Evaluation of selected quantitative structure permeability relationship (QSPR) based mathematical models for the prediction of skin permeability of *Camellia sinensis* (tea) compounds

Sudeshna Kundu\(^a\), Anusua Singh\(^a,b\), Sharmistha Samanta\(^a\), Dipankar Chattopadhyay\(^c\), Raghwendra Mishra\(^b\*\)
Roshnara Mishra\(^a\*)

**ABSTRACT**

*Background:* Skin permeability coefficient (log\(K_p\)) is a major determinant for topical drugs. *In-vitro* and *ex-vivo* determination of log\(K_p\) is expensive, time and labor-intensive, and difficult to apply to large databases. QSPR models derived from the statistical correlation between descriptors of the compounds with the *in-vitro* or *ex-vivo* permeation data are used extensively. The vast number of QSPR equations makes the selection of a particular equation to screen a database difficult. *Objective:* This study has evaluated common descriptor-based equations to select the best suitable equation for screening the phytochemical library of *Camellia sinensis*. *Methods:* Seven QSPR-based models were used to estimate and compare the log\(K_p\) of tea compounds. The best method was selected with respect to the gold standard log\(K_p\). *Result:* The model of Potts and Guy showed close proximation with the gold standard log\(K_p\) and had the highest association along with least RMSE value and least deviation. According to this method, approximately 37.38, 35.35 and 27.27% of the tea compounds were found to have high, good and poor log\(K_p\) respectively. *Conclusion:* Potts and Guy equation can be effectively used to screen the phytochemical library of *Camellia sinensis*. This study has potential applications in the field of topical medicine and cosmetics.

*Keywords:* Skin permeation coefficient, QSPR, topical, *Camellia sinensis*, tea, phytochemical.

Indian Journal of Physiology and Allied Sciences (2023); DOI: 10.55184/ijpas.v75i03.143  
**ISSN:** 0367-8350 (Print)

**INTRODUCTION**

Drug delivery through the skin offers an attractive and alternative route of drug administration over oral and parenteral drug delivery. It bypasses the first-pass metabolism and overcomes the limitations of oral drugs like, GI degradation, hepatic clearance etc. Moreover, this non-invasive route is convenient and preferred over the parenteral route of administration.\(^1\) Unlike systemic application, local administration of topical drug maximizes therapeutic efficacy by increasing local tissue concentration and minimizing adverse effects of nonspecific targeting.\(^2\) In spite of such advantages, poor skin permeation of drug candidates is a major limitation for topical route. The stratum corneum, the thickest layer containing numerous coverings of keratinized corneocytes, is the primary barrier for drug permeation through the skin.\(^3\) However, some provisions exist for transferring natural compounds across the skin, including the intercellular, intracellular and follicular pathways. The intercellular path facilitates the transmission of hydrophilic drugs, whereas the intracellular path is suitable for the transport of lipophilic drugs. The follicular or trans-appendageal path allows the rapid transfer of drugs directly to the infundibulum region.\(^4\) The barrier function of skin, imparted by the unique arrangement of hydrophilic keratin filaments compactly packed with hydrophobic lamellar lipids,\(^5\) presents a challenge to the study of skin permeation. Different types of skin permeation measurements have been developed to assess the dermal absorbed dose. Among them the most common are: percent absorbed and permeability coefficient (log\(K_p\)). Measurement of the former requires careful surface area control and is highly dependent on the magnitude and duration of exposure. Therefore, being independent of time, volume, exposure concentration and cross-laboratory comparison, log\(K_p\) is increasingly used.\(^6\) Log\(K_p\), the major determinant of the bioavailability of topical drug candidates, is derived from Fick’s law and is determined...
Selection of QSPR model for screening tea compounds

by several in-vitro and ex-vivo methods using polymer membranes or excised mammalian skin. Besides ethical considerations, these methods are expensive and time and labor-intensive. While for individual drug candidates, these methods are still worth the trouble; these are, therefore, difficult to apply for large databases. Quantitative structure permeability relationship (QSPR) based mathematical models for logK_p determinations are used extensively to screen compound libraries. The output of the QSPR model is in form of an equation constructed from the statistical correlation between physicochemical or molecular descriptors of the compounds with the in-vitro or ex-vivo permeation data.

In the last few decades, many QSPR based mathematical models for predicting skin permeability with different degrees of accuracy and limitations have been reported and validated. These models have used different combinations of mechanistic and empirical descriptors, in-vitro or ex-vivo permeation data and different statistical or machine learning tools for correlation study. However, every model has its limitations. No model can be used universally as the complexity of skin permeation physiology is too great to be bound by the assumptions on which a model is built. This presents a unique challenge for application of QSPR models in large databases.

This study has evaluated seven QSPR based models that have used common descriptors and common assumptions and are used extensively by in silico tools to select the equation best suitable for screening the phytochemical library of *Camellia sinensis*.

Though more famous as a beverage, tea has been used topically since ancient times. Traditional tea baths were famous in ancient China for skin and hair and were also used in detoxification. *Camellia sinensis* and its compounds are used as topically applied cosmetics and have been reviewed in detail. Black tea dressing has been reported to be beneficial in facial dermatitis. Topical formulations of green and black tea have been reported to be effective against skin cancer, UV-induced damage, wound healing, as well as have anti-aging effects. The antibacterial effect of caffeine has been clinically tested in psoriasis. The benefits of tea, not only as a beverage but also as a topical remedy and cosmetic, warrant assessment of the skin permeability profile of its compounds library. The present study is designed to evaluate tea compounds’ skin permeability by using well-documented mathematical models based on a common set of limited and easily obtainable parameters, compare them, and select the best suitable one based on the standard methodology.

**MATERIALS AND METHODS**

Construction of Phytochemical library of *Camellia sinensis* and Preparation of Ligands for Qikprop Analysis

Tea phytochemicals were enlisted after extensive literature searching. Beside this, databases such as IMMPAT (https://cb.imsc.res.in/imppat/), Dr. Duke’s Phytochemical and Ethnobotanical database (https://phytochem.nal.usda.gov/) and PubChem (https://pubchem.ncbi.nlm.nih.gov/) were also utilized to screen out a phytochemical library of 693 compounds (Figure 1). Their SDF structures were downloaded from PubChem and subsequently prepared using the LigPrep module of Schrödinger suite version 2020-3, considering all possible stereochemical, ionization and tautomeric variations using the OPLS3e force field. All the possible ionization and tautomeric states between pH 6.8 and 7.2 were generated using Epik module and the optimized ligands were used for QikProp studies.

**Prediction of Skin Permeability using Different Mathematical Models**

Skin permeability coefficient (logK_p) was assessed to investigate cutaneous absorption of tea compounds. The QSPR equations for prediction of logK_p Table 1 which use either or both octanol-water partition coefficient (logK_ow) and molar mass (MW), two easily obtainable descriptors, were selected for this study. Log transformation of predicted K_p values of Mitragotri and Hatanaka equation was carried out for common scaling of data. The unit conversion, from cm/hr to cm/sec, was carried out in case of Vecchia and Bunge, QikProp and DERMWIN™ predicted logK_p values. The predicted skin permeation coefficient (logK_p) using different mathematical models and skin permeation coefficient (QPlogK_p) yielded by QikProp (referred as QikProp predicted logK_p value hereafter) were used for further analysis.

For Mitragotri’s equation, the second factor molecular radius (r), was calculated using MW, as mentioned by Lian et al. (22) as, r = 0.91 × MW.

**Selection of Mathematical equation for Prediction of logK_p**

Due to the lack of any agreeable gold standard method for the determination of logK_p value, and a pairwise analysis of reliability and agreement between the selected methods
Selection of QSPR model for screening tea compounds

Table 1: Selected mathematical QSPR models and their equations

<table>
<thead>
<tr>
<th>QSPR based models</th>
<th>Equations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hatanaka et al. (^{18})</td>
<td>(K_p = 4.78 \times 10^{-7} \times K_{ow}^{0.509} + 8.33 \times 10^{-8})</td>
</tr>
<tr>
<td>Potts and Guy (^{19})</td>
<td>(\log K_p = -6.3 + 0.71 \times \log K_{ow} - 0.0061 \times MW)</td>
</tr>
<tr>
<td>Mitragotri (^{20})</td>
<td>(K_p = 5.6 \times 10^{-6} \times K_{ow}^{(-0.467)})</td>
</tr>
<tr>
<td>Vecchia and Bunge (^{21})</td>
<td>(\log K_p = -2.44 + 0.514 \times \log K_{ow} - 0.005 \times MW)</td>
</tr>
<tr>
<td>Lian et al. (^{22})</td>
<td>(\log K_p = -5.2 + 0.7 \times \log K_{ow} - 0.072 \times MW^{0.75})</td>
</tr>
<tr>
<td>DERMWIN (^{TM23})</td>
<td>(\log K_p = -2.8 + 0.66 \times \log K_{ow} - 0.0056 \times MW)</td>
</tr>
</tbody>
</table>

Where, \(K_{ow}\): octanol/water partition coefficient; \(MW\): the molecular weight of the compound; \(r\): molecular radius of the compound. DERMWIN is a module for estimation of skin permeability coefficient in the EPI Suite (https://www.epa.gov/tsca-screening-tools/epi-suitetm-estimation-program-interface).

The prepared phytochemical library of *Camellia sinensis* comprises 693 compounds belonging to different chemical classes, with molecular weight ranging between 31 to 1322 Da and \(K_{ow}\) between -7 to +18. The tea phytochemicals library, classified as per Cumming and Rucker., 2017 \(^{33}\) was found to be consisting of 41 (5.92%) extremely hydrophilic phytochemicals (\(K_{ow} \leq -3\)), 13 (1.88%) extremely hydrophobic phytochemicals (\(K_{ow} \geq +10\)) and remaining 639 (92.2%) compounds with \(K_{ow}\) values ranging between -3 to +10.

Using the selected equations, the \(\log K_p\) values were then calculated from the QikProp generated molecular weight and \(Kow\). The distribution of the predicted \(\log K_p\) values of the tea compounds obtained from different mathematical equations, QikProp and the average \(\log K_p\) were depicted in Figure 2. Except for the equations given by Hatanaka et al. and DERMWIN \(^{TM}\), all the predicted \(\log K_p\) values showed similar distribution patterns.

The mean, standard deviation and 95% confidence interval of the mean of predicted \(\log K_p\) were tabulated in Table 2. The average \(\log K_p\) value was calculated as the arithmetic mean of the predicted \(\log K_p\) values of the seven QSPR methods and used as the gold standard for further comparison. As the Bland and Altman plot analysis was conducted and the 95% limit of agreement (mean ± 1.96 x standard deviation ranges of the respective differences) between predicted \(\log K_p\) values compared to the gold standard value was evaluated. \(^{28,31}\) The coefficient of variation (%CV), a measure of relative dispersion, of the difference between the QSPR predicted \(\log K_p\) values and standard \(\log K_p\) value was also estimated. The reliability of the predicted \(\log K_p\) values was estimated using the intraclass correlation coefficient (ICC) (2-way mixed-effect model), in term of both consistency and absolute agreement. \(^{32}\) A two-tail \(p<0.05\) was considered as statistically significant. All the analyses were carried out using MS Excel (version 2019) and statistical program packages OriginPro 2021b, OriginLab Corporation, Northampton, MA, USA.

**Result**

The statistical analysis of the results is presented in Table 2. The predicted \(\log K_p\) values were summarized as mean, standard deviation and 95% confidence interval of mean (95% CI). The accuracy of the prediction of the gold standard \(\log K_p\) value was estimated using the coefficient of determination \((R^2)\) and root mean square error (RMSE), a residual statistic, which gives a good idea of both bias and spread of the data. Finally, the statistical analyses will result in a cumbersome output (21 pairs of different statistical analyses) which will be hard to interpret, thus, there is a need of a consensus standard for comparison. Also, the experimental methods do suffer from inherent limitations (in this case, pH, temperature, skin type etc.) and as there is a lack of experimental data of such a large chemical set, we have to look for other feasible options for better understanding and interpretation of the results.
Selection of QSPR model for screening tea compounds

The consensus gold standard average logK<sub>p</sub> value showed significantly lower logK<sub>p</sub> compared to all other methods (p<0.001). The predicted logK<sub>p</sub> value obtained in case of Hatanaka et al. equation was found to be significantly higher compared to all other methods (p<0.001) followed by the equation given by Vecchia and Bunge (p=0.132–<0.001), whereas DERMWIN™ yields significantly lower logK<sub>p</sub> compared to all other methods (p<0.001).

The scatter plot analysis of predicted logK<sub>p</sub> values for different pair of equations (not shown) revealed that they have monotonic association, i.e. the data points are moving in the same relative direction but not with a constant rate (as in case of linear relation). Hence Spearman rho (ρ) correlation analysis was conducted to evaluate the association between the predicted logK<sub>p</sub> values obtained from different equations and the selected gold standard logK<sub>p</sub> value (Figure 3). All predicted logK<sub>p</sub> values showed significant positive association (Spearman ρ≥0.658, p<0.001) with one another. Potts and Guy, Mitragotri, Lian et al. and Vecchia and Bunge predicted logK<sub>p</sub> values had ρ values of 0.996 and above with each other and the consensus average logK<sub>p</sub> value, whereas DERMWIN™ predicted logK<sub>p</sub> value showed strong correlation with all the predicted logK<sub>p</sub> values (ρ=0.878, p<0.001) with all others. The consensus gold standard selected for this study, i.e. the average logK<sub>p</sub> value showed strong correlation with all the predicted logK<sub>p</sub> values (p=0.878-0.998, p<0.001).

From the R<sup>2</sup> values for standard logK<sub>p</sub>, as predicted from the QSPR equations, it was evident that the average logK<sub>p</sub> value could be explained well by all the predictors (Table 3), which was in accordance to the correlation data. The lowest RMSE value for predicted average logK<sub>p</sub> was observed from the equation of Potts and Guy (RMSE= 0.005) followed by Lian et al. (RMSE=0.007), whereas Hatanaka et al. equation showed highest RMSE value (RMSE= 0.125).

The result of correlation and regression analysis indicated that there was a significant association between the criterion (standard logK<sub>p</sub> values) and predictors (the predicted logK<sub>p</sub> values from selected equations) but failed to provide any information regarding the limit of agreement and reliability of the data compared to the gold standard method. Bland and Altman plot analysis was conducted in order to assert the agreement between the predicted logK<sub>p</sub> values from the gold standard one. The consensus gold standard average logK<sub>p</sub> value showed significant difference compared to all the methods used for the prediction of logK<sub>p</sub> values (Games-Howell post hoc analysis, p<0.001). The coefficient of variation (%CV), a measure of dispersion of dataset, of the difference between the predicted and average logK<sub>p</sub> values (represented on the y-axis of Figure 4) showed appreciable variation (21.15–90.59%) across the QSPR methods. The least dispersion was noted in case of Lian et al. (%CV=21.15%) closely followed by Potts and Guy (%CV= 21.31%). Despite the varied dispersion across the data set, the predicted logK<sub>p</sub> values showed good agreement with average logK<sub>p</sub> as indicated by the number of compounds falling between 95% limit of agreement, the predicted logK<sub>p</sub> and standard average logK<sub>p</sub> values (Table 4) ranging 86.75–97.98%. The predicted logK<sub>p</sub> values and standard logK<sub>p</sub> values showed excellent consistency.

### Table 2: Predicted logK<sub>p</sub> values of the *Camellia sinensis* phytochemical library from different equations and average logK<sub>p</sub>

<table>
<thead>
<tr>
<th>LogK&lt;sub&gt;p&lt;/sub&gt;</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>95% Confidence Interval of mean</th>
<th>p-value (Welch ANOVA)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lower Bound</td>
<td>Upper Bound</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>9.83</td>
<td>3.831</td>
<td>-10.111</td>
<td>9.539</td>
</tr>
<tr>
<td>Hatanaka et al.</td>
<td>5.42</td>
<td>1.651</td>
<td>-5.545</td>
<td>5.299</td>
</tr>
<tr>
<td>Potts and Guy</td>
<td>-8.20</td>
<td>3.686</td>
<td>-8.475</td>
<td>-7.925</td>
</tr>
<tr>
<td>Vecchia and Bunge</td>
<td>-7.63</td>
<td>2.841</td>
<td>-7.843</td>
<td>-7.419</td>
</tr>
<tr>
<td>Lian et al.</td>
<td>-8.40</td>
<td>3.646</td>
<td>-8.668</td>
<td>-8.124</td>
</tr>
<tr>
<td>DERMWIN™</td>
<td>-12.877</td>
<td>5.592</td>
<td>-13.294</td>
<td>-12.460</td>
</tr>
<tr>
<td>QikProp</td>
<td>-8.05</td>
<td>2.979</td>
<td>-8.271</td>
<td>-7.827</td>
</tr>
</tbody>
</table>

![Figure 3: Correlogram depicting the Spearman p values among the predicted logK<sub>p</sub> from different models and the gold standard logK<sub>p</sub> (average logK<sub>p</sub>) values.](Image)
Selection of QSPR model for screening tea compounds

Table 3: Accuracy of standard logKp value obtained from each formula compared to the selected gold standard logKp value

<table>
<thead>
<tr>
<th>QSPR Equations</th>
<th>Average logKp (n=693)</th>
<th>R²</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hatanaka et al. (18)</td>
<td>0.5624</td>
<td>0.125</td>
<td></td>
</tr>
<tr>
<td>Potts and Guy (19)</td>
<td>0.9930</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td>Mitragotri (20)</td>
<td>0.9953</td>
<td>0.008</td>
<td></td>
</tr>
<tr>
<td>Vecchia and Bunge (21)</td>
<td>0.9954</td>
<td>0.011</td>
<td></td>
</tr>
<tr>
<td>Lian et al. (22)</td>
<td>0.9954</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>DERMWIN™</td>
<td>0.8887</td>
<td>0.080</td>
<td></td>
</tr>
<tr>
<td>QikProp</td>
<td>0.8436</td>
<td>0.010</td>
<td></td>
</tr>
</tbody>
</table>

(iICC=0.975, mean measurement: K=8, consistency, 2-way mixed-effect model) and excellent agreement (ICC=0.929, mean measurement: K=8, absolute agreement, 2-way mixed-effect model) and was in accordance with the above results (21). Except for Hatanaka predicted logKp values (ICC=0.410), all other QSPR equation predicted logKp values showed good to excellent reliability (0.845≤ICC≤0.963), mean measurement: K=2, absolute agreement, 2-way mixed-effect model) compared to the consensus standard (Table 4).

The intraclass correlation coefficient was calculated based on mean measurement (K=2) of the used QSPR equation and the consensus standard, for absolute agreement using a 2-way random mixed-effect model.

DISCUSSION

Skin, the body’s largest organ, forms a unique and flexible interface between the body’s internal milieu and the external environment; as a potential barrier, skin protects the body from foreign compounds. Stratum corneum, the outermost permeability barrier of the skin, is made of multilayers of hydrophilic keratin filaments embedded in a hydrophobic lamellar lipid matrix. The type and the amount of lipid in the stratum corneum depends on the site of the body and, it is generally accepted that skin permeability is affected by this lipid layer.34,35 Michaels et al., 197536 showed that several drugs had significant skin permeability and determined their stratum corneum diffusion coefficients. The main limiting factor for this process is the slow diffusion through the dead layer of stratum corneum. Several investigators have used the published human stratum corneum permeability coefficient (Kp, often expressed as logKp) data to predict the skin permeability and examined the effect of the structural parameters of penetrants on the permeability19,37 which led to the development of QSPR models. QSPRs are useful in predicting the behavior of novel compounds and provide insights into mechanisms of activity. A current trend in QSPR studies is the use of theoretical molecular descriptors that can be calculated directly from molecular structure.

The descriptors used in the QSPR development are mostly measurable and easily obtainable physicochemical properties like- molecular weight (MW), melting point (MP), and logKow. The MW and logKow are often the key- and in most cases the only- descriptors in the correlation-based QSPRs, such as the Potts and Guy method (38). Though using descriptors on the basis of ease of measurement undermines important influencers, these methods prevail in practice. Most in silico tools (such as- DERMWIN, QikProp, SwissADME etc.) use Potts and Guy equation or equation based on the same dataset used by them.23,24,39

This study chose six different QSPRs consisting of only MW and logKow as descriptors, along with QikProp generated logKp values, to evaluate the most suitable QSPR to predict the skin permeability. The profile of selected seven QSPR based models is given in Table 5. Hatanaka et al. provided an equation for predicting drugs’ steady state permeation...
rate based on their model of two parallel skin permeation pathways of lipid and pore. They had mentioned that the permeability coefficient is correlated to the partition coefficient and proposed an equation based on $K_{ow}$. Potts and Guy provided a mechanistically based model, preferable for the compounds ranging between MW 18-750Da and $K_{ow}$ -3 to +6. They considered the lipid matrix as the pathway of skin permeation. Based on the analytical solutions of diffusion Mitragotri proposed a mechanistic model. In this model, four pathways were taken for consideration- free-volume diffusion through lipid bilayers, lateral diffusion along lipid bilayers, diffusion through pores, and diffusion through shunts. Mitragotri’s equation describes free volume diffusion of the lipophilic chemicals ($logK_{ow}>1$) which was found to be in perfect agreement with the Potts Guy method. Mitragotri also explained the limitation of Potts Guy QSPR mechanistically and mathematically by showing the permeation of hydrophilic solutes by diffusion through aqueous pores and giving a correction on Potts and Guy equation. Lian et al., 2008 provided a modified form of Mitragotri’s equation by substituting the solute radius of a molecule with the molecular weight for the diffusion in lipid bilayer and given the equation resemblance to the equation of Potts and Guy. Vecchia and Bunge established a model based on $logK_{ow}$ and MW, providing a simple equation for reasonably estimating the stratum corneum permeability coefficient. They had presented diverse MW ranging from 18-584Da, and $logK_{ow}$ ranging from -3.1 to +4.6 to develop the equation. Two software-based models, DERMWIN™ and QikProp taken in this study, developed their QSPRs using the dataset of Potts and Guy but they did not provide enough information in their user guidelines. There is a great discrepancy between the predicted $logK_p$ using different models as well as the superiority of a particular QSPR model basically due to their empirical nature and the experimental conditions under which data were collected. Hence, while selecting a particular method, its agreement, reliability and reproducibility with some gold standard method, is warranted. Simple association statistics, which are a measure of relationship but not the differences alone, fails to assess the comparability. In the present study, along with association statistics, Bland-Altman plot analysis was conducted for agreement and intraclass correlation coefficient, a reliability index, was estimated to select an appropriate QSPR method for the prediction of $logK_p$ values having better consistency and agreement with the gold standard $logK_p$ value. Based on the overall performance of different statistical parameters of the predicted $logK_p$ values in relation to the gold standard $logK_p$ value, the QSPR equation proposed by Potts and Guy was selected as the method of choice for the prediction of $logK_p$ value based on lowest RMSE value and %CV (dispersion of its difference from standard $logK_p$ value) associated with better overlap with other predicted $logK_p$ values, 95% limit of agreement, and comparable Spearman $p$, $R^2$, consistency and absolute agreement, for further use with the tea phytochemical library. The Potts and Guy equation predicted $logK_p$ values of the tea phytochemicals

<table>
<thead>
<tr>
<th>Sl. No.</th>
<th>Models</th>
<th>Reference range of compounds</th>
<th>Database used</th>
<th>$R^2$ value of the equation</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Hatanaka et al. (18)</td>
<td>MW:130.08 to 375.85Da; $logK_{ow}$ : -4.7 to +4.4</td>
<td>In-vitro dataset of hairless rat skin and artificial membrane (n=17)</td>
<td>Not reported</td>
<td>Could not predict the permeation of the compounds which are hydrophilic and high molecular weighted</td>
</tr>
<tr>
<td>2</td>
<td>Potts and Guy (19)</td>
<td>MW: 18 to &gt;750Da; $logK_{ow}$ : -3 to +6</td>
<td>In-vitro dataset of human epidermis (n=90) from Flynn., 1990 (37)</td>
<td>0.676</td>
<td>Under predicts the skin permeability of hydrophilic compounds</td>
</tr>
<tr>
<td>3</td>
<td>Mitragotri (20)</td>
<td>MW: 18Da to 150KDa; $logK_{ow}$ : -6.9 to +5.49</td>
<td>In-vitro dataset of mammalian epidermis (n=120) from Jhonson et al., 1997 (41)</td>
<td>0.698</td>
<td>Under predicts the skin permeability of highly hydrophilic compounds ($K_{ow}&lt;0.01$) and ignored hydrophilic pathway</td>
</tr>
<tr>
<td>4</td>
<td>Vecchia and Bunge (21)</td>
<td>MW: 18 to 584Da; $logK_{ow}$ : -3.1 to +4.6</td>
<td>In-vitro dataset of human skin (n=127) gathered from multiple dataset</td>
<td>0.551</td>
<td>More permeability coefficient data needed to decide the mechanism of permeation of hydrophilic compounds</td>
</tr>
<tr>
<td>5</td>
<td>Lian et al. (22)</td>
<td>MW: 18 to 765Da; $logK_{ow}$ : -3.7 to +5.49</td>
<td>In-vitro dataset of human skin (n=124) gathered from multiple dataset</td>
<td>0.698</td>
<td>Modification of Mitragotri’s method only</td>
</tr>
<tr>
<td>6</td>
<td>DERMWIN™</td>
<td>Not mentioned</td>
<td>Used data set from Potts and Guy., 1992 (19)</td>
<td>Not mentioned</td>
<td>Insufficient information (40)</td>
</tr>
<tr>
<td>7</td>
<td>QikProp</td>
<td>Not clearly mentioned</td>
<td>Used data set from Potts and Guy., 1992 (19)</td>
<td>0.78</td>
<td>Insufficient information</td>
</tr>
</tbody>
</table>
Figure 5: Pie-diagram depicting the percentage of tea compounds shown to have logK<sub>p</sub> above naringenin (excellent), between naringenin and naringin (good) and below naringin (poor) were then scaled on the basis of predicted logK<sub>p</sub> values of the selected skin permeable standard phytochemicals (based on the report of Chuang et al., 2017) (30)– naringenin (logK<sub>p</sub>= 6.17153; suggestive for high skin permeability boundary) and naringin (logK<sub>p</sub>= -10.1535; suggestive for poor permeability boundary) and their percutaneous administration profile were constructed (Figure 5). It was observed that, out of 693, 259 (37.38%) tea phytochemicals have a skin permeation rate higher than that of naringenin, whereas 245 (35.35%) components were found to have their skin permeability between the range of naringenin and naringin. As more than two-thirds of the tea phytochemicals possess appreciable skin permeability, as per in silico logK<sub>p</sub> predicted value, it can be considered as a potent candidate for topical application and formulation. Agreement and reliability facilitated the selection of one particularly successful model to predict logK<sub>p</sub> using the mathematical equation provided by Potts and Guy. This model transforms the easily obtainable descriptors like, molar mass and partition coefficient of a known dataset as a resource of information to a more beneficial model to replace permeation testing for a wide range of compounds with an unknown dataset. Using this model, our prepared library of tea phytochemicals was found to have 72.73% skin-permeable compounds. For the rest of the compounds with poor permeability, carriers should be used for their successful topical and transdermal application. This study is limited on the small set of QSPR based models and their validation with the data obtained from the computational study. Therefore, further study is warranted along with the experimental data set for a better understanding of these models’ performance.

REFERENCES

Selection of QSPR model for screening tea compounds


PEER-REVIEWED CERTIFICATION

During the review of this manuscript, a double-blind peer-review policy has been followed. The author(s) of this manuscript received review comments from a minimum of two peer-reviewers. Author(s) submitted revised manuscript as per the comments of the assigned reviewers. On the basis of revision(s) done by the author(s) and compliance to the Reviewers’ comments on the manuscript, Editor(s) has approved the revised manuscript for final publication.