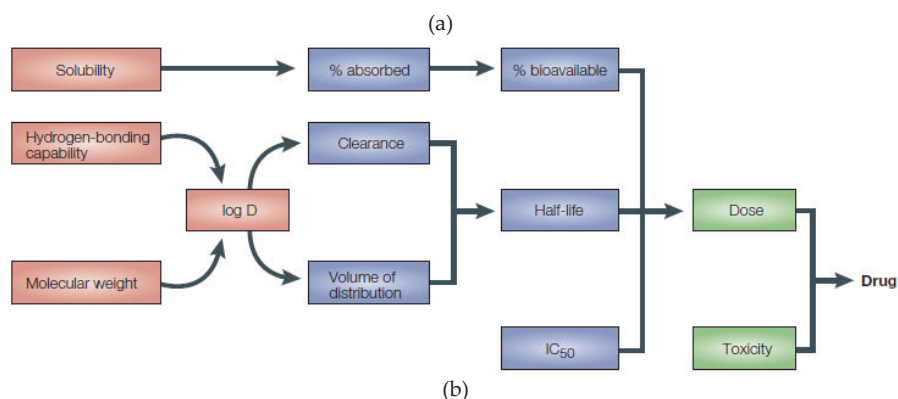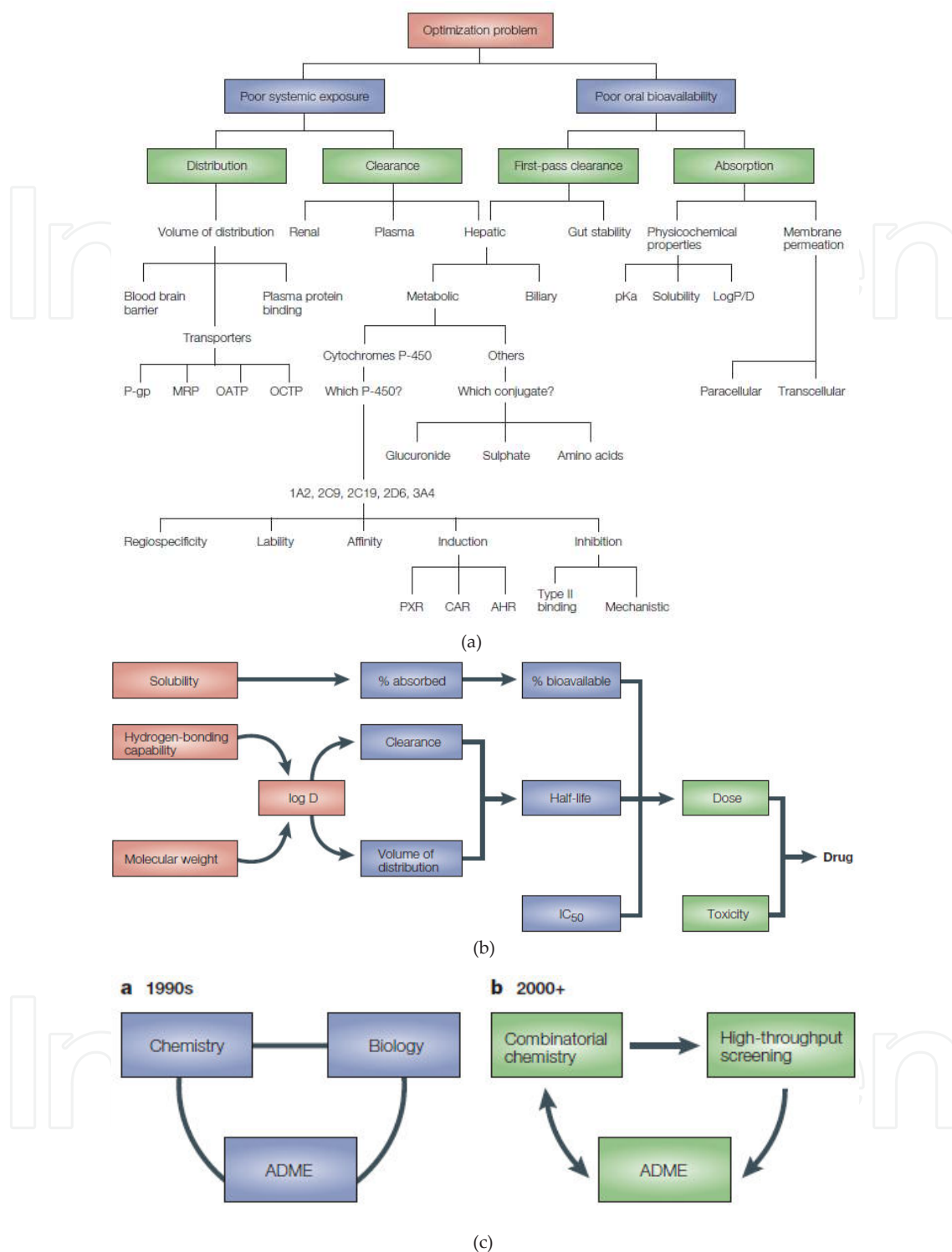
Figure 2. a. An analysis of the crucial ADME processes for which predictive models are available or are being developed (Carlson & Segall, 2002). b. Prediction paradigm ranging from models for simple physicochemical properties (H-bonding capacity, Molecular mass, Solubility and lipophilicity (log*D*)) to models for ADME properties (% drug absorbed & bioavailability, clearance, volume of distribution and half-life ($t_{1/2}$)) to complex endpoints (binding (IC_{50}), to molecular target of new drug and its required dose and toxicity potential (Waterbeemd & Gifford, 2003). c. The evolution of drug discovery and the changing role of ADME studies I (Waterbeemd & Gifford, 2003)

In figure 2c (Waterbeemd & Gifford, 2003), the form *a* depicting the classical project-collaboration approach between chemistry, biology and drug metabolism (ADMES) groups in the 1990s is shown to *b* which is a much more automated world at the start of this millennium in which combinatorial chemistry (*CombiChem*), high throughput screening and ADME studies are linked together in a streamlined fashion.

It should be noted that all three activities shown in figure 3 can even be carried out by separate companies of research units or even researchers in health departments or for health interest. A good example is the Medical Research Council (MRC) of South Africa which supports unit health research projects towards a nationally planned and prioritized health sectors. Furthermore, the wide introduction of *in silico* and high-speed in vivo methods could redefine the traditional meaning of ADME to Automated Decision-Making Engine. To reflect, one sees that considerable effort has equally been devoted into the development of *in silico* models for the prediction of oral absorption (Veber et al., 2002; Agoram et al., 2001; Parrott & Lavé, 2002; Yu et al., 2000). This range from the simplest models based on a single descriptor, such as log P or log *D*, or polar surface area, which is a descriptor of hydrogen-bonding potential to combo or meta- or QSAR models.

2. Analytical methods in health research

As analytical methods become increasingly sophisticated and capable for the detection of each component in a sample (Barnes & Dourson, 1988), including biological systems (Kote-Jarai et al. 2011), it is critical to separate and quantify them. Methods such as mass spectrometry (MS) and high performance chromatography (HPLC) routinely ensure these in many laboratories around the world (Thorp et al., 2011). HPLC instrumentations provide crucial analytical data (Beitler, 1995) used to calculate or predict drug's affinity constants ($log\ K_a^{HSA}$) (Packard et al., 1996; Endo et al., 1982). For nearly half a century, analytical and testing methods as opposed to empirical approaches have played a key role in the identification of key diseases causative toxins (Barnes & Dourson, 1988; Bathija, 2003) and drug-like compounds to cure them (Kote-Jarai et al., 2011; Moore & Carpenter, 1999).

In this process, as tons of data are being produced with the analytical chemists struggling to make sense out of the bunch, the health researcher and health practitioners are faced with a constant task to make better and faster decisions in the area of disease treatment and prevention based on laboratory results. In the midst of all this, there is the requirement not only to produce efficient drugs, in enough quantities, to cure diseases but their development at the pace at which pandemics are spreading around the globe is also required, for example, cancer and HIV & AIDS.

As far as the identification of data needs is crucial in clinical laboratories, the quest for methods to determine biomarkers of exposure and effect of diseases in the public health is also growing fast. Hence, analysis of metabolites of drugs in humans or animals can provide a biomarker of exposure that is sensitive to low levels of exposure and correlates well with exposure concentrations. Methods for determining biomarkers of exposure in hu-

mans are needed to determine background levels in the population and levels at which bi‐
ological effects occur. For example, Abdel-Rahman et al. (1980a & 1980b) developed a
method to quantitatively and qualitatively measure the metabolites of chlorine dioxide
(e.g., ClO_2^-, and ClO^-) in biological fluids. These biomarkers may be used to indirectly
measure chlorine dioxide exposure.

In the absence of sensitive and reliable methods for determining diseases vector-borne me‐
tabolites and biomarkers of exposure, mechanistic models of tissue distribution of drug
compounds have been used (Rowley et al., 1997) to assess levels at which biological effects
occur in the population and mitigate disease occurrence. Poulin et al. (Poulin & Theil, 2002;
Poulin & Krishna, 1995) developed tissue composition-based equations for calculating tis‐
sue-plasma partition coefficients ($P_{t:p}$).

The following expressions are used (equations 1 and 2) (Yamashita & Hashida, 2004):

$$P_{t:p\,nonadipose} = \frac{P_{o:w}(V_{nlp} + 0.3V_{php}) + (V_{wt} + 0.7V_{pht})}{P_{o:w}(V_{nlp} + 0.3V_{php}) + (V_{wp} + 0.7V_{php})} \times \frac{f_{Up}}{f_{Ut}} \tag{1}$$

$$P_{t:p\,adipose} = \frac{D^*_{vo:w}(V_{nlt} + 0.3V_{pht}) + (V_{wt} + 0.7V_{pht})}{D^*_{vo:w}(V_{nlp} + 0.3V_{php}) + (V_{wp} + 0.7V_{php})} \times \frac{f_{Up}}{1} \tag{2}$$

$P_{o:w}$ is the n-octanol:buffer partition coefficient of non-ionized species at pH 7.4.

$D^*_{vo:w}$ is the olive oil:buffer partition coefficient of both the nonionized and ionized species
at pH 7.4, V is the fractional tissue volume content of neutral lipids (*nl*), phospholipids (*ph*),
and water (*w*), t is the tissue, p is the plasma and f_u is the unbound fraction.

These equations are based on the assumption that each tissue and plasma is a mixture of lip‐
ids, water and plasma proteins in which the drug can be homogeneously distributed.

The first term of these equations is based on the drug Lipophilicity-hydrophilicity balance
of tissues and plasma due to their lipid and water contents, while the second term of the
equation considers the binding to common proteins present in plasma and plasma intersti‐
tial space.

3. Use of *In-silico* techniques and chemical informatics in health research

Even though disease mapping has been done for over a hundred years, historically, the focus
in health research has been on person and conventional medicinal chemistry targeting specif‐
ic disease treatment with little regard for the implications of *de novo* molecular design costs.
Also, the need for high throughput screening in drug discovery research for public health in‐
terest has surfaced as a top priority in the last decade (Norris et al., 2000; Green et al, 1974).

However, the testing of much lead drug-like candidates often fails because of unsatisfactory ADME properties (Zmuidinavicius et al., 2003). In handling this challenge, ADME studies employing *in silico* techniques to improve the rate of success in the more costly downstream stages of drug development before clinical trials has gain tremendous interest in public health research methodologies. In this light, *in silico* ADME studies use various models developed for predicting ADME properties of compounds from their chemical structures and integrating them to simulate the kinetics at the organ or body levels Carlson & Segall, 2002; Podlogar et al., 2001; Ekins et al., 2001). Data-based approaches such as quantitative structure-activity relationship (QSAR) (Zhaou et al., 2001, Yoshida & Topliss, 2000), similarity searches, 3D-QSAR (Jones et al., 1996) and structure-based methods such as ligand-protein docking and pharmacophore modeling are amongst approaches currently used. In addition, several methods of integrating ADME properties to predict pharmacokinetics at the organ or body level have been studied (Agatonovic-Kustrin, 2001; Van de Waterbeemd, 2001; Ho et al., 2000; Jacoby et al., 2009; Lave et al., 1999). All these effort is to reduce the risk of late-stage attrition of drug development and to optimize screening and testing by looking at only the promising compounds. Recently, researchers have equally develop keen interest in *de novo* molecular design (Good et al., 1995; Clark et al., 1995), predictive modeling, graph theory, molecular similarity and diversity, virtual ligand docking, scaffold hopping (Langdon et al., 2010), multi-objective optimization Nicolaou et al., 2007), molecular descriptors, bioisosteric replacement, machine learning and evolutionary algorithms (Gillet, 2008).

In the last decade, a wide variety of descriptors used in QSAR studies have been developed (Khan et al., 2009; Miners et al., 2006). A subset of these descriptors is potentially useful for predicting ADME properties. Many QSAR studies on BBB permeation of drugs have been published recently. In the big junk of these works (Wichmann et al., 2007; Zhao et al., 2007; Cuadrado et al., 2007; Katritzky et al., 2006; Garg & Verma, 2006; Hemmateenejad et al., 2006; Narayanan & Gunturi, 2005) experimental data are represented as log*BB* constants or as qualitative (binary) index subdividing all compounds into 'CNS positive' and "CNS negative" classes according to presence or lack of central nervous system (CNS) activity. Table 1 summarizes the most notable QSAR models (log*PS*) representing blood/brain partitioning coefficients at equilibrium conditions (log*BB*)(Goodwin & Clark, 2005; Abbott, 2004) showing that typical data sets involved only 20-30 compounds.

Log*PS* is based on *in vivo* kinetic permeability measurements using intravenous administration (Oldendorf, 1971), brain uptake index (Bickel, 2005; Oldendorf, 1971) and *in situ* perfusion in rat or mouse (Bickel, 2005; Dagenais et al., 2000; Takasato et al., 1984;). Where, P (cm s^{-1}) is observed permeability across BBB, whereas S (cm^2/g) is surface area of brain capillary endothelium which equals to ~ 100 – 130 cm^2 in rats (Abraham, 2004; Bodor & Buchwald, 1999). PS product can be calculated from Kety-Renkin-Crone equation of capillary transport (K_{in} = F. (1- $e^{-PS/F}$)), and by its physical meaning, PS is equal to the unidirectional influx rate constant (K_{in}) corrected for cerebral blood flow (F). Earlier attempts of log*PS* prediction were largely restricted by the lack of high quality data. A subset of statistical techniques can deal with larger sets of molecular descriptors aimed at finding relationships or patterns in data

sets. For examples, multiple linear regressions (MLR) and partial least square (PLS) (Norinder & Österberg, 2001).

References	Descriptors	*N*	*R²*	RMSE
Levin[19]	Log(P.MW$^{-1/2}$)	22	0.83	-
Abraham and coworkers[10,24]	Solvation parameters (A, B, E, S, V$_x$)	18	0.95	0.48
		30	0.87	0.52
Bodor and Buchwald[11]	Log*P*(log*D*)	58	0.90	0.62
Liu et al.[5]	Log*D*, PSA, vsa base	23	0.74	0.50
Luco and Marchevsky[25] (review of earlier studies)	Log*P*, different *MW* functions	7-37	0.80-0.96	-
This literature	Log*P*, HD, HA, V$_x$, ion fractions (p*Ka* function)	125[a]	0.84	0.48
		53[b]	0.82	0.49

[a]Training set; [b]Validation set.

Table 1. Summary of QSAR Models for Predicting LogPS (Lanevskij et al., 2008)

These illustrations show how *in silico* chemistry or cheminformatics and high-throughput screening have increased the possibility of finding new lead compounds at much shorter time periods than conventional medicinal chemistry. With judicious selection of lead compounds and constant monitoring of physical properties (especially Lipophilicity (equation 1) or other major physchem parameter) during optimization, medicinal chemists have an opportunity to help alleviate the appalling attrition rates, estimated at 93–96% (Norris et al., 2000) in clinical drug development (Bhal et al.,. This means that physicochemical properties in small-molecule drug discovery are completely under the control of medicinal chemists and can easily be calculated before chemical synthesis. It is, however, important to emphasize here that when interpreting results from prediction models, that the predictions are only as good as the dataset used to create the model. So when we calculate a prediction, if the training set does not contain chemical structures that are similar to the particular compound in question, the predicted result may not be reliable, regardless of the actual result of the prediction.

From various literature sources (Bhal, 2007; Sazonovas et al., 2010), it is reported that ADME and toxicity prediction models can be a valuable part of many different research workflows, including virtual screening, metabolite identification, impurity analysis and chemical safety, reliability index (*RI*) value (0 – 1) in addition to the predicted probability result (Japertas et al., 2010) which is an indicator of how well spatial chemical space around a particular compound is represented within the training set of the model (Fig. 3)
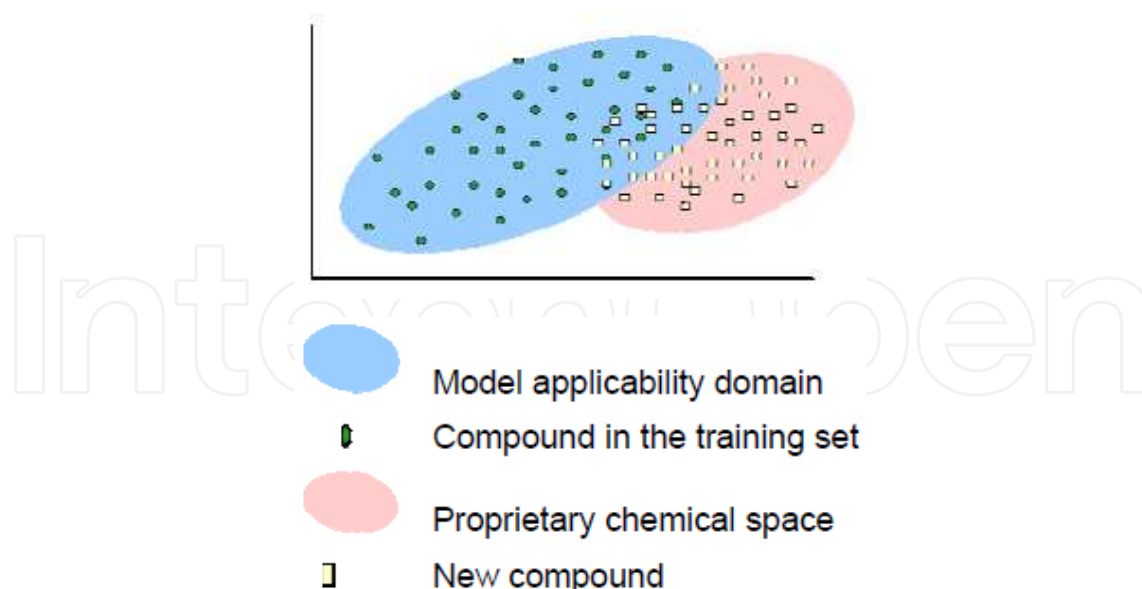
Figure 3. Model applicability domain versus proprietary chemical space, predicted using ACD/Log*D* Suite (Bhal & Coworkers, 2007)

The development of *in silico* technologies or chemical informatics over the last 20 years has, thus, provided a more powerful and rapid ability to examine clinical decision-making and answers to critical public research health issues based on scientifically validated models from an un biased an empirical question based on a minimally biased appraisal of all the relevant empirical studies. This, in turn, has fostered the discussion of policy of drug discovery relevant to health issues as well as health services and planning, in conjunction with the use of clinical investigations and disease surveillance. Such reviews aim to improve ethically relevant decisions in public healthcare research or policy. With this in mind, identification of a lead compound or biochemical starting point for a drug discovery program has been highlighted recently as a critically important activity, reflected by lead generation strategies being widely implemented in the pharmaceutical industry (Jupertas, 2007).

For a more robust process, calculated quantitative parameters will provide further information though slightly different from the core predictive pharmacokinetic data. These parameters show great inter relation. Such parameters include the drug's affinity constants (*log* K_a^{HSA}) to human serum albumin (the major carrier protein in plasma). Experimental data come from direct chromatographic determination of binding strength to that particular protein. These parameters are usually calculated as follows (Bhal et al. 2007):

$$\log K_a^{HSA} = \log\left(\frac{[LA]}{[L[A]]}\right). \tag{3}$$

Where [LA] is the concentration of ligand bound to albumin, [L] is that of free ligand, and [A] is the concentration of free albumin which, estimated at ~ 0.6 *mM* in human plasma.

%PPB values represent the overall fraction of drug bound in human plasma, i.e. accounts for interactions with different proteins: albumin, α-1-acid glycoprotein, liproteins, SHBG, transortion, etc. In vitro measurements of the extent of plasma protein binding usually involve equilibrium dialysis, ultrafiltration or ultracentrifugation methods. The supplementary distribution/V_d module calculates apparent volume of distribution of drugs in human body expressed in litres per *Kg* body weight (*L/Kg*). This expression is given as: *%PPB* = (1-f_u) x 100%, Where f_u is a fraction of free (unbound) drug in plasma (0 -1). Note that predictive models for %PPB and $\log_a^{HSA}$ are derived using GALAS (Global, Adjusted Locally According to Similarity) modeling methodology consisting of global (baseline) statistical model (based on *PLS* with multiple bootstrapping, using a predefined set of fragmental descriptors), and local correction to baseline prediction (based on the analysis of model performance for similar compounds from the training set = self-training library). Because of the percentage (*%PPB*) scale platform of the overall protein binding in plasma, values are linearized prior to modeling and converted to apparent serum affinity constants, *log K_{app},* the main parameter used in modeling. With this, final prediction can be converted back to *%PPB* using the following expression (Bhal et al., 2007):

$$\%PPB = \frac{\log K_{app}}{1 + \log K_{app}} x100\% \qquad (4)$$

To illustrate the performance of *%PPB* and $\log_a^{HSA}$ models, validation set compounds within Model Applicability Domain (MAD) (*RI ≥ 0.3*) is taken from Bhal et al. (2012) and shown in figures 4. In equation (y), the local part of the model provides the basis for estimating reliability of prediction by means of calculated Reliability Index (*RI*) values that range from $0-1$. 0 means unreliable prediction and 1 fully reliable or ideal prediction. The *RI* values can also be used for interpreting prediction results, for example in setting the cut-off point (*RI = 0.3*), meaning that if compound falls outside of the MAD, the respective prediction should be totally discarded from further analysis irrespective of *%PPB* and $\log_a^{HSA}$.
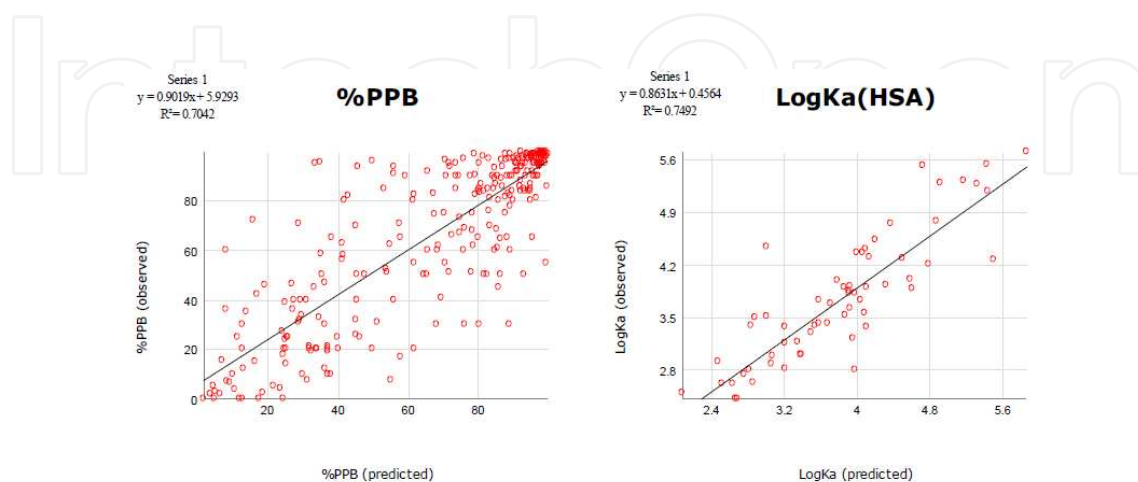

Figure 4. Predicted *versus* observed *%PPB* and $\log_a^{HSA}$ values for validation set compounds (ACD/Log*D* Suite).
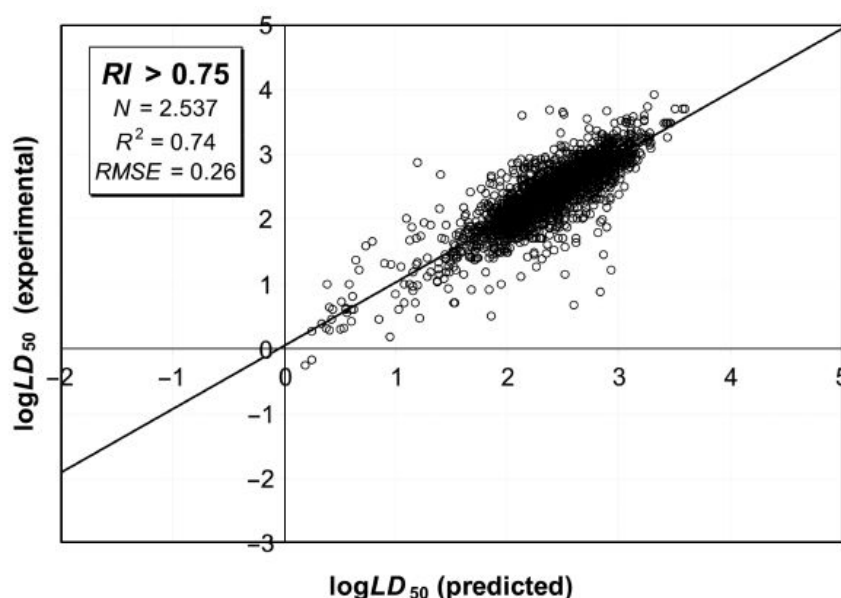
Figure 5. Predicted *versus* experimental values of *logLD₅₀* (with a very high *RI > 0.75*) %PPB and logₐ^HSA values for validation set compounds (Bhal et al., 2012).

In figures 5 and 6, both models produce highly accurate results, while even better statistical characteristics are observed if only predictions of moderate and high reliability index are considered.

In order to understand the behavior of drug compounds in the real world, Bhal et al. (2012) has used ACD/Predictors such as log$P$, log$D$, solubility, and pKa to evaluate and predict the likely behavior of 5-Methoxy-2-(1-piperidin-4-ylpropyl)pyridine (Fig. 7), a compound prior to its synthesis. Lipophilicity is represented by the descriptors log$P$ (also known as K_{ow} or P_{ow}) and log$D$, and is used, for example, to help predict *in-vivo* permeability of active compounds in drug discovery and the behavior of compounds in many other areas of health related research. The partition coefficient, P, is a measure of the differential solubility of a compound in two immiscible solvents and the most commonly used commonly used solvent system is octanol-1-ol/water.
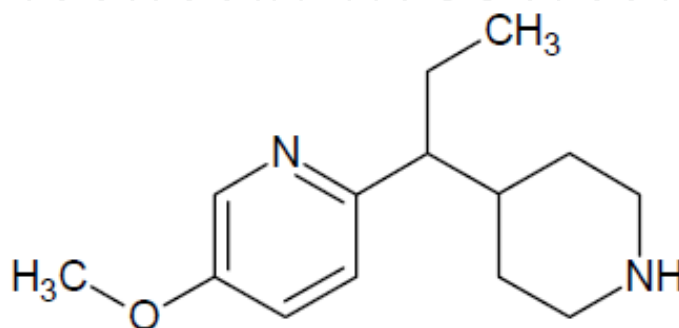

Figure 6. Chemical structure of 5-Methoxy-2-(1-piperidin-4-ylpropyl)pyridine, A

The *P* descriptor is a lipophilicity descriptor for neutral compounds, or where the compounds exist in a single form.

$$\text{Partition coefficient, } p = \frac{[Compound]_{octanol}}{[Compund]_{water}} \text{ and Log} P = \log_{10}(\text{Partition Coefficient})$$

Log$P$ = 1 can be explained as a ratio of 10 :1 organic:aqueous phase.

For ionizable solutes, the compounds may exist as a variety of different species in each phase at a given pH. D, typically used in the logarithmic for (lo$D$) representing the distribution coefficient which is the appropriate descriptor for ionizable compounds since it is a measure of the pH-dependent differential solubility of all species in the octanol/water system (Levin, 1980).

$$\text{Distribution coefficient, } D = \frac{\sum [Microspecies]_{octanol}}{\sum [Microspecies]_{water}}.$$

Bhal (2012) used methylamine to illustrate the difference between these two descriptors as follows: $MeNH_3^+ \rightleftharpoons MeNH_2 + H^+$, meaning that,

$$P_{MeNH_2} = \frac{[MeNH_2]_{octanol}}{[MeNH_2]_{water}} \text{ and/or } P_{MeNH_3^+} = \frac{[MeNH_3^+]_{octanol}}{[MeNH_3^+]_{water}}.. \text{ So, keeping the definition of D above in}$$

mind, we can write the expression of D as: $D = \frac{[MeNH_2]_{octanol} + [MeNH_3^+]_{octanol}}{[MeNH_{water}] + [MeNH_3^+]_{water}}.$

To accurately predict a compound's lipophilicity based on predicted molecular physical properties, it was imperative that the author applied the correct descriptor in an appropriate manner. In this context, log$D$ instead is used in lieu of log$P$ as the use of the latter to address lipophilicity concerns by drug research and manufacturing companies in the past (around 1980's and 90's) had resulted in incorrect conclusions for ionizable compounds. A good example would be to reproduce data generated by the previous author employing ACD/Log$D$ in his application to discuss the significance of applying log$D$ instead of log$P$ using drug discovery as an example where lipophilicity is correlated with *in vivo* permeability. Using the sample molecule in figure 6 (5-Methoxy-2-(1-piperidin-4-ylpropyl)pyridine), data presented in table were calculated.

pKa	Ionization centre
4.8	Pyridine
10.9	Piperidine

Table 2. pKa values of 5-Methoxy-2-(1-piperidin-4-ylpropyl)pyridine (Bhal et al., 2012)

The pH dependence of log$D$ for this sample molecule (5-Methoxy-2-(1-piperidin-4-ylpropyl)pyridine) is shown in figure 7 showing a plot of log$D$ versus pH, while figure 8 displays the changing ionic forms of molecule A.

Looking at the plot in figure 8, and according to Bhal el al. (2012), we can confirm that ionization of the compound greatly affects octanol-water partitioning and that lipophilicity cannot

be simplified to a constant. This is very so as lipophilicity of the compound is low below pH 12 when the majority of the compound exists in an ionized form. This would definitely be contradictory, of course, if log$P$ was examined alone (Predicted log$P$ is 2.7 ± 0.3 by author for comparison sake).
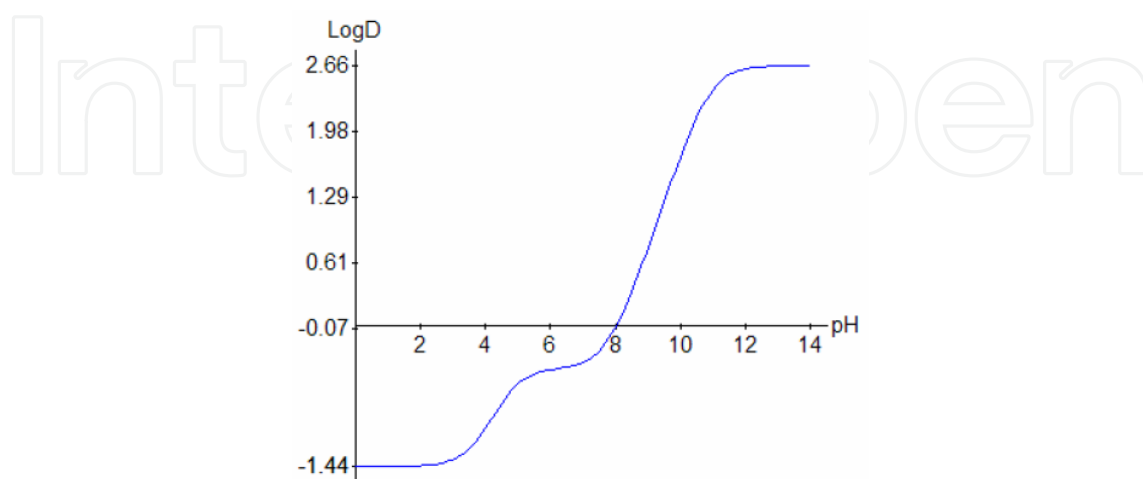

Figure 7. The logD curve of 5-Methoxy-2-(1-piperidin-4-ylpropyl)pyridine (ACD/LoD Suite), Bhal et al., 2012)

The author concludes that the negative values of log$D$ (-1.44 to 0) in the physiologically relevant pH range (pH 1–8) lead us to conclude that this compound would be more susceptible to higher aqueous solubility and of lower lipophilicity in the body. As a result we would expect membrane permeability to be poor. Also, it is true as seen from graph in figure 8 the neutral form of molecule A is almost non-existent at physiologically relevant pH (1–8). This neutral form possibly dominates at ~pH 13. Conclusively, compound A is highly associated with the lipid phase (>30 fold affinity for octanol over water), and thus will likely permeate biological membranes spontaneously.
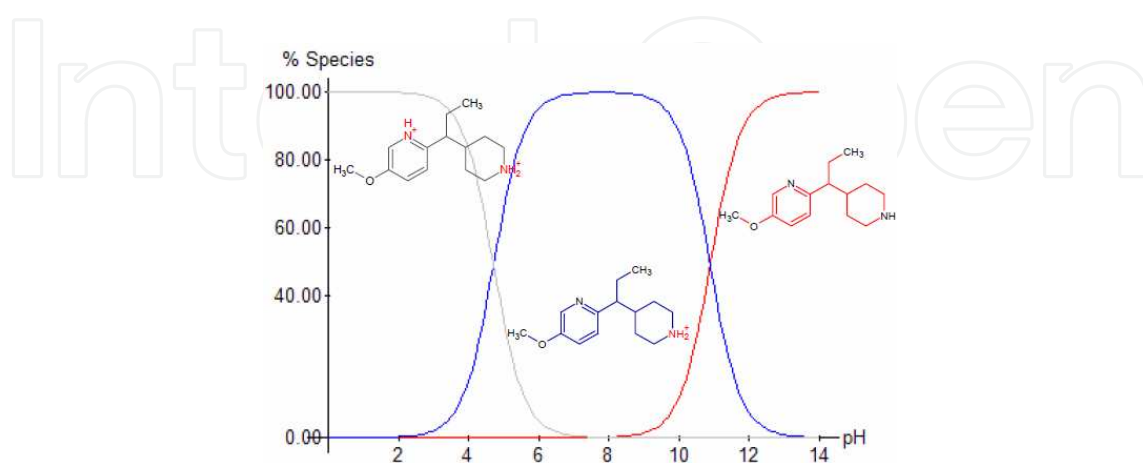

Figure 8. A graph illustrating the changing ionic forms of 5-Methoxy-2-(1-piperidin-4-ylpropyl)pyridine with pH (Bhal et al., 2012)

Figure 9 shows a schematic representation of the changing pH environments that an orally administered compound is likely to encounter in the gastrointestinal (GI) tract. From figure 9, we can observe that there is, thus, no constant pH in the body and it is therefore essential that we consider an appropriate pH when predicting the *in vivo* behavior of a drug candidate.
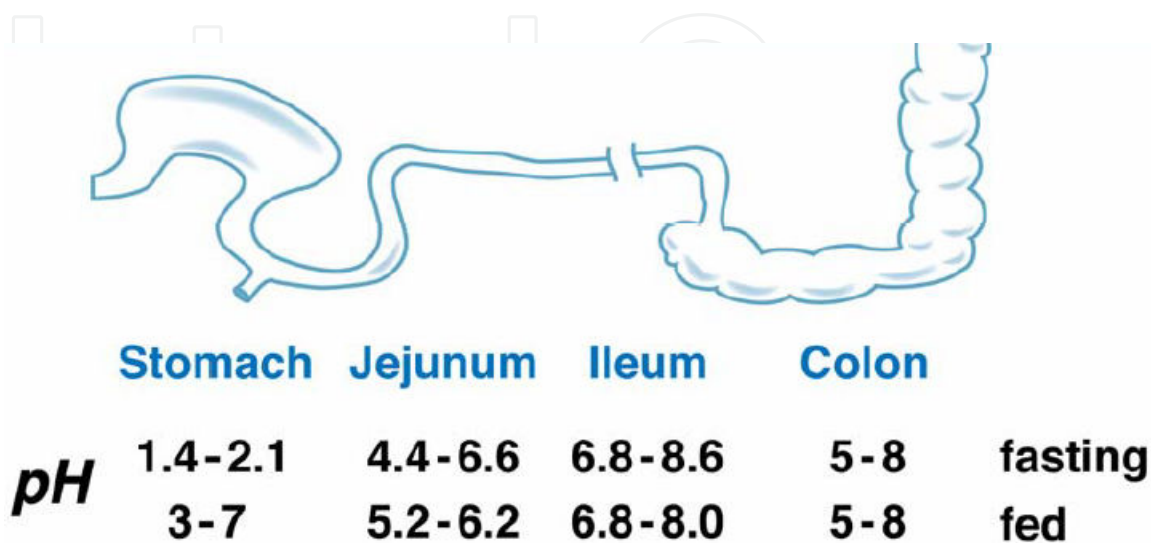

Figure 9. The pH environment of the human gastrointestinal tract (Bhal et al., 2012)

4. Implications of analytical methods and *in-silico* techniques in public health

The outcomes from a global network on the development of standardized analytical methods for the public and environmental health directly impacts on the quality of health of mankind worldwide. High quality data necessitates the development of harmonized study approaches and adequate reporting of data (Bouwmeester et al., 2011).

Priority public health scale can only be based on well-characterized dose-response relations derived from a systematic study of the bio-kinetics and bio-interactions of drugs or drug-like molecules at both organism and (sub)-cellular levels using validated analytical methods and pharmacokinetic studies. The ADME (absorption, distribution, metabolism and excretion) and toxicity effects is crucial to declare a particular molecule safe for the treatment of a particular disease and often clinical trials to arrive at a conclusive release of a new drug very costly, sometimes in the range of millions of dollars covering the cost of fundamental research through clinical trials or testing to manufacturing. Multiple content databases, data mining and predictive modeling algorithms, visualization tools, and high-throughput data-analysis solutions are being integrated to form systems-ADME/Tox (Ekins et al., 2005). More so, Ekins and co-authors (2005) reported that the functional interpretation and relevance of

complex multidimensional data to the phenotype observed in humans is the focus of current research in toxicology.

In fact, increased effort is needed to develop and validate analytical methods to determine ADMET effects in complex matrices such as the human body. This implies the use of validated analytical methods and *in-silico* chemistry to reduce time and cost at developmental stage. In addition, this would allow for systematic study of sets of drug-like molecules' reactivities with specific aim to generate a data set which would allow the establishment of dose-response relations. This approach is commonly referred to as quantitative structure-activity relationship (QSAR) and has a practical implication in the process of *in-silico* algorithms optimizations. A more general but direct implications of analytical methods and *in-silico* techniques in public health is that international cooperation and worldwide standardization of terminology, reference materials, models and protocols are needed to make progress in establishing lists of essential globally acceptable drug formulation methods, clinical trials, quality control and assurance. In this way, traditional pharmacopeia and those exploiting natural biodiversity can benefit from proven methods at the grassroots' levels, especially in developing countries.

Metabolomics, metabonomics, proteomics, pharmacogenomics and toxicogenomics, are groups of latest experimental approaches that are combined with high-throughput molecular screening of targets to provide a view of the complete biological system that is modulated by a compound with direct or indirect implications of analytical methods and *in-silico* techniques development for application in the public health sector (Ekins et al., 2005). In conclusion, it is widely recognized in industry-orientated research and development of APIs (active pharmaceutical ingredients) that predicting or determining the ADME/Tox properties of molecules would help to prevent failure of many of the compounds targeted before they reach the clinic. Many authors (Bhal et al., 2012; O'Donnell et al., 2012; Ndibewu et al., 2012; Thorp, 2011; Ekins et al., 2005) agree that this has, undoubtedly, been as a result of considerable research into developing better *in-silico*, *in vitro* and *in vivo* methods and models.

5. General conclusion

Driven by the changes in the working paradigm in the pharmaceutical and biotechnology, and now in environmentally health-related research, *in-silico* approaches will inevitable find their place. Some insilicoids have even mentioned that this approach will save the world. This is, probably, owing to the gains recorded in terms of cost reduction and efficiency for early stage drug discovery research to the point of manufacturing of some important drugs or vaccines to which the future of mankind mercilessly rely like terminal illnesses such as cancer or HIV-AIDS. All these diseases are unquestionable public health pandemonium requiring quick systematically robust holistic research approaches to halt their effects on human evolution.

Conclusively, ADMET data is tackled in three ways, namely: first, a variety of *in vitro* assays are further automated through the use of robust laboratory integrated system (LIM); secondly, *in silico* models are used to assist in the selection of both appropriate assays, as well as in the selection of subsets of compounds to go through these screens. And thirdly, predictive models have are then developed that might ultimately become sophisticated enough to replace *in vitro* assays and/or *in vivo* experiments. So, the need for ADMET information should start with the design of new compounds. Finally, this information shall influence the decision to proceed with synthesis either via traditional medicinal chemistry or combinatorial chemistry strategies. Although, it but obvious that at this point computational approaches are the most appropriate option to get this information, we should never, however, forget that predictions can be imperfect if flawed with errors. So, optimization should be the way forward once a lead compound is obtained and further optimized using more robust mechanistic models towards clinical trials before final manufacturing and authorization for absorption as a public health drug.

6. Recommendations

Since most models are rule-based and may use descriptors that are not easily understood by the chemist or not easily translated into better molecular structures, it is important to constantly train models of datasets. A combo approach, combining first generation (basic predictive descriptors) and second generation (meta-models) computational ADMET technologies would be the best way to go.

To then get value for your money, it is clearly demonstrated that ADME predictive tools is imperative, nowadays, in the health research programs in order to cut costs and propose reliable lead drug-like compounds. It is though highly desirable and recommendable to add in-house data in the prediction models whenever available. Sensitive and reliable high throughput instrumentations are a prerequisite in generating in-house analytical data necessary for efficient and useful predictive processes. Training data sets in models would be an added advantage for a wide range of investigations in health related research.

Acknowledgements

Author details

Peter Papoh Ndibewu* and Prince Ngobeni

*Address all correspondence to: ndibewup@tut.ac.za

Tshwane University of Technology, Faculty of Science – Arcadia Campus, Department of Chemistry, Pretoria, South Africa

References

[1]   Abbott, N. J. Prediction of blood–brain barrier permeation in drug discovery from in vivo, in vitro and *in silico* models. Drug Discov Today Technol. 1,407–416 (2004).

[2]   Abdel-Rahman, M. S, Couri, D., Jones, J. D. Chlorine dioxide metabolism in rat. J Environ Pathol Toxicol 3:421-430 (1980b).

[3]   Abdel-Rahman, M. S., Couri, D., Bull. R. J. Kinetics of ClO_2 and effects of ClO_2, ClO_2^-, and ClO_3 in drinking water on blood glutathione and hemolysis in rat and chicken. J Environ Pathol Toxicol 3,431-449 (1980a).

[4]   Abraham, M. H. The factors that influence permeation across the blood-brain barrier. Eur J. Med. Chem. 39, 235–240 (2004).

[5]   ACD/Log*D* Suite. www.acdlabs.com/logdsuite/ Advanced Chemistry Development, Inc., Toronto, ON, Canada. 24 April, (2007).

[6]   Agatonovic-Kustrin, S., Beresford, R. & Yusof, A. P. M.Theoretically-derived molecular descriptors important in human intestinal absorption. *J. Pharm. Biomed. Anal.* 25, 227–237 (2001).

[7]   Agoram, B., Woltosz, W. S. & Bolger, M. B. Predicting the impact of physiological and biochemical processes on oral drug bioavailability. *Adv. Drug Del. Rev.* 90, S41–S67 (2001).

[8]   Barnes, D. G, Dourson, M. Reference dose (RfD): Description and use in health risk assessments. Regul. Toxicol. Pharmacol. 8, 471-486 (1988).

[9]   Bathija, A. T. Health Risk to Fetuses, infants and children from stage 1 disinfectants and disinfectant by-products (D/DBPs). Toxicologist. 72(S-1):28 (2003).

[10] Beitler, M. K, Chin, H. B. Improved determination of chlorite and chlorate in rinse water from carrots and green beans by liquid chromatography and amperometric and conductivity detection. J. AOAC Int. 78(3),878-883 (1995).

[11] Bhal, S. K. Lipophilicity Descriptors: Understanding When to Use Log*P* & Log*D*.

[12] Bickel, U. How to measure drug transport across the blood-brain barrier. NeuroRx, 2, 15–26 (2005).

[13] Bodor, N., Buchwald, P. 1999. Recent advances in the brain targeting of neuropharmaceuticals by chemical delivery systems. Adv. Drug. Deliv. Rev. 36,229–254 (1999).

[14] Brown, N. Algorithms for Chemoinformatics. *Wiley Interdisciplinary Reviews: Computational Molecular Science.* 1(5), 716-726 (2011).

[15] Carlson, T. J. & Segall, M. D. Predictive, computational models of ADME properties. *Curr. Drug Disc.* 3, 34–36 (2002).

[16] Clark, D. E., Frenkel, D., Levy, S. A., Li, J., Murray, C. W., Robson, B., Waszkowycz, B., David R. Westhead, D. R. PRO LIGAND: An approach to de novo molecular design. 1. Application to the design of organic molecules. Journal of Computer-Aided Molecular Design. 9(1) 13-32 (1995).

[17] Cuadrado, M. U. Ruiz, I. L., Gomez-Nieto, M. A. QSAR models based on isomorphic and nonisomorphic data fusion for predicting the blood brain barrier permeability. J. Comput. Chem. 28, 1252–1260 (2007).

[18] Dagenais, C., Rousselle, C., Pollack, G. M., Scherrmann, J. M. Development of an in situ mouse brain perfusion model and its application to mdr1a P-glycoprotein-deficient mice. J. Cereb. Blood Flow Metab. 20, 381–386 (2000).

[19] Ekins, S., Nikolsky, Y., Nikolskaya, T. Techniques: Application of systems biology to absorption, distribution, metabolism, excretion and toxicity. Trends in Pharmacological Sciences. 26(4), 202-209 (2005).

[20] Ekins, S., Waller, C. L., Swaan, P. W., Cruciani, G., Wrighton, S. A. & Wikel, J. H. Progress in predicting human ADME parameters *in silico. J. Pharmacol. Toxicol. Methods* 44, 251–272 (2001).

[21] Endo, S., Hayashi, H., Wada, A. Affinity chromatography without immobilized ligands: a new method for studying macromolecular interactions using high-performance liquid chromatography. Anal. Biochem. 124,372-379 (1982).

[22] Garg, P. Verma, J. *In silico* prediction of blood brain barrier permeability: An Artificial Neural Network model. J. Chem. Inf. Model. 46, 289–297 (2006).

[23] Gillet, V. J. (2008) De Novo Molecular Design, in Evolutionary Algorithms in Molecular Design (ed D. E. Clark), Wiley-VCH Verlag.

[24] Good, A. C., Ewing, T. J. A., Gschwend, D. A., Kuntz, I. R. New molecular shape descriptors: Application in database screening. *Journal of Computer-Aided Molecular Design*, 9(1), 1-12 (1995).

[25] Goodwin, J. T., Clark, D. E. *In silico* predictions of blood-brain barrier penetration: Considerations to ''Keep in Mind''. J. Pharmacol. Exp. Ther. 315,477–483 (2005).

[26] Green, J. P., Johnson, C., Kang, S. Application of Quantum Chemistry to Drugs and Their Interactions. Annual Review of Pharmacology. 14 (4) 319-342 (1974).

[27] Han Van de Waterbeemd, H., Gifford, E. ADMET *IN SILICO* MODELLING:TO-WARDS PREDICTION PARADISE? www.nature.com/reviews/drugdisc 192 2(3) 192 -204 (2003).

[28] Hans Bouwmeester, H., Lynch, I., Marvin, H. J. P., Dawson, K. A., Berges, M., Braguer, D., Byrne, H. J., Casey, A., Chambers, G., Clift, M. J. D., Elia, G., Fernades, T. F., Fjellsbø, L. B., Hatto, P., Juillerat, L., Klein, C., Kreyling, W. G., Nickel, C., Riediker, M., Stone, V. Minimal analytical characterization of engineered nanomaterials needed for hazard assessment in biological matrices. Nanotoxicology, 5(1), 1-11(2011).

[29] Hemmateenejad, B., Miri, R., Safarpour, M. A., Mehdipour, A. R. 2006. Accurate prediction of the bloodbrain partitioning of a large set of solutes using ab initio calculations and genetic neural network modeling. J. Comput. Chem. 27,1125–1135 (2006).

[30] Ho, N. F. H., Raub, T. J., Burton, P. S., Barsuhn, C. L., Adson, A., Audus, K. L. & Borchardt, R. T. Quantitative approaches to delineate passive transport mechanisms in cell culture monolayers. *Drugs Pharm. Sci.* 102 (Transport Processes in Pharmaceutical Systems), 219–316 (2000).

[31] Jacoby, E., Boettcher, A., Mayr, L. M., Brown, N., Jenkins, J. L., Kallen, J., Engeloch, C., Schopfer, U., Furet, P., Masuya, K., Lisztwan, J. Knowledge-Based Virtual Screening: Application to the MDM4/p53 Protein-Protein Interaction. In: Jacoby, E. (ed.), *Chemogenomics, Methods in Molecular Biology, vol.* 575. Humana Press, 2009.

[32] Japertas, P. Fragmentiniu̢ ir statistiniu̢ metodu̢ naudojimas biologiškai aktyviu̢ junginiu̢ strukturos-aktyvumo sa̢ryšio tyrimuose. Ph.D. thesis, Vilniaus universitetas, Vilnius (2007).

[33] Japertas, P., Dapkunas, J. Sazonovas, A. Probabilistic Model for the Prediction of the Human Liver Microsomal Metabolism Regioselectivity. QSAR, 18th Euro QSAR. Sep 19, (2010).

[34] Japertas, P., Dapkunas, J. Sazonovas, A. Trainable In-Silico Screening Filter for Various Human Cytochrome P450 Isoforms Inhibition Liability. QSAR, 18th Euro QSAR. Sep 19, (2010).

[35] Japertas, P., Kubilius, R., Sazonovas, A., Lanevskij, K. ACD/Labs, Inc., A.Mickeviciaus g. 29, LT-08117 Vilnius, Lithuania *In Silico* Techniques in Lead Optimization ACD/Labs, Inc., A.Mickeviciaus g. 29, LT-08117 Vilnius, Lithuania, ACS Fall, August 29, (2011).

[36] Jones, J. P., He, M., Trager, W. F., Rettie, A. E. 3-dimensional QSAR for inhibitors of cytochrome P4502C9. *Drug. Metab. Dispo.*, 24, 1-6(1996).

[37] Katritzky, A. R., Kuanar, M., Slavov, S., Dobchev, D. A., Fara, D. C., Karelson, M., Acree, W. E., Jr., Solov'ev, V. P., Varnek, A. Correlation of blood-brain penetration using structural descriptors. Bioorg. Med. Chem. 14,4888–4917 (2006).

[38] Khan, Mahmud, T.H. Sylte, Ingebrigt. Multivariate Linear Regression Models Based on ADME Descriptors and Predictions of ADMET Profile for Structurally Diverse Thermolysin Inhibitors. Letters in Drug Design & Discovery. 6(6) 428-436(9) (2009).

[39] Kote-Jarai, Z.; Amin Al Olama, A.; Leongamornlert, D.; Tymrakiewicz, m.; Saunders, E.; Guy, M.; Giles, G. G.; Severi, G.; Southey, M.; Hopper, J. L.; Sit, K. C.; Harris, J. M.; Batra, J.; Spurdle, A. B.; Clements, J. A.; Hamdy, F.; Neal, D.; Donovan, J.; Muir, K.; Pharoah, P. D. P.; Chanock, S. J.; Brown, N.; Benlloch, S.; Castro, E.; Mahmud, N.; O'Brien, L.; Hall, A.; Sawyer, E.; Wilkinson, R.; Easton, D. F.; Eeles, R. A. Identification of a novel prostate cancer susceptibility variant in the *KLK3* gene transcript. *Hum. Genet.* 129(*6*), 687-694(2011).

[40] Lanevskij, K., Japertas, P., Didziapetris, R., Petrauskas, A. Drug Discovery Interface. Ionization-Specific Prediction of Blood–Brain Permeability. *Wiley InterScience* (www.interscience.wiley.com). 4 (3) (2008).

[41] Langdon, S. R., Ertl, P., Brown, N. Bioisosteric Replacement and Scaffold Hopping in Lead Generation and Optimization. *Mol. Inf.* 29(*5*), 366-385 (2010).

[42] Langdon, S. R.; Brown, N.; Blagg, J. Scaffold Diversity of Exemplified Medicinal Chemistry Space. *J. Chem. Inf. Model.* 51(9), 2174-2185 (2011).

[43] Lave, T., Coassolo, P., Reigner, B. Prediction of hepatic metabolic clearance based on interspecies allometric scaling techniques and in vitro-in vivo correlations. *Clin. Pharmacokinet.,* 36, 211-231 (1999).

[44] Leeson, P. D., Springthorpe, B. The influence of drug-like concepts on decision-making in medicinal chemistry. Natuure reviews, *drug discovery*, 6, (9), 881-890 (2007).

[45] Levin, V. A. Relationship of octanol/water partition coefficient and molecular weight to rat brain capillary permeability. J. Med. Chem. 23, 682–684 (1980).

[46] Lipinski, C. A. Lead- and drug-like compounds: the rule of five revolution. *Drug Discov. Today Technol.* 1, 337–341 (2004).

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

[48] Lu, A. Y. H.,Wang, R.W., and Lin, J. H. Cytochrome P450 in vitro reaction phenotyping: a re-evaluation of approaches used for P450 isoform identification. Drug Metab Dispos 31, 345–350 (2003).

[49] Miners, J. O., Knights, K. M., Houston, J. B., Mackenzie, P.I. *In vitro-in vivo* correlation for drugs and other compounds eliminated by glucuronidation in humans: pitfalls and promises. *Biochem. Pharmacol.* 71(11), 1531-9 (2006).

[50] Moore, D. A., Carpenter, T. A. Spatial Analytical Methods and Geographic Information Systems: Use in Health Research and Epidemiology Epidemiologic Reviews 21, (2), 143-161 (1999).

[51] Narayanan, R., Gunturi, S. B. *In silico* ADME modelling: Prediction models for blood-brain barrier permeation using a systematic variable selection method. Bioorg. Med. Chem. 13,3017–3028 (2005).

[52] Ndibewu, P. P., LL Sibali, L. L., Sieliechi, J. Mnisi, L. R. Application of *in-silico* Chemistry in understanding health risk and bioaccumulation Mechanism of POPs. *ENVIRONMIN2012*, Loskop Dam, Forever Resorts, Middleburg, South Africa. 11 – 15 March (2012)

[53] Nicolaou, C. A., Brown, N., Pattichis, C. K. Molecular Optimization Using Multi-Objective Methods. *Curr. Opin. Drug. Discov. Devel.* 10*(3)*, 316-324 (2007).

[54] Norinder, U. & Österberg, T. Theoretical calculation and prediction of drug transport processes using simple parameters and partial least squares projections to latent structures (PLS) statistics. The use of electrotopological state indices. *J. Pharm. Sci.* 90, 1076–1085 (2001).

[55] Norris, D. A., Leesman, G. D., Sinko, P. J. & Grass, G. M. Development of predictive pharmacokinetic simulation models for drug discovery. *J. Contr. Rel.* 65, 55–62 (2000).

[56] O'Donnell, III, J. J., O'Donnell, J. T., Juettner, V. M. Clinical Pharmacology for the Primary Care Physician. DM, 10,558-610 (2012).

[57] Oldendorf, W. H. Brain uptake of radiolabeled amino acids, amines and hexoses after arterial injection. Am. J. Physiol. 221, 1629–1639 (1971).

[58] Packard, B. Z., Toptygin, D. D., Komoriya, A., Brand, L. Profluorescent protease substrates: Intramolecular dimmers described by the exciton model. Proc. Natl. Acad. Sci. 93,11640-11645 (1996).

[59] Parrott, N. & Lavé, T. Prediction of intestinal absorption: comparative assessment of GastroPlus and iDEA. *Eur. J.Pharm. Sci.* 17, 51–61 (2002).

[60] Podlogar, B. L., Muegge, I. & Brice, L. J. Computational methods to estimate drug development parameters. *Curr. Opin. Drug Disc. Dev.* 4, 102–109 (2001).

[61] Poulin, P., Krishnan, K. A biologically-based algorithm for predicting human tissue: blood partition coefficients of organic chemicals. *Hum. Exp. Toxicol.*, 14,273-280 (1995).

[62] Poulin, P., Theil, F. P. Prediction of pharmacokinetics prior to *in vivo* studies. 1. Mechanism-based prediction of volume of distribution. *J. Pharm. Sc.,* 91,129-156 (2002).

[63] Pradeep, S. Bioinformatics Infrastructure FacilityNorth-East Institute of Science & Technology. http://www.scribd.com/salampradeep/d/14651445-Chemical-Informatics. http://biotech-job.blogspot.com/ (2009). [Accessed 01 June 2012].

[64] Rowley, M., Kulagowski, J. J., Watt, A.P., Rathbone, D., Stevenson, G. I., Carling. R. W., Baker, R., Marshall. G. R., Kemp, J. A., Foster, A. C, Grimwood, S., Hargreaves, R., Hurley, C., Saywell, K. L., Tricklebank, M. D., Leeson, P. D. Effect of plasma protein binding on in vivo activity and brain penetration of glycine/NMDA receptor antagonists. J Med. Chem. 40, 4053–4068 (1997).

[65] Sazonovas, A., Japertas, P., and Didziapetris, R. Estimation of reliability of predictions and model applicability domain evaluation in the analysis of acute toxicity (LD50). SAR QSAR *Environ Res*. 21, 127–148 (2010).

[66] Takasato, Y., Rapoport, S. I., Smith, Q. R. An in situ brain perfusion technique to study cerebrovascular transport in the rat. Am. J. Physiol. 247, H484–H493 (1984).

[67] Thorp, D. Powerful Analytical Chemistry Software Solutions for CROs Through ACD/Labs Subscription Service Advanced Chemistry Development, Inc. Toronto, ON, Canada, www.acdlabs.com. Feb 23, (2011).

[68] Vallero, D. A. Ethical Decisions in Emergent Science, Engineering and Technologies, In: Emerging informatics – Innovative Concepts and Applications. Edited by Miah, S. J. InTech Publisher, Janeza Trdine 9, 51000 Rijeka, Croatia, 4, 21 (2012).

[69] Van de Waterbeemd, H. Quantitative structure-absorption relationships. In: *Pharmacokinetic Optimization in Drug Research: Biological, Physicochemical and Computational Strategies*, Testa, B., Van de Waterbeemd, H., Folkers, G. & Guy, R. (Eds), Verlag HCA, Zurich (2001), pp. 499–511.

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

[71] Wichmann, K., Diedenhofen, M., Klamt, A. Prediction of blood-brain partitioning and human serum albumin binding based on COSMO-RS sigma-moments. J. Chem. Inf. Model. 47, 228–233 (2007).

[72] Yamashita, F., Hashida, M. Insilico Approaches for predicting ADME properties of drugs. *Drug Metab. Pharmacokin*. 19(15), 327-338 (2004).

[73] Yoshida, F., Topliss, J. G. QSAR models for drug human oral bioavailability. *J. Med. Chem*. 43, 2575-2585 (2000).

[74] Yu, L. X., Gatlin, L. & Amidon, G. L. Predicting oral drug absorption in humans. *Drugs Pharm. Sci.* (Transport Processes in Pharmaceutical Systems), 102, 377–409 (2000).

[75] Zhao, Y. H., Abraham, M. H., Ibrahim, A., Fish, P. V., Cole, S., Lewis, M. L., de Groot, M. J., Reynolds, D. P. Predicting penetration across the blood-brain barrier from simple descriptors and fragmentation schemes. J. Chem. Inf. Model. 47,170–175 (2007).

[76] Zhao, Y. H., Le, J., Abraham, M. H., Hersey, A., Eddershaw, P. J., Luscombe, C. N., Boutina, D., Beck, G., Sherborne, B., Cooper, I. & Platts, J. A. Evaluation of human intestinal absorption data and subsequent derivation of a quantitative structure-activity relationship (QSAR) with the Abraham descriptors. *J. Pharm. Sci.* 90, 749–784 (2001).

[77] Zmuidinavicius, D., Japertas, P., Petrauskas, A., Didziapetris, R. Progress in Toxinformatics: the Challenge of Predicting Acute Toxicity. *Curr. Top. Med. Chem.* 9(1), 2003.