In silico quantitative structure pharmacokinetic relationship modeling on cardiovascular drugs: Half-life

Chouhan Jagtar Singh, Jat Rakesh Kumar and Paul Yash

Abstract
An estimate of Half-life ($t_{1/2}$) is a very vital pharmacokinetic parameter to assess safety and efficacy of any drug molecule, which helps to treat serious cardiovascular diseases i.e., hypertension, arrhythmias. This study was conducted to develop Quantitative Structure Pharmacokinetic Relationship (QSPkR) for the prediction of $t_{1/2}$ in men for congeneric series of twenty two cardiovascular derivatives, using computer assisted Hansch approach. The QSPkR correlations were duly analyzed using a battery of apt statistical procedures and validated using leave-one-out (LOO) approach. Analysis of several hundreds of QSPkR correlations developed in this study revealed high degree of cross-validated coefficients ($Q^2$) using LOO method ($p<0.001$). The overall predictability was found to be high half-life ($t_{1/2}$) ($R^2=0.9352$ $F=15.88$ $S^2=5.29$ $Q^2=0.7932$ $p<0.001$). Half-life ($t_{1/2}$) in the present QSPkR investigations was found to depend upon electrostatic and topological etc. Half-life ($t_{1/2}$) does not seem to have any dependence on lipophilic and electronic parameters indicating that the hydrophobic and ionic bonding of Cardiovascular drugs is negligible.

Keywords: Quantitative structure pharmacokinetic relationships (QSPkR), Half-life, in silico ADME, cardiovascular drugs

Introduction
Drug discovery and development is an intense, lengthy and an interdisciplinary endeavor. Drug discovery is mostly portrayed as a linear, consecutive process that starts with target and lead discovery, followed by lead optimization and pre-clinical in vitro and in vivo studies to determine if such compounds satisfy a number of pre-set criteria for initiating clinical development. For the pharmaceutical industry, the number of years to bring a drug from discovery to market is approximately, 12-14 years and costing up to $1.2 - 1.4$ billion dollars. Nearly 40% of the drug candidates fail during the clinical trials owing to their poor pharmacokinetic properties. This is an economic disaster as the failed drugs have been in the pipeline for several years with high expenditure of efforts, time and money invested in their development. More recently in silico ADME modelling has been investigated as a tool to optimize selection of the most suitable drug candidate for development. The use of computational models in the prediction of ADME properties has been growing rapidly in drug discovery as they provide immense benefits in throughput and early application of drug design [1].

The major aim of in silico QSPkR is to enable the drug designer to modify the chemical structure of a pharmacodynamically active drug so that its pharmacokinetic property may be altered without compromising pharmacodynamic potential. An early assessment of ADME properties will help pharmaceutical scientist to select the best drug candidate for development and as well as to reject those with a low plausibility of success. In silico QSPkR technique tends to save considerable amount of time, money, animal life and involvement of “normal, healthy and drug –free volunteers” required for conducting the experimental pharmacokinetic studies [2].

Half-life ($t_{1/2}$) is a vital pharmacokinetic parameter because it is directly related to the bioavailability and can be used in assessing the efficacy of drug. Hence it is important to predict the values of half-life ($t_{1/2}$) during drug discovery, so that compounds with acceptable rate of absorption can be identified and those with poor bioavailability can be eliminated. The current study was conducted to investigate in silico QSPkR amongst Cardiovascular drugs for half-life. Cardiovascular drugs were chosen for QSPkR as this category of drugs has extensively been used in cardiovascular disorders in the treatment of hypertension, arrhythmias and angina pectories etc. Moreover, Cardiovascular drugs consist of significant number of
The Pharma Innovation Journal

compounds thoroughly investigated for their pharmacokinetic performance particularly Half-life (t_{1/2}) (n=18) Further, the congeners in this class have many common pharmacokinetic characteristics, mechanism and degree of affinity with body tissues [1].

Application
1. As an instrument for prediction
   Estimation of physicochemical properties using subsistent Constants

   ![](image)

   Fig 1: Quantitative Structure Pharmacokinetic Relationship (QSPkR) modeling

   Pharmacokinetic parameters
   - Absorption
   - Distribution
   - Protein Binding
   - Metabolism
   - Renal Excretion

   Structural descriptors
   - Lipophilic Parameters
   - Topological Parameters
   - Quantum chemical Parameters
   - Geometrical Parameters
   - Electronic parameters
   - Electrostatic Parameters
   - Steric parameters
   - Constitutional parameters

   Multivariate Statistical Analysis

   Methods
   QSPkR was conducted amongst cardiovascular drugs employing extra-thermodynamic Multi Linear Regression Analysis (MLRA or Hansch) approach. The general steps for developing QSPkR model include data set selection, chemical structure entry, 3D structure generation and descriptor calculation, model construction that involves selection of descriptors and validation of testing set using a Pentium dual core (Intel, USA), Desktop (IBM, USA) with 1GB RAM and 160 GB Hard Disk.

   Dataset Selection
   21 Cardiovascular drugs with known human half-life (t_{1/2}) values were selected from literature. In order to ensure that experimental variations in determining half-life (t_{1/2}) do not significantly affect the quality of our datasets. Half-life (t_{1/2}) values obtained from healthy adult males after oral administration of drug were used for constructing the dataset. Half-life (t_{1/2}) value of each of these compounds was also log-transformed (Log Vd) to normalize the data to reduce unequal error variance.

   Molecular structure and descriptors
   Chemical structures were drawn using suitable templates under Chem draw 7.0 software (Cambridge Soft Corporation, Cambridge, MA) and energy minimization was carried out using Chem3D pro 3.5 software and the files were saved as MDL molfiles. Molfiles generated by Chem3D were exported to CODESSA 2.0 software (Semiichem, Shawnee, USA) for calculation of more molecular descriptors.

   Reduction of the number of compounds to be synthesized
   Faster detection of the most promising compounds
   Avoidance of synthesis of compounds with same activity

   As a diagnostic instrument
   - Information on possible types of interaction forces
   - Information on the nature of receptor
   - Information on the mechanism of fraction

   Detection of exceptions (outlier)

   Multivariate statistical analyses
   Attempts were made to correlate various descriptors with the half-life (t_{1/2}) values. The initial regression analysis was carried out using heuristic analysis followed by best MLRA (RGMS) options of CODESSA software. All the descriptors were checked to ensure that value of each descriptor was available for each structure and there is a significant variation in these values. Descriptors for which values were not available for every structure in the data in question were discarded. Thereafter, the one and multiple parameter correlation equations for each descriptor were calculated.

   Pharmacokinetic data of Half-life (t_{1/2}) parameter available for 21 cardiovascular drugs was analyzed, limiting the ratio of descriptors: drug to 4:1. As a final result, the heuristic method yields a list of the best ten correlations each with the highest r^2 and F-values. Many such attempts were carried out to obtain significant correlations for cardiovascular drugs. A set of important descriptors found to significantly ascribe the variation of half-life, was constructed. Further, a search for the multi-parameter regression with the maximum predicting ability was performed. A number of sets of descriptors were thus made and MLRA performed with half-life. Regression plots of each correlation thus attempted were examined. Residual plots were also studied for absence of randomization and distinct patterns to eliminate chance correlations.

   Validation of Testing Set
   The predictability of the final models was tested by LOO method. Briefly, the descriptors of one compound are removed, the model is redefined and the target properties of
the removed compound are predicted. This process is repeated until all target properties have been predicted once for each drug. A value of cross-validated $R^2$, commonly called $Q^2$, is then computed analogous to the conventional $R^2$ according to equation no.1:

$$Q^2 = 1 - \frac{\sum (y_{\text{pred}} - y_{\text{obs}})^2}{\sum (y_{\text{obs}} - \bar{y}_{\text{obs}})^2}$$  \hspace{1cm} …(1)$$

A model with good predictive performance has a $Q^2$ value close to 1, models that do not predict better than merely chance alone can have negative values.

The F-values were computed according to Equation no.2:

$$F = \frac{S_1^2}{S_2^2}$$  \hspace{1cm} …(2)$$

Where, $S_1$ is variance between samples and $S_2$ variance within samples.

The values of computed F-ratio were compared with the critical values tabulated in statistical texts and levels of significance discerned. The correlations found to be statistically significant were compiled from CODESSA software.

**Results and Discussion**

The value of half-life ($t_{1/2}$) of all cardiovascular drugs was found to depend upon various electrostatic parameters. As seen from Table 1, the correlations of $t_{1/2}$ with various descriptors were statistically significant ($p<0.001$) with very good prediction power of the best correlation $R^2(0.9352-0.9140)$ and $Q^2(0.7930-0.8116)$. Logarithmic transformations $R^2(0.9350-0.9020)$ $Q^2 (0.7739-0.7930)$ slightly rather reduce the degree of correlations. There was quite significant reduction is $S_2$ values, attributable to reduction in the magnitude of the property values. The values were found to be highly predictable ($p<0.001$) during the QSPR studies above. As lipophilic and electronic parameters were not observed to be considerably significant, the diffusional interactions tend to outweigh the permeation ones.

Dependence of biological half-life on electrostatic descriptors has also been reported in literature. Also, the values of half-life have been shown depends upon the topological parameters.

**Table 1**: Significant linear, logarithmic relationships for a series of 21 cardiovascular drugs using Half-life ($t_{1/2}$) as pharmacokinetic parameter

<table>
<thead>
<tr>
<th>Eqns.s</th>
<th>M</th>
<th>$R^2$</th>
<th>$Q^2$</th>
<th>SSE</th>
<th>F</th>
<th>P&lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_{1/2}$ = - 5.25 + 1.66FPSA+3.15MPCC + 4.51HDCA+5.0MPCH-3.8RNCS</td>
<td>5</td>
<td>0.9352</td>
<td>0.7930</td>
<td>5.2397</td>
<td>15.88</td>
<td>0.001</td>
</tr>
<tr>
<td>$t_{1/2}$ = - 5.85 + 4.27FPSA+2.64MPCC + 4.33HDCA+5.54MPCH-3.6RNCS</td>
<td>5</td>
<td>0.9343</td>
<td>0.8116</td>
<td>5.3166</td>
<td>15.64</td>
<td>0.001</td>
</tr>
<tr>
<td>$t_{1/2}$ = - 4.85 + 2.07FPSA+2.52MPCC + 4.08HDCA+4.98MPCH-3.8RNCS</td>
<td>5</td>
<td>0.9140</td>
<td>0.8114</td>
<td>5.3388</td>
<td>15.57</td>
<td>0.001</td>
</tr>
<tr>
<td>Log $t_{1/2}$ = - 5.72 + 4.37FPSA+2.74MPCC + 3.75HDCA+5.38MPCH-3.61RNCS</td>
<td>5</td>
<td>0.9320</td>
<td>0.7739</td>
<td>5.4998</td>
<td>15.08</td>
<td>0.001</td>
</tr>
<tr>
<td>Log $t_{1/2}$ = - 6.77 + 5.28ASIC + 3.45 HDCA -4.55RNCS-3.56ASIC-3.06PP</td>
<td>5</td>
<td>0.9023</td>
<td>0.7803</td>
<td>6.2129</td>
<td>18.46</td>
<td>0.001</td>
</tr>
<tr>
<td>Log $t_{1/2}$ = - 5.25 + 4.17ASIC+2.37ACIC-2.18PP-1.5AIC-1.43MPCO</td>
<td>5</td>
<td>0.9352</td>
<td>0.7930</td>
<td>5.2397</td>
<td>15.88</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Figure 1 shows the linear and residual plots between the values of untransformed $t_{1/2}$, as reported in literature and those predicted using multi parameter QSP&K investigations for a Series of 21 cardiovascular drugs. Figure 2 show the corresponding plots for log- transform of half-life

![Fig 1: Plot between the predicted and reported values of Half-life ($t_{1/2}$) for QSPR of Cardiovascular drugs. The inset shows the corresponding residual plot](image-url)
Conclusions
Analysis of several hundreds of QSPkR correlations and consequent profiles in the current investigations on Cardiovascular drugs revealed that: The quantitative relationships for various pharmacokinetic parameters were highly predictable in most cases ($p<0.001$).

Half-life ($t_{1/2}$) in the present QSPkR investigations was found to depend upon various topological parameters influencing $t_{1/2}$ encompassed AIC-2, ASIC-2, ACIC-1, ASIC-3, Kier shape indices, Kier Hall indices, etc. The vital electrostatic parameters encompassed FPSA-3, RNCS, HDCA-2, MPCNA etc., as is vivid from Table 1.

It is a duly accepted fact that the pharmacokinetic performance of a drug is not merely a function of its physicochemical nature, but of the complexities of biological system(s). The list of biological variants embodies the somatic (age, sex, weight, etc.) psychological, pathological (nature and degree of disease), environmental, nutritional, genetic, hereditary and diurnal (chronopharmacokinetics) status of the human subjects. This causes a great deal of variation in pharmacokinetic profiles amongst the patients/volunteers undergoing study. The literature values of the pharmacokinetic parameters taken up in the present investigations, pertain to diverse subject populations, hailing from different age groups, gender, races, nutritional and physical attributes, etc., studied in different geographical regions under different weather conditions. Considering these potentially high inter-subject and intra-subject variations amongst pharmacokinetic parameters, the correlations in QSPkR studies even with moderate statistical significance ($p<0.05$) cannot even be overlooked. Accordingly, the QSPkR results ($p<0.001$) should be taken up very high level of credence and confidence. It is expedient to render deeper insight for future studies on such *in silico* ADME predictive relationships of very high statistical significance.

References