JPP
Journal of
Pharmacy and Pharmacology

JPP 2009, 61: 1147–1153

Received December 18, 2008
Accepted June 1, 2009
DOI 10.1211/jpp/61.09.0003
ISSN 0022-3573

Research Paper

The application of Gaussian processes in the prediction of percutaneous absorption

Gary P. Moss[a], Yi Sun[b], Maria Prapopoulou[a], Neil Davey[b], Rod Adams[b], W. John Pugh[c] and Marc B. Brown[a]

[a]The School of Pharmacy, [b]The School of Engineering & Information Sciences, University of Hertfordshire, Hatfield, UK and [c]The Welsh School of Pharmacy, Cardiff University, UK

Abstract

Objectives The aim was to assess mathematically the nature of a skin permeability dataset and to determine the utility of Gaussian processes in developing a predictive model for skin permeability, comparing it with existing methods for deriving predictive models.
Methods Principal component analysis was carried out in order to determine the nature of the dataset. MatLab software was used to assess the performance of Gaussian process, single linear networks (SLN) and quantitative structure–permeability relationships (QSPRs) using a range of statistical measures.
Key findings Principal component analysis showed that the dataset is inherently non-linear. The Gaussian process model yielded a predictive model that provides a significantly more accurate estimate of skin absorption than previous models, particularly QSPRs (which were consistently worse than Gaussian process or SLN models), and does so across a wider range of molecular properties. Gaussian process models appear particularly capable of providing excellent predictions where previous studies have shown QSPRs to fail, such as where penetrants have high log P and high molecular weight.
Conclusions A non-linear approach was more appropriate than QSPRs or SLNs for the analysis of the dataset employed herein, as the prediction and confidence values in the prediction given by the Gaussian process are better than with other methods examined. Gaussian process provides a novel way of analysing skin absorption data that is substantially more accurate, statistically robust and reflective of our empirical understanding of skin absorption than the QSPR methods so far applied to skin absorption.
Keywords Gaussian process; machine learning methods; percutaneous absorption; quantitative structure–permeability relationships

Introduction

In the new age of combinatorial chemistry and high throughput screening there are often thousands of potential drug candidates to select from in order to choose one potential drug for pharmaceutical development. The first part of this selection process frequently depends on the preferred route of delivery but often involves some form of in-silico prediction of bioavailability and efficacy. As such, the predictive modelling of percutaneous absorption has been a subject of significant interest and debate. For example, Flynn compiled and published his dataset, the findings of which are summarised in Table 1.[1] This allowed, for the first time, semi-quantitative estimates of skin absorption to be proposed, and indicated that lipophilicity (log P) and molecular weight (MW) were the main determinants of percutaneous absorption. Subsequently, Potts and Guy proposed a quantitative structure–permeability relationship (QSPR) based on multiple linear regression analysis:[2]

$$\log K_p = 0.71 \log P - 0.0061\, MW - 6.3 \quad (1)$$

(number of observations = 93; correlation coefficient r^2 = 0.67; standard error of the estimate and Fisher's statistic not reported), where log K_p (as cm/h) is the permeability coefficient, log P is the lipophilicity (described as log P_{KNOWN} by Potts and Guy) and MW is the molecular weight of the penetrant. This model quantified Flynn's findings into an equation that has been consistently validated, and modified, in the intervening years.[3–12] Potts and Guy[13] also investigated the role of hydrogen bonding and developed predictive

Correspondence: Gary P. Moss, The School of Pharmacy, University of Hertfordshire, College Lane, Hatfield Al10 9AB, UK.
E-mail: g.p.j.moss@herts.ac.uk

Table 1 Algorithms for calculating permeability coefficient (K_p). Adapted from Flynn 1990[1]

Range	Low MW (<150 Da)	High MW (>150 Da)
Log P < 0.5	Log K_p = –3	Log K_p = –5
0.5 ≤ Log P ≤ 3.0	Log K_p = log P – 3.5	
0.5 ≤ Log P ≤ 3.5		Log K_p = log P – 5.5
Log P > 3.0	Log K_p = –0.5	
Log P > 3.5		Log K_p = –1.5

models using the permeability coefficients of 37 non-electrolyte compounds selected from the compilation of Flynn.[1] This resulted in the development of a QSPR model without a hydrophobic term but including molecular volume and hydrogen bonding terms. The various equations discussed in these papers have been summarised elsewhere.[14] Gullick and colleagues[15] also compared the accuracy of each model with the experimentally measured permeability coefficients of a series of novel compounds. Generally, they found that the models provided a good fit to the experimental data at low log P values (normally, where log P was less than 2.0), but that increasingly lipophilic molecules did not fit well to any of the models. This is perhaps not surprising, given Flynn's qualitative estimates of skin absorption and the nature and distribution of the data used by Flynn and its basis for most QSPR-based studies of skin absorption that have followed. Nevertheless, it exposes a significant limitation to the use of QSPR-based approaches in generating a single holistic model for percutaneous absorption that is relevant to a wide range of potential penetrants.

In essence, it is understood that extremes of relevant molecular properties, such as log P and MW, will have a substantial impact on the percutaneous absorption of exogenous chemicals. If, for example, lipophilicity is too low or too high (in broad terms, a log P less than 1 or greater than 3), then absorption will in general be poor, and significantly lower than that of molecules with intermediate lipophilicities. Both ends of what is effectively a Gaussian-type distribution are not effectively modelled by the QSPRs discussed above, and as such these models fail to fully represent the apparently non-linear relationship between a molecule's physicochemical parameters and its permeability across skin. Interestingly, it has been shown that the greatest difference between experimental and predicted values was found at the highest log P values, and that the greatest inaccuracies were found with the Potts & Guy model.[14,15]

There is therefore compelling qualitative evidence suggesting that the non-linear response to skin absorption observed experimentally and that is missing from the linear QSPR equations (such as those derived from Potts and Guy style analyses) should be modelled by alternative means. While this is complicated by the interrelationship of the skin permeability coefficient, K_p, and flux (J_{max}), it suggests that non-linear modelling might provide a prediction of percutaneous absorption for molecules with a wide range of log P values with a single, relevant mathematical formula. Gaussian models are inherently non-linear and, by the application of techniques such as principal component analysis (PCA) and machine learning methods, can interrogate the fundamental nature of the datasets used to generate predictions of percutaneous absorption, thus allowing selection of the most suitable mathematical technique. The aim of this study was therefore to investigate the use of Gaussian processes (GPs) on a skin permeability dataset and to map this against key physicochemical parameters of the penetrants.

Methods

Dataset

The dataset used in this study was collated from a range of literature sources. It predominantly consisted of the Flynn dataset, used by Potts and Guy, and others, and was supplemented by a dataset collated from the Edetox database (available at www.ncl.ac.uk/edetox/index.html) and data published elsewhere.[12,16] It also contained several additions whose origins are described by Moss and colleagues.[14] The final dataset consisted of 142 compounds, each defined by six molecular descriptors and experimental values for K_p for permeation across human skin. The descriptors are melting point (MPt), MW, solubility parameter (SP – a measure of solubility in the stratum corneum, relative to the nature of this membrane), log P (experimental values from the literature as described above, or predicted values derived from the KOWWIN resource, available via http://syrres.us/esc/kowwin.htm), HA and HD, which are counts of the number of hydrogen bonding acceptor (HA) and donor groups (HD) that can be found on a molecule.

Theoretical background – single layer networks and Gaussian processes

The process starts with a set of N data items (i.e. measured or predicted permeability data), each of which is a vector, normally known as the input set (denoted as a typical vector $\mathbf{x}_n$). Each input vector has a corresponding output value, y_i. The aim in regression analysis is to model the relationship between the inputs and outputs. For example, there are 100 compounds and six attributes (physicochemical descriptors described above) associated with each of the compounds. The output is the skin permeability coefficient, K_p. In this case, therefore, the task is to infer a function that relates the input descriptors to the corresponding skin permeability, and then to predict the skin permeability coefficient for a new compound. The process of determining the model on the basis of the dataset is called, in this context, learning or training. Since the goal is to find the model with the best performance on new data, the simplest approach to the comparison of different models is to evaluate the error function using data that are independent of the data used for training. Various models are trained by minimisation of an appropriate error function defined with respect to a training data set. The performance of each algorithm should then be evaluated using an independent test set.

Three models were applied to the data: a naive model, a single layer neural (SLN) network and a GP model. GP modelling is a non-parametric method and does not produce an explicit functional representation of the data. Here it is assumed that the underlying function, $f(\mathbf{x})$, that produces the outputs will remain unknown (i.e. this represents a 'black box'

approach to modelling), but that the data are produced from a (infinite) set of functions, with a Gaussian distribution in the function space. Normally, a GP is completely characterised by its mean and covariance function. A more comprehensive discussion of the background to GPs is provided by Rasmussen and Williams.[17]

Performance measures employed to assess models of percutaneous absorption

N_{trn} and N_{tst} are given as training and test input-target pairs $(\mathbf{x}_n^{trn}, y_n^{trn})$ and $(\mathbf{x}_n^{tst}, y_n^{tst})$, respectively. Model prediction, given a test input $\mathbf{x}_n^{tst}$, is denoted by $\hat{y}_n$. Mean squared error (MSE) measures the average squared difference between model predictions $\hat{y}_n$, and the corresponding targets y_n^{tst}. Herein, the normalised mean squared error (NMSE) is reported, which is to normalise MSE by the variance of target values. Furthermore, the degree of improvement of the model over the Naive predictor can be quantified by the Improvement over Naive (ION) measure, described in equation 2:

$$ION = \frac{MSE_{naive} - MSE}{MSE_{naive}} \times 100\% \tag{2}$$

where MSE_{naive} denotes MSE of a naive model. When the results obtained from a GP are investigated, the average negative log estimated predictive density, NLL, given by equation 3 needs to be considered:

$$NLL = \frac{1}{n}\sum_{n=1}^{Ntst} -\log\, p(y_n|\mathbf{x}_n) \tag{3}$$

Where $-\log\, p(y_n|\mathbf{x}_n) = \frac{1}{2}\log(2\pi\sigma_*^2) + \frac{(y_n - E[f_*])^2}{2\sigma_*^2}$, in which case σ_*^2 is the predictive variance. Normally, in the in-vitro modelling of skin permeability (i.e. K_p), the correlation coefficient (CORR, or r^2) between targets and predictions is also employed to assess the quality of the models. This is also consistent with a criterion often employed in QSPR modelling, where reporting of the CORR is commonplace. This will readily provide contextualisation and comparison between GP and QSPR methods. Therefore, when results are analysed, the aim is to obtain a model on the test set that provides low values on both NMSE and NLL, and high values on both ION and CORR.

Computational method

Classic PCA is a projection method in which a linear transformation is used to map data to a lower dimensional space. It is one of the most popular techniques for pre-processing and visualising high-dimensional data. In practice, in order to map vectors $\mathbf{x}_n$ in a D-dimensional space onto a lower d-dimensional space (where $d < D$), the data should be pre-processed so that the mean is zero. Next the covariance matrix of the data is computed and its principal components (PCs) are obtained as the eigenvectors of this matrix. The data are projected onto the first L principal components corresponding to the largest L variances in the new set of coordinates. Compounds were plotted using the corresponding log K_p values and the first two principal components for representing the variation in the molecular descriptors of the members of the dataset. This technique was initially used to visualise the data, in order to determine whether its underlying nature was linear or non-linear.

In each experiment, the dataset was randomly divided into training and test sets 10 times. On each occasion there were 130 compounds in the training set, while the test set consisted of the remaining 12 compounds. Initially, two regression methods – SLNs and GPs – were applied using two molecular descriptors, log P and MW. These were repeated using six descriptors.

For GPs, Rasmussen and Williams' GP Toolbox[17] was applied to the dataset. The toolbox was implemented in MatLab (version 7.6), a high-level computing language used for algorithm development, numeric computation, data visualisation and analysis. A GP regression function (*gpr.m*) was employed, with a covariance function being the sum of squared exponential process and an independent noise process. The log of the hyperparameters was initialised thus: the logarithm of characteristic length-scale to 0, the logarithm of signal variance to 1.0, and noise variance to 0.1. Next, the negative log marginal was minimised with respect to the log of the hyperparameters by using the function *minimiz.m*.[17] When training compounds were included in the arguments to the *gpr.m* function they returned predictive mean values and variances as outputs. The Netlab Toolbox was used for SLNs, also implemented in MatLab. An SLN can be considered as a simple generalised linear model and is implemented using the function *glm.m*. The iterative reweighted least squares algorithm is applied to set the weights in the generalised linear model, which is implemented using the function *glmtrain.m*. After training the model, the function *glmfwd.m* is implemented to compute prediction values for the test compounds. Note that before any models are trained, the data are normalised into z-scores, by subtracting the mean and dividing by the standard deviation (SD). The test set is normalised by the mean and SD of the training set.

Results

Visualisation of skin data

Before predicting the skin permeability coefficients, the underlying data distribution was investigated by means of data visualisation technology, which projects data into a low-dimensional space. The dataset was initially visualised using PCA projections. The first principal component, $PC1$, is shown by equation 4:

$$PC1 = 0.272\, MW + 0.336\, MPt + 0.306\, SP - 0.313\, \log P + 0.567\, HA + 0.548\, HD\,, \tag{4}$$

where MW, MPt, SP, log P, HA and HD are the normalised variables. It can be seen that all six features contribute to this component, as all coefficients are relatively large. Similarly, the second principal component, $PC2$, is given by equation 5:

$$PC2 = -0.668\, MW - 0.201\, MPt + 0.396\, SP - 0.566 \log\, P - 0.173\, HA + 0.089\, HD. \tag{5}$$

The first two principal components account for 66.2% of the total variance. Figure 1 shows that there is no linear

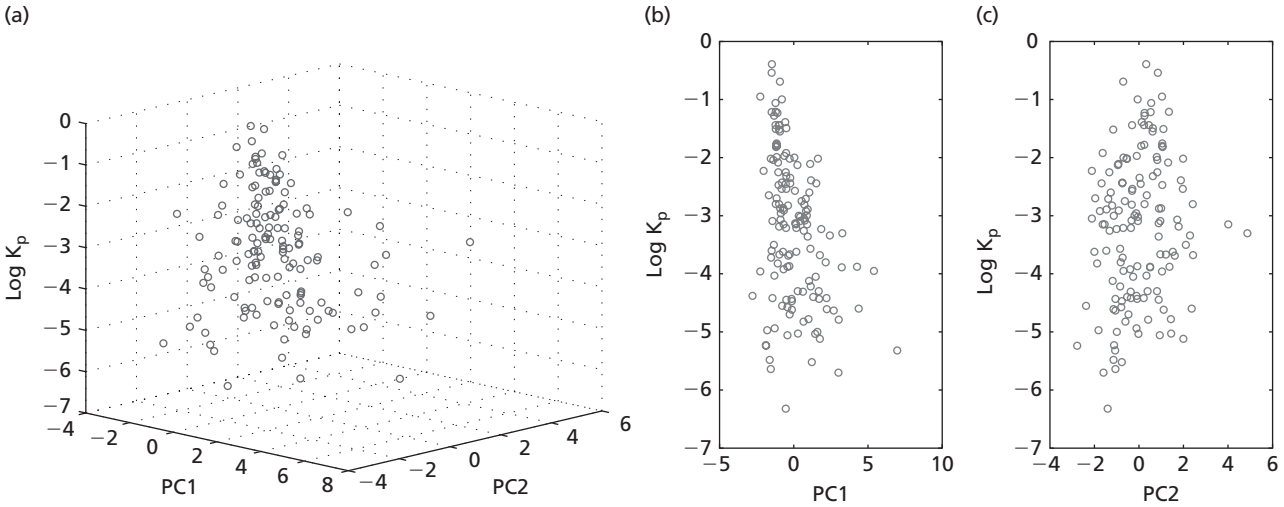

Figure 1 Initial analysis of dataset by principal component analysis, shown as (a) a three-dimensional scatter plot and (b, c) simplified two-dimensional plots which show the relationship of PC1 and PC2 to log K_p. PC, principal component.

relation between log K_p and the compound features, suggesting that there may be more complex non-linear structures in the data. Figure 2 shows the distribution of each descriptor across the whole dataset, which again suggests a non-linear distribution in the data.

Analysis of the dataset

Before machine learning methods were employed, quantitative structure–activity relationships (QSAR)-type models (i.e. Potts & Guy[2]) were applied to the whole dataset. Predictions of log K_p, MW and log P were plotted, together with target values (Figure 3). It can be seen that there is no linear relationship among MW, log P and log K_p (as shown using stars in Figure 3), while predictions of log K_p using the QSAR models show a linear shape with MW and log P along the diagonal.

Results in Table 2 show that the naive model shows better predictivity than the QSAR.[2] Table 3 shows that the trainable SLN has better predictivity than the Potts & Guy model and the naive model, and that the GP provides better predictivity than the SLN. These results suggest that the non-linear model is significantly better at predicting permeability than the linear QSAR model examined. Table 3 also shows that the GP model with six molecular features provides an even more accurate model than the GP with two features, suggesting that the addition of four more features significantly improves the predictive capabilities of the GP model.

Tables 2 and 3 demonstrate that both SLN and GP with either two or six features produce statistically more robust models than the QSAR model, which shows significantly worse predictivity than the naive model. The machine learning methods showed that models with six features had better predictivity than those obtained using only two features, when considering NMSE, ION and CORR. As for NLL, the mean of GP with six features is slightly worse than the one obtained with two features and has a larger SD.

Discussion

Classically, mathematical models describing percutaneous absorption have been developed from linear regression models. This includes, for example, the Potts and Guy equation,[2] and models derived from it. More recent iterations of this approach have veered towards more complex descriptions of absorption, using higher power terms to describe skin penetration.[12] Generally, it has been proposed that this approach reflects several issues, particularly the ability of non-linear regression methods to readily fit to any dataset presented for analysis.[18] While many models have since been developed, the model proposed by Potts & Guy (1992),[2] was used in this study because it is still the most frequently cited – and used – model for the prediction of percutaneous absorption, and as such still provides a substantial benchmark in this field. In addition, the Potts & Guy model and the subsequent development of a QSPR model relating permeability to hydrogen bonding and molecular volume[13] was key in the selection of parameters used in this study.

However, an empirical understanding of skin absorption is necessary to place this issue into perspective. Based on experimental findings from some 30 years of research in percutaneous absorption, it is generally understood that to penetrate the skin in an appreciable concentration, an exogenous chemical should possess a combination of lipophilic and hydrophilic properties. Hence, if a molecule is too hydrophilic (i.e. log P < 1.0) or too hydrophobic (i.e. log P > 3.5) it will not penetrate the skin in large enough quantities to render it therapeutically viable in a pharmaceutical context. This, of course, should be considered in the context of the required therapeutic dose. Furthermore, molecules possessing a 'suitable' intermediate solubility profile would exhibit properties that allow partition into and across the lipophilic stratum corneum skin barrier and the hydrophilic underlying viable tissues of the epidermis and dermis.[1,19–21] Thus while representing the issue in

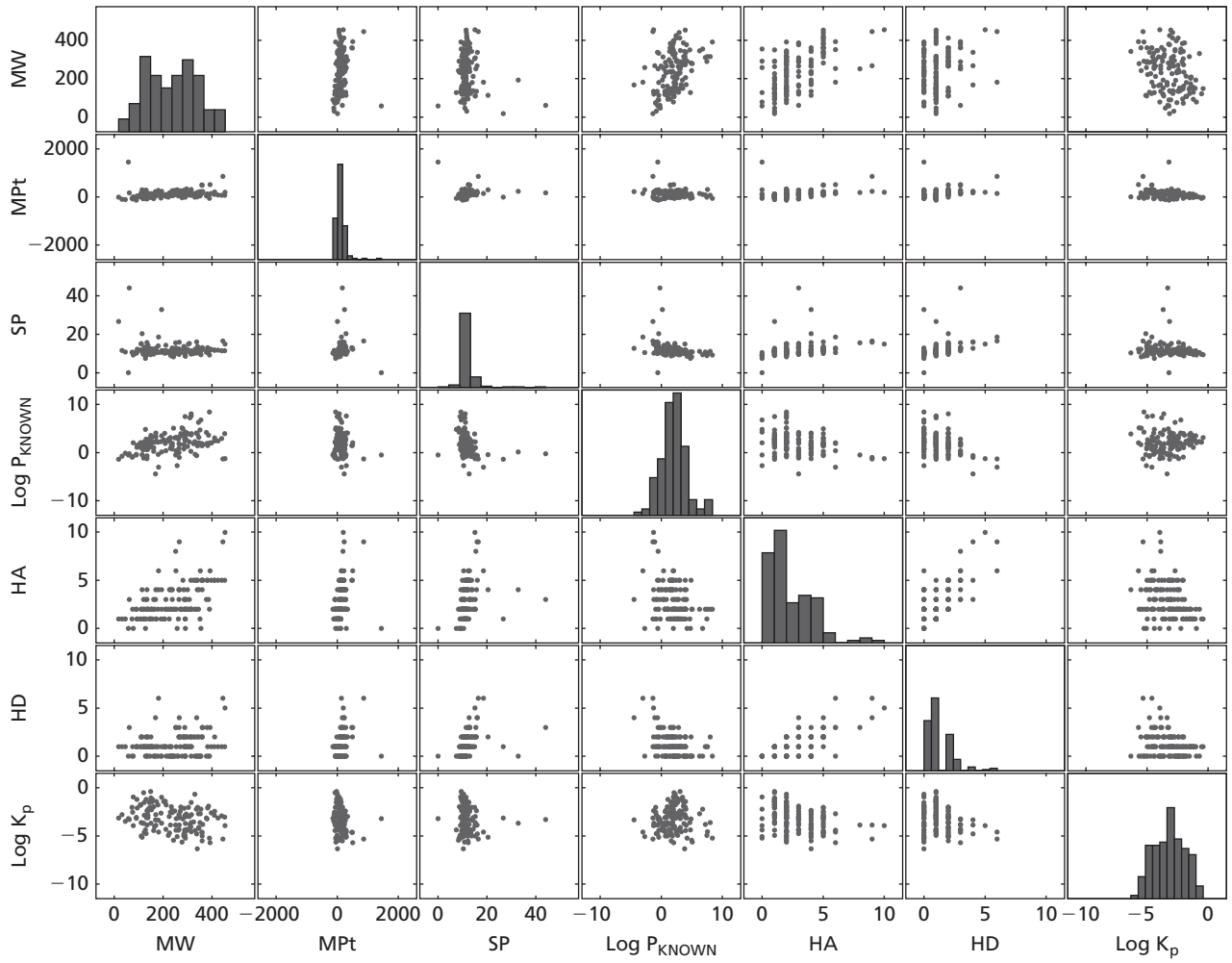

Figure 2 A scatter plot matrix of the skin dataset. The diagonal from top left to bottom right shows the shape of the distribution of each molecular descriptor examined in this study. MW, molecular weight; MPt, melting point; SP, solubility parameter, log $P_{KNOWN,}$ lipophilicity reported by Potts and Guy[2]; HA, number of hydrogen bonding acceptor groups; HD, number of hydrogen bonding donor groups; log K_p, permeability coefficient.

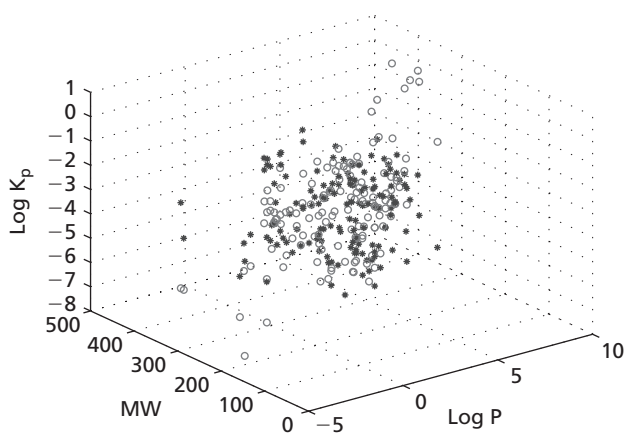

Figure 3 Comparison of predictions (circles) of permeability coefficient (log K_p), molecular weight (MW) and lipophilicity (log P) made by the Gaussian process and target (experimentally derived) values (stars).

qualitative terms, this description would suggest a bell-shaped or Gaussian curve when considering permeability as a function of key molecular descriptors such as log P. This is not surprising, as one would not usually extrapolate meaning outside the range of the dataset. For example, Gullick and colleagues[15] compared the accuracy of a range of QSPR models with experimentally determined results. Prediction at low log P was reasonable, but in every case the models failed to provide a reasonable estimate of the experimental results at high log P (i.e. > 2.5). The original model by Potts & Guy provided the least accurate fit compared with more recent models, a finding that is supported by this study (Tables 2 and 3), where the Potts and Guy model provides – in statistical terms – the least accurate model. It should also be noted that while the GP models appear to be more widely applicable to a larger dataset than, for example, the Potts and Guy (1992) model,[2] it is difficult to quantify this difference, or to establish specific ranges of applicability, because of the nature of the GP process and the use of a different dataset. Furthermore, while other QSPR models are

Table 2 Results on test sets with two (log P, MW) and six (log P, MW, number of hydrogen acceptor (HA) and hydrogen donor (HD) groups, solubility parameter, melting point) physicochemical descriptors: the naive model and quantitative structure–activity relationships (QSAR)

Inputs	NMSE	ION (%)	CORR
Naive model	1.16 ± 0.12	0	–
QSAR	1.48 ± 0.23	–35.55 ± 21.15	0.36 ± 0.08

Values are means ± SD for analysis of the whole dataset. CORR, correlation coefficient (r^2); ION, percentage improvement over naive model; NMSE, normalised mean square error.

Table 3 Results from single linear network and Gaussian process modelling employing two features (log P, MW) and six features (log P, MW, number of hydrogen acceptor (HA) and hydrogen donor (HD) groups, solubility parameter, melting point)

Inputs	NMSE	ION (%)	CORR	NLL
Single linear network				
Two features	1.02 ± 0.10	11.20 ± 5.00	0.38 ± 0.07	–
Six features	1.00 ± 0.10	11.77 ± 6.30	0.43 ± 0.08	–
Gaussian process				
Two features	0.84 ± 0.09	25.48 ± 6.82	0.53 ± 0.07	1.60 ± 0.06
Six features	0.72 ± 0.10	35.51 ± 8.30	0.59 ± 0.08	1.61 ± 0.13

Values are means ± SD for analysis of the whole dataset. CORR, correlation coefficient (r^2); ION, percentage improvement over naïve model; NMSE, normalised mean square error; NLL, average negative log estimated predictive density.

comparatively poor compared with the GP model,[8,11] they are statistically more applicable than the Potts & Guy model in terms of NMSE for example (Tables 2 and 3). However, it may be argued that such unwarranted extrapolations are exactly what have been developed with recent iterations of the Potts & Guy models.[12] The dataset from which many of these equations have been derived is essentially that used by Flynn,[1] which contains elements that have log P values predominantly below 2.2. It is possible that this distribution of data could skew the analysis or interpretation of this dataset. For example, molecules that are often discarded as outliers[11] may in fact be relevant to the analysis. Such molecules may not, however, be present in sufficient quantities in the dataset to impact statistically upon the dataset and its analysis. Therefore, a more balanced dataset that has an equal distribution of data points representing a true range of all major descriptors (e.g. MW, log P, solubility) may be required in order to develop a model with a true representation of skin absorption for as wide a range of penetrants as possible. The nature of the model also reflects the parameter being measured. For example, the use of K_p and not J_{max} is common in the models discussed, but the former clearly depends on the solvent used and the solubility of the solute therein.

Generally, it is common for a QSPR to yield an equation in a form that is readily comprehensible and, in some cases, easily applied by other researchers. However, it is also apparent that these equations are often incomprehensible and not easily accessed by other researchers, as they contain complex descriptors that require calculation by expensive software, limiting their use by workers in related fields.[11,22] The misuse of such equations is potentially an issue if researchers apply models without understanding their limitations. Gullick and colleagues showed predictions of skin permeability from a range of models, and indicated that conventional QSPR models clearly failed to provide accurate predictions of experimental results at higher log P values.[15] This may reflect a limitation of the dataset – and its inherent skew towards hydrophilic molecules – or a misapplication of those models. It may also suggest that, as the relationship between skin permeability and molecular properties is non-linear, the use of non-linear models is not only valid but wholly appropriate, reflecting the true nature of skin absorption for a wide range of exogenous penetrants. This is clearly reflected in the findings of the current study, where PCA – and visualisation of the dataset, shown in Figures 1 and 3 – clearly shows the underlying non-linear nature of the dataset.

However, the QSPR approach normally yields a very good fit to the data, with some studies returning r^2 values in excess of 0.8.[11,14,16] This may be due to the nature of the dataset being analysed. For example, if the Flynn dataset is mostly comprised of molecules with a log P of predominantly less than 2.5, then it – and any models derived from it – will not fully represent the whole nature of skin absorption, and as such any models so derived will be substantially limited. In essence, it might be argued that the clustering of data observed in this study and others (e.g. Patel and colleagues[16]) completes only one slope of a traditional Gaussian plot, and that a substantially expanded dataset, as used in this study, will improve the predictive accuracy of permeability for lipophilic penetrants.

By comparison, the Gaussian model employed herein essentially acts as a 'black box' into which data are input and out of which emerges the prediction of permeability. While this limits transparency of the model, misuse is not possible in such approaches because the parameters input to the model can be carefully screened and controlled prior to use, ensuring that the output is generated only from data within the range of the model. Given that the model can be easily and rapidly improved by the addition of new data points, the GP can effectively learn every time additions are made to the dataset. Furthermore, the model developed in this study is based on a wider dataset than previous studies, collated from a wide range of sources and covering a broader range of penetrants; as such, the issue of skewed data may not be as applicable to this dataset as with previous datasets.

The six parameters used to model the data in this study (log P, MW, MPt, HA, HD and SP) were chosen because they have consistently been found to be of significance in studies of skin permeability, both qualitatively and quantitatively. While it is a limitation of this study that a wider range of molecular descriptors (e.g. as used by Cronin and colleagues[8] or Patel and colleagues[16]) has not been used, it is also an advantage with regard to the portability and ease of use of the model by a range of researchers. The parameters used are readily determined and do not require specialist software to calculate them.

Conclusions

The GP model – utilising an expanded dataset – has yielded a predictive model that provides a significantly more accurate estimate of skin absorption than previous models, and does so across a wider range of molecular properties (particularly MW and log P). It is particularly capable of providing excellent predictions where previous studies have shown QSPRs to fail: at high log P and MW of penetrants. The implications of these findings are particularly significant in determining *how* skin absorption should be modelled; specifically, this infers that, statistically, a non-linear approach is more appropriate for analysis of the dataset employed herein than linear techniques (e.g. QSPR methods), as the GP method gives a prediction and a confidence value in the prediction that is better than other methods. Fundamentally, this work provides a novel way of analysing skin absorption data that is substantially more accurate, and reflective of our empirical understanding of skin absorption, than the QSAR/QSPR methods so far applied to skin absorption. In analysing the patterns in the data, and in the distribution of data, it is apparent that these data are fundamentally non-linear in nature. This is in agreement with many experimental studies that show a decrease in percutaneous absorption at high log P values. It also indicates that GP methods may be more directly applicable to the prediction of percutaneous absorption than the widely used QSPR-based methods. The veracity and consistency of the skin dataset remains a key issue however. Nevertheless such a model, in combination with efficacy and preformulation data, may allow the more accurate prediction and selection of lead molecules for development of topical formulations.

Declarations

Conflict of interest

The Author(s) declare(s) that they have no conflicts of interest to disclose.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

References

1. Flynn GL. Physicochemical determinants of skin absorption. In: Gerrity TR, Henry CJ, eds. *Principles of Route-to-route Extrapolation for Risk Assessment*. New York: Elsevier, 1990: 93–127.
2. Potts RO, Guy RH. A predictive algorithm for skin permeability: the effects of molecular size and hydrogen bond activity. *Pharm Res* 1995; 12: 1628–1633.
3. Abraham MH *et al*. The factors that influence skin penetration of solutes. *J Pharm Pharmacol* 1995; 47: 8–16.
4. Abraham MH *et al*. Hydrogen bonding part 46. A review of the correlation and prediction of transport properties by an LFER method: physicochemical properties, brain penetration and skin permeability. *Pesticide Sci* 1999; 55: 78–88.
5. Abraham MH *et al*. Algorithms for skin permeability using hydrogen bond descriptors: the problem of steroids. *J Pharm Pharmacol* 1997; 49: 858–865.
6. Barratt MD. Quantitative structure–activity relationships for skin permeability. *Toxicol in Vitro* 1995; 9: 27–37.
7. Brown SL, Rossi JE. A simple method for estimating dermal absorption of chemicals in water. *Chemosphere* 1989; 19: 1989–2001.
8. Cronin MTD *et al*. Investigation of the mechanism of flux across human skin *in vitro* by quantitative structure-permeability relationships. *Eur J Pharm Sci* 1999; 7: 325–330.
9. Fiserova-Bergerova V *et al*. Dermal absorption potential of industrial chemicals: criteria for skin notation. *Am J Indust Med* 1990; 17: 617–635.
10. McKone TE, Howd RA. Estimating dermal uptake of non-ionic organic chemicals from water and soil: I Unified fugacity-based models for risk assessment. *Risk Anal* 1992; 12: 543–557.
11. Moss GP, Cronin MTD. Quantitative structure-permeability relationships for percutaneous absorption: re-analysis of steroid data. *Int J Pharm* 2002; 238: 105–109.
12. Wilschut A *et al*. Estimating skin permeation – the validation of 5 mathematical skin permeation models. *Chemosphere* 1995; 30: 1275–1296.
13. Potts RO, Guy RH. Predicting skin permeability. *Pharm Res* 1992; 9: 663–669.
14. Moss GP *et al*. Design, synthesis and characterization of captopril prodrugs for enhanced percutaneous absorption. *J Pharm Pharmacol* 2006; 58: 167–177.
15. Gullick DR *et al*. Variation of predicted permeability using calculated and experimental log P values: a comparison of quantitative models for skin absorption. In: Brain KR, Walters KA, eds. *Perspectives in Percutaneous Penetration, 10th International Conference*. France: La Grande Motte, 2008: 86.
16. Patel H *et al*. Quantitative structure-activity relationships (QSARs) for the prediction of skin permeation of exogenous chemicals. *Chemosphere* 2002; 48: 603–613.
17. Rasmussen CE, Williams CKI. *Gaussian Processes for Machine Learning*. Cambridge MA, USA: The MIT Press, 2006.
18. Moss GP *et al*. Quantitative structure-permeability relationships (QSPRs) for percutaneous absorption. *Toxicol In Vitro* 2002; 16: 299–317.
19. Roberts MS *et al*. Epidermal permeability–penetrant structure relationships: 2. The effect of H-bonding groups in penetrants on their diffusion through the stratum corneum. *Int J Pharm* 1996; 132: 23–32.
20. Roberts MS *et al*. Epidermal permeability–penetrant structure relationships: 1. An analysis of methods of predicting penetration of monofunctional solutes from aqueous solutions. *Int J Pharm* 1995; 126: 219–233.
21. Pugh WJ *et al*. Epidermal permeability–penetrant structure relationships: 3. The effect of hydrogen bonding interactions and molecular size on diffusion across the stratum corneum. *Int J Pharm* 1996; 138: 149–165.
22. Cronin MTD *et al*. An investigation of the mechanism of flux across polydimethylsiloxane membranes by use of quantitative structure-permeability relationships. *J Pharm Pharmacol* 1998; 50: 143–152.