

OPEN
ANALYSIS


Predicting skin permeability using HuskinDB

Laura J. Waters   & Xin Ling Quah

A freely accessible database has recently been released that provides measurements available in the literature on human skin permeation data, known as the 'Human Skin Database – HuskinDB'. Although this database is extremely useful for sourcing permeation data to help with toxicity and efficacy determination, it cannot be beneficial when wishing to consider unlisted, or novel compounds. This study undertakes analysis of the data from within HuskinDB to create a model that predicts permeation for any compound (within the range of properties used to create the model). Using permeability coefficient (K_p) data from within this resource, several models were established for K_p values for compounds of interest by varying the experimental parameters chosen and using standard physicochemical data. Multiple regression analysis facilitated creation of one particularly successful model to predict K_p through human skin based only on three chemical properties. The model transforms the dataset from simply a resource of information to a more beneficial model that can be used to replace permeation testing for a wide range of compounds.

Introduction

Permeation of a compound through human skin is an increasingly important delivery route in pharmaceutical applications as well as being a vital property to consider in risk assessment analysis for any compound that may come in to contact with skin. *In vivo* analysis can be a good predictor of properties, such as bioavailability^{1,2}, where human volunteers are used for stratum corneum (SC) sampling but often *in vitro* techniques are used instead. It is apparent that discrepancies often occur between the *in vivo* reality and *in vitro* based predicted data. One reason for this is the variety of experimental data, in terms of skin origin or age and experimental conditions. Even when SC sampling is used the situation can be complex³. In the vast majority of such published studies experimental data facilitates calculation of the permeability coefficient (K_p) whereby a compound permeates through skin under steady-state conditions from an aqueous vehicle. To minimise potential variability in data from the source of skin, alternatives have been investigated including work in our group using poly(dimethylsiloxane) as a skin replacement^{4–6} and alternative analytical techniques including micellar liquid chromatography⁷. However, all of these experimental systems require extensive laboratory work to allow measurement of K_p . To avoid this time-consuming process, predictive models are frequently used for modelling in drug development, such as for pharmacokinetic applications⁸ and of relevance to this work, for skin permeability. These models utilise the calculated K_p value to formulate a quantitative structure-permeability relationship (QSPR) that relates permeability to identified physicochemical properties of the compound undergoing permeation, such as hydrogen bond activity⁹. With over thirty years of research in this area there are numerous datasets available for skin absorption modelling studies, ranging in many parameters including the number of compounds considered, the source and thickness of skin used, experimental temperature, pH and vehicle composition^{10–13}. Prior to this study, the vast majority of data used in creating permeation models is based on animal skin studies which are renowned for being poor mimics for human skin, a phenomenon known as the inter-species translational gap^{14,15}. One study in particular found that only four of the thirty three QSPRs available at the time of publication were deemed 'acceptable' according to their stated four criteria with only one of these providing 'reasonable' predictions¹⁶. Since then, many QSPRs have been created varying in complexity with a particularly well-known model published in 1992 by Potts and Guy¹⁷ using a compounds partition coefficient ($\log P$) and molecular weight (MW) to create the following (relatively simple) Eq. (1):

School of Applied Sciences, University of Huddersfield, Queensgate, Huddersfield, HD1 3DH, UK.  e-mail: l.waters@hud.ac.uk

$$\log K_p = 0.71 \log P - 0.0061 MW - 6.3$$

$$n = 93, r^2 = 0.67 \quad (1)$$

Other, often more elaborate models have incorporated additional physicochemical properties to predict K_p , including hydrogen bond acceptor/donor activity, solubility, charge, melting point, polarisability and vehicle formulation¹⁸. However, the variability in experimental conditions undertaken to create these datasets has been noted, such as if the solute is analysed under finite or infinite dosing conditions¹⁹ and the chosen experimental volume²⁰. These effects can result in variability in dataset K_p values which will ultimately affect the suitability and success of the model created^{21–23}. It has therefore been stated that ideally values should be obtained under the same experimental conditions²⁴ yet this is difficult to achieve in reality. Over the years many datasets have been employed to predict percutaneous absorption and in some cases combined to create even more detailed datasets²⁵.

In 2020, a new dataset was created (HuskinDB²⁶) that removed the uncertainty associated with previous models which was revolutionary in its approach²⁷ and, since a recent update, has been expanded even further. The freely accessible database now lists skin permeation values (K_p values) for 253 compounds analysed with human skin yet it also includes experimental parameters such as skin source site, skin layer used, preparation technique, storage conditions along with experimental conditions such as temperature of the acceptor and donor solutions, pH and solution compositions (where available) for each K_p value quoted. Since its release this database has been used by regulatory agencies for dermal risk assessment²⁸ and is becoming known as a useful resource to researchers²⁹.

This study utilises this in-depth dataset to create unique models that take in to account this additional information, highlighting the influence of experimental parameters on data analysis and leading to a highly specific and optimised model for use when investigating new compounds for permeability.

Results and Discussion

With 550 K_p values in total, HuskinDB is a significant source of data for those wishing to know the extent of skin permeation for any of the 253 compounds included. It has many benefits to those using the dataset, firstly that all included data was obtained using human skin thus variability is limited compared with other datasets that have included animal and other non-human membranes in the analysis. Furthermore, a variety of experimental parameters are included for each K_p value allowing the researcher to obtain a specific K_p value under whatever specific conditions are of interest, such as temperature and donor concentration. However, this work takes the dataset much further and utilises the data to create models that then permit *prediction* of K_p for other compounds of interest rather than limited to only those in the dataset. This is particularly beneficial for several applications, including when considering compounds that currently exist (but are not already in the dataset) or, have not yet even been synthesised.

The 27 scenarios where data was available from the dataset were each analysed to create a QSPR model and are listed in Table 1.

With respect to skin source, i.e. anatomical site, it is known that the source can affect permeation³⁰. Only five scenarios were analysed using breast skin and six with thigh skin, leaving the majority (sixteen) using abdomen skin. This is as expected as skin from the abdomen is frequently used in analysis for convenience reasons³¹. For permeation analysis, skin can be separated into layers to allow researchers to focus on permeation through only the epidermis or dermis, both epidermis and dermis combined or the stratum corneum. Nine of the scenarios analysed the epidermis only, three the dermis only, thirteen the epidermis and dermis and two the stratum corneum. It could be argued that permeation through the stratum corneum is the most important layer to consider as this is the first stage of the process and will therefore dictate subsequent permeation. However, as permeation must also be achieved through the entire epidermis and then dermis it is also arguable that analysis should consider both layers combined, as was the case for the majority of the scenarios. With respect to donor concentration, twenty two of the scenarios involved a diluted solute concentration in the donor phase with the remaining five as neat (saturated) solutions. This finding is particularly interesting as it is more usual in permeation analysis to apply saturated solutions to the skin to maintain sink conditions throughout the experiment³². Finally, experimental donor solution temperature was particularly variable throughout the dataset thus a decision was made to divide the experimental data into four options to simplify analysis. Results appeared equally split in that seven scenarios involved an experimental donor solution temperature between 20 and 25 °C, six between 26 and 30 °C, seven between 31 and 35 °C, with seven between 36 and 40 °C. This finding was surprising if the data entered in the dataset was acquired for *in vivo* prediction as the surface of skin is usually approximately 33 °C, and internal body temperature 37 °C^{33,34}. Therefore, the *in vivo* permeation process will occur between these temperatures and the latter two temperature options of the four listed would be the most suitable choices rather than the two lower temperature options.

Originally, 96 scenarios were considered using the four variables discussed yet a lack of data (where no compounds fit the criteria) for 69 scenarios reduced the number of models created to 27. Of these 27 remaining scenarios, 19 had a limited number of compounds ($n = \leq 15$) which was deemed too low for consideration as a suitable QSPR model. The eight remaining scenarios therefore contained 16 or more compounds with a maximum number of 45 compounds.

Along with ensuring a suitable number of compounds had K_p values available to create the QSPR model, the coefficient of determination (R^2) was an important factor for consideration with a value approaching 1 sought. This concept, whereby the value is as close to 1 as possible, has often been the focus of discussions surrounding the suitability of models for permeability prediction. Although absolute limits on what can be classed as an

Skin Source	Skin Type	Donor Conc.	Exp. Temp (°C)	No. of cmpds	R ²	Equation
Breast	Epidermis	Diluted	36–40	9	0.9839	LogK _p = -2.869 - 0.406 LogP - 0.143 TPSA + 0.021 MV
Breast	Epidermis + Dermis	Saturated	36–40	6	0.8321	LogK _p = -3.083 - 0.728 LogP - 0.168 TPSA + 0.043 MV
Breast	Epidermis + Dermis	Diluted	20–25	4	0.9703	LogK _p = -4.993 - 0.808 LogP - 0.168 TPSA + 0.033 MV
Breast	Epidermis + Dermis	Diluted	31–35	20	0.5464	LogK _p = -4.406 - 0.514 LogP - 0.004 TPSA - 0.006 MV
Breast	Epidermis + Dermis	Diluted	36–40	5	0.8747	LogK _p = -6.297 - 3.151 LogP - 0.424 TPSA + 0.128 MV
Abdomen	Epidermis	Saturated	20–25	10	0.9595	LogK _p = -5.203 + 7.125 LogP + 0.525 TPSA - 0.236 MV
Abdomen	Epidermis	Saturated	26–30	8	0.9772	LogK _p = -6.875 + 1.410 LogP + 0.206 TPSA - 0.059 MV
Abdomen	Epidermis	Diluted	20–25	36	0.8545	LogK _p = -6.052 + 0.777 LogP - 0.004 TPSA - 0.008 MV
Abdomen	Epidermis	Diluted	26–30	2	N/A	LogK _p = -6.477 + 1.467 LogP + 0.220 TPSA - 0.051 MV
Abdomen	Epidermis	Diluted	31–35	36	0.7102	LogK _p = -5.788 + 0.099 LogP - 0.028 TPSA + 0.008 MV
Abdomen	Epidermis	Diluted	36–40	43	0.1619	LogK _p = -6.231 + 0.211 LogP - 0.006 TPSA - 0.001 MV
Abdomen	Dermis	Saturated	20–25	8	0.9737	LogK _p = -6.277 + 1.778 LogP + 0.226 TPSA - 0.077 MV
Abdomen	Dermis	Diluted	20–25	16	0.8334	LogK _p = -4.584 + 0.931 LogP + 0.019 TPSA - 0.020 MV
Abdomen	Dermis	Diluted	31–35	6	0.9393	LogK _p = -1.784 + 2.499 LogP + 0.040 TPSA - 0.081 MV
Abdomen	Epidermis + Dermis	Saturated	26–30	4	0.5268	LogK _p = -7.599 - 1.743 LogP - 0.113 TPSA + 0.045 MV
Abdomen	Epidermis + Dermis	Diluted	20–25	4	0.9916	LogK _p = -2.586 + 2.453 LogP - 0.082 TPSA - 0.053 MV
Abdomen	Epidermis + Dermis	Diluted	26–30	8	0.8166	LogK _p = -5.361 + 0.400 LogP - 0.025 TPSA - 0.001 MV
Abdomen	Epidermis + Dermis	Diluted	31–35	45	0.1422	LogK _p = -6.129 + 0.193 LogP + 0.004 TPSA - 0.006 MV
Abdomen	Epidermis + Dermis	Diluted	36–40	14	0.9734	LogK _p = -6.960 + 0.975 LogP - 0.004 TPSA - 0.005 MV
Abdomen	Stratum corneum	Diluted	26–30	3	1	LogK _p = -5.695 - 1.147 LogP - 0.137 TPSA + 0.047 MV
Abdomen	Stratum corneum	Diluted	31–35	3	1	LogK _p = -6.009 - 0.043 LogP - 0.036 TPSA + 0.008 MV
Thigh	Epidermis	Diluted	31–35	3	1	LogK _p = -13.211 - 1.366 LogP + 0.031 TPSA + 0.031 MV
Thigh	Epidermis	Diluted	36–40	3	1	LogK _p = -11.084 - 1.472 LogP - 0.144 TPSA + 0.043 MV
Thigh	Epidermis + Dermis	Diluted	20–25	5	0.9344	LogK _p = 2.734 + 3.503 LogP + 0.007 TPSA - 0.088 MV
Thigh	Epidermis + Dermis	Diluted	26–30	17	0.6139	LogK _p = -6.826 + 0.656 LogP + 0.006 TPSA - 0.006 MV
Thigh	Epidermis + Dermis	Diluted	31–35	3	1	LogK _p = -4.001 + 0.309 LogP - 0.085 TPSA + 0.031 MV
Thigh	Epidermis + Dermis	Diluted	36–40	21	0.3329	LogK _p = -6.876 + 0.566 LogP - 0.002 TPSA - 0.002 MV

Table 1. QSPR models for skin permeability (K_p) coefficient prediction using data extracted from HuskinDB, as related to a compound's partition coefficient ($\log P$), topological surface area (TPSA) and molecular volume (MV).

'acceptable threshold' do not exist, researchers have previously described values below 0.3 as poor¹⁶, around 0.6 as significant³² and above 0.8 as good²⁵. Values in this study for R^2 ranged from 0.1422 (i.e. very little correlation) up to 0.8545 (i.e. an acceptable correlation). An ideal model would combine the greatest number of compounds possible with the highest R^2 value yet in reality this is not always possible. As a consequence, a compromise between these two factors was applied and the most suitable model from those available deemed to be that which included 36 compounds with an R^2 value of 0.8545. For further confirmation of the performance of this model, the total dataset ($n = 36$) was subdivided into two groups: a training set ($n = 29$) and a test set ($n = 7$) with the latter chosen at random then checked to ensure it included a range of $\log P$, TPSA and MV values. Equation (2) displays the equation created as a result of this process with the training and test set coefficients of determination (R^2) and root mean square error values (RMSE) specified.

$$\log K_p = -6.136 + 0.818 \log P - 0.005 \text{TPSA} - 0.007 \text{MV}$$

$$\text{Training set: } n = 29, R^2 = 0.8428, \text{RMSE} = 0.30$$

$$\text{Test set: } n = 7, R^2 = 0.8949, \text{RMSE} = 0.35 \quad (2)$$

Interestingly, this particular scenario was not for full thickness skin but epidermis only, with a diluted donor phase and at the lowest of the four donor solution temperature ranges considered. Why this particular model achieved the best performance of all the models created is unclear at this time. However, the high level of control over skin choice, anatomical site, skin thickness, donor phase concentration and experimental temperature do prove that removing variability in data can lead to a model with high predictive ability.

Although the derived R^2 value is deemed adequate, it could be argued that the comparatively small dataset utilised may reduce the acceptability of the model for permeation prediction in a more general context. To consider an alternative approach (whereby a larger dataset was used) an additional QSPR model was created to investigate how this compares with Eq. (2). In this additional model any compound with a K_p value was included although if multiple values were available for a compound, four experimental variables were used to reduce the number to one. These were set as: abdomen site, epidermis and dermis layers, concentrated solute, experimental donor solution temperature 30–35 °C, as well as an experimental pH between 7 and 7.5. Using these criteria all 253 compounds were analysed and found to have a low coefficient of determination where $R^2 = 0.2308$. This could be improved somewhat by removing any predicted $\log K_p$ values that were more than ± 1.5 from

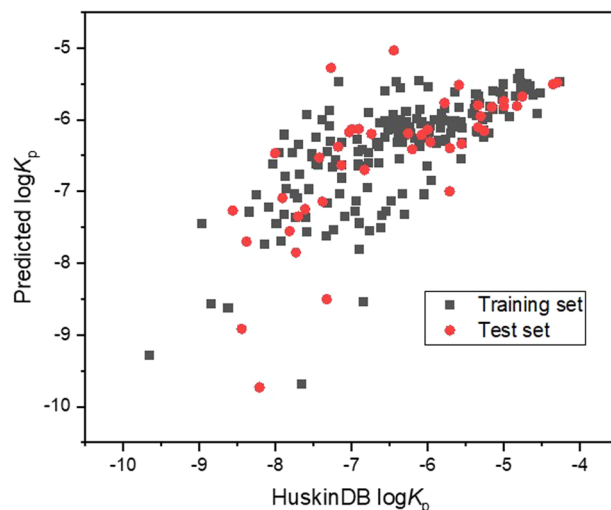


Fig. 1 Predicted (from Eq. (3)) and experimental (HuskinDB) $\log K_p$ values for the training and test sets.

the dataset value, i.e. the extreme outliers, to produce a more acceptable model (Eq. (3)) with 214 compounds included and a coefficient of determination of $R^2 = 0.5044$. The vast majority of the 39 compounds that were deemed ‘outliers’, and therefore removed to create Eq. (3), were at the extremities of the K_p values considered. As before, the total dataset ($n = 214$) was subdivided into two groups: a training set ($n = 171$) and a test set ($n = 43$) with the latter chosen at random and then checked to ensure it included a range of $\log P$, TPSA and MV values. Equation (3) displays the equation created as a result of this process with the training and test set coefficients of determination (R^2) and root mean square error values (RMSE) specified.

$$\log K_p = -5.820 + 0.319 \log P - 0.001 \text{TPSA} - 0.005 \text{MV}$$

$$\text{Training set: } n = 171, R^2 = 0.5042, \text{RMSE} = 0.73$$

$$\text{Test set: } n = 43, R^2 = 0.5057, \text{RMSE} = 0.84 \quad (3)$$

Figure (1) displays the relationship between the predicted and experimental $\log K_p$ values for the 214 compounds analysed using Eq. (3) based upon HuskinDB logarithmic K_p values expressed in cm/s.

Although Eq. (3) is superior in that a far larger dataset was included, the lower coefficient of determination (and higher RMSE) indicates that it would be better to use Eq. (2) rather than Eq. (3) when attempting to predict permeability coefficients.

Many models already exist for predicting skin permeation, including the DERMWIN™ model (<https://www.epa.gov/tsc-screening-tools/epi-suitetm-estimation-program-interface>). This model is based on an equation similar to that proposed by Potts and Guy¹⁷ in that it uses the same physicochemical properties to predict $\log K_p$ values and is often used by researchers for comparison with newly proposed models^{21,35,36}. For comparative analysis in this work, $\log K_p$ values for the 214 compounds were analysed using DERMWIN™ and the values obtained compared with those from HuskinDB (as selected for Eq. (3)). It was found that the coefficient of determination between these two sets of $\log K_p$ values was lower than both the training and test set values presented in Eq. (3) (0.4351 for DERMWIN™ vs. HuskinDB and 0.5042 then 0.5057 for Eq. (3) vs. HuskinDB), along with a higher RMSE (1.04 for the former and 0.73 then 0.84 for the latter). It can therefore be concluded that Eq. (3) provides a superior model when predicting $\log K_p$ values for human skin permeation compared with the DERMWIN™ model.

In summary, HuskinDB is an exciting and useful new database providing permeability data for a large range of compounds. This extensive dataset can be of even more use by creating models using the plethora of experimental information available about each K_p value. It would appear that the most successful QSPR model utilised a total of 36 compounds with four specified experimental conditions to create an *in silico* method for predicting permeation for any compound of interest. In comparison, a larger dataset can be considered with less focus on experimental variable selection to create an alternative model yet with a lower degree of correlation achievable. In both cases, this expansion of HuskinDB to allow prediction of permeation for compounds not included in the dataset is an exciting development in permeation prediction. This takes the database from being a limited resource only for included compounds to a way of predicting permeation for any compound of interest. As further experimental data becomes available in literature over the following years (with the required experimental parameters listed) then it will be possible to expand the dataset even further, thus potentially creating an even more successful model for prediction of permeation than that proposed in this study.

Methods

All K_p values included in this study were extracted from HuskinDB²⁷, expressed as logarithmic K_p values measured in cm/s. Four experimental variables were selected that were deemed particularly influential on the K_p value obtained (and encompassed a comparatively large total quantity of compounds), namely: skin source (breast/abdomen/thigh), skin layer used (epidermis/dermis/epidermis + dermis/stratum corneum), concentration of donor solution (neat/diluted) and donor solution temperature (divided into 20–25/26–30/31–35/36–40 °C). This created a total of 96 experimental scenarios with associated K_p values encompassing all possible combinations (3 for skin source × 4 for skin layer × 2 for concentration and × 4 for temperature). However, of the 253 compounds in the dataset, 71 were excluded as they did not have specified experimental conditions for at least one of the four variables under investigation. Furthermore, of the 96 scenarios only 27 included one or more K_p values (i.e. $n \neq 0$) reducing the dataset further. When more than one K_p value remained (even after applying the four variables) then all values were included in the analysis. These K_p values were then analysed against three physicochemical properties for each compound: partition coefficient (logP), topological polar surface area (TPSA) and molecular volume (MV) as these are known to be influential properties when determining K_p values³⁷, particularly logP and MV as discussed by Tsakovska *et al.*¹¹. Data for these three properties was extracted from Molinspiration (www.molinspiration.com, Molinspiration Cheminformatics, 2022). Multiple linear regression analysis (using Excel Solver) for each scenario created a series of unique QSPR equations with associated coefficients of determination (R^2). Two equations in particular were then analysed in more detail whereby the data was divided into a training and test set for further validation of their performance. In both cases the total number of K_p values were randomly divided into two groups with ~ 80% in the training set and the remaining ~ 20% as a test set, ensuring a range of logP, TPSA and MV values were included in all cases.

Data availability

The datasets analysed during the current study are freely accessible from <https://huskindb.drug350.design.de> or <https://doi.org/10.7303/syn2199888126>.

Code availability

No custom code was used to generate or process the data described in this manuscript.

Received: 10 June 2022; Accepted: 14 September 2022;

Published online: 23 September 2022

References

- Pensado, A. *et al.* Assessment of dermal bioavailability: predicting the input function for topical glucocorticoids using stratum corneum sampling. *Drug Delivery and Translational Research* **12**, 851–861, <https://doi.org/10.1007/s13346-021-01064-8> (2022).
- Cordery, S. F. *et al.* Topical bioavailability of diclofenac from locally-acting, dermatological formulations. *Int J Pharm* **529**, 55–64, <https://doi.org/10.1016/j.ijpharm.2017.06.063> (2017).
- Tabosa, M. A. M. *et al.* Skin pharmacokinetics of diclofenac and co-delivered functional excipients. *Int J Pharm* **614**, 121469, <https://doi.org/10.1016/j.ijpharm.2022.121469> (2022).
- Sabo, S. & Waters, L. J. Poly(dimethylsiloxane): A Sustainable Human Skin Alternative for Transdermal Drug Delivery Prediction. *Journal of Pharmaceutical Sciences* **110**, 1018–1024, <https://doi.org/10.1016/j.xphs.2020.11.028> (2021).
- Waters, L. J. & Sabo, S. Permeation of Pharmaceutical Compounds Through Silanized Poly(dimethylsiloxane). *Journal of Pharmaceutical Sciences* **109**, 2033–2037, <https://doi.org/10.1016/j.xphs.2020.03.007> (2020).
- Waters, L. J., Finch, C. V., Bhuiyan, A. K. M. M. H., Hemming, K. & Mitchell, J. C. Effect of plasma surface treatment of poly(dimethylsiloxane) on the permeation of pharmaceutical compounds. *Journal of Pharmaceutical Analysis* **7**, 338–342, <https://doi.org/10.1016/j.jpha.2017.05.003> (2017).
- Waters, L. J., Shahzad, Y. & Stephenson, J. Modelling skin permeability with micellar liquid chromatography. *European Journal of Pharmaceutical Sciences* **50**, 335–340, <https://doi.org/10.1016/j.ejps.2013.08.002> (2013).
- Pelkonen, O., Turpeinen, M. & Raunio, H. *In vivo-in vitro-in silico* pharmacokinetic modelling in drug development: current status and future directions. *Clin Pharmacokinet* **50**, 483–491, <https://doi.org/10.2165/11592400-000000000-00000> (2011).
- Potts, R. O. & Guy, R. H. A predictive algorithm for skin permeability: the effects of molecular size and hydrogen bond activity. *Pharm Res* **12**, 1628–1633, <https://doi.org/10.1023/a:1016236932339> (1995).
- Pecoraro, B. *et al.* Predicting Skin Permeability by Means of Computational Approaches: Reliability and Caveats in Pharmaceutical Studies. *J Chem Inf Model* **59**, 1759–1771, <https://doi.org/10.1021/acs.jcim.8b00934> (2019).
- Tsakovska, I. *et al.* Quantitative structure-skin permeability relationships. *Toxicology* **387**, 27–42, <https://doi.org/10.1016/j.tox.2017.06.008> (2017).
- Degim, I. T. New tools and approaches for predicting skin permeability. *Drug Discov Today* **11**, 517–523, <https://doi.org/10.1016/j.drudis.2006.04.006> (2006).
- Mitragotri, S. *et al.* Mathematical models of skin permeability: an overview. *Int J Pharm* **418**, 115–129, <https://doi.org/10.1016/j.ijpharm.2011.02.023> (2011).
- Neil, J. E., Brown, M. B. & Williams, A. C. Human skin explant model for the investigation of topical therapeutics. *Scientific Reports* **10**, 21192, <https://doi.org/10.1038/s41598-020-78292-4> (2020).
- Souci, L. & Denesvre, C. 3D skin models in domestic animals. *Veterinary Research* **52**, 21, <https://doi.org/10.1186/s13567-020-00888-5> (2021).
- Bouwman, T., Cronin, M., Bessems, J. & van de Sandt, J. Improving the applicability of (Q)SARs for percutaneous penetration in regulatory risk assessment. *Human & Experimental Toxicology* **27**, 269–276, <https://doi.org/10.1177/0960327107085829> (2008).
- Potts, R. O. & Guy, R. H. Predicting Skin Permeability. *Pharmaceutical Research* **9**, 663–669, <https://doi.org/10.1023/A:1015810312465> (1992).
- Karadzovska, D., Brooks, J. D., Monteiro-Riviere, N. A. & Riviere, J. E. Predicting skin permeability from complex vehicles. *Adv Drug Deliv Rev* **65**, 265–277, <https://doi.org/10.1016/j.addr.2012.01.019> (2013).
- Selzer, D., Abdel-Mottaleb, M. M., Hahn, T., Schaefer, U. F. & Neumann, D. Finite and infinite dosing: difficulties in measurements, evaluations and predictions. *Adv Drug Deliv Rev* **65**, 278–294, <https://doi.org/10.1016/j.addr.2012.06.010> (2013).
- Karadzovska, D., Brooks, J. D. & Riviere, J. E. Experimental factors affecting *in vitro* absorption of six model compounds across porcine skin. *Toxicol In Vitro* **26**, 1191–1198, <https://doi.org/10.1016/j.tiv.2012.06.009> (2012).

21. Grégoire, S. *et al.* Cosmetics Europe evaluation of 6 *in silico* skin penetration models. *Computational Toxicology* **19**, 100177, <https://doi.org/10.1016/j.comtox.2021.100177> (2021).
22. Kladt, C., Dennerlein, K., Göen, T., Drexler, H. & Korinth, G. Evaluation on the reliability of the permeability coefficient (K(p)) to assess the percutaneous penetration property of chemicals on the basis of Flynn's dataset. *Int Arch Occup Environ Health* **91**, 467–477, <https://doi.org/10.1007/s00420-018-1296-5> (2018).
23. Cronin, M. T. D. *et al.* A review of *in silico* toxicology approaches to support the safety assessment of cosmetics-related materials. *Computational Toxicology* **21**, 100213, <https://doi.org/10.1016/j.comtox.2022.100213> (2022).
24. Baba, H., Takahara, J. & Mamitsuka, H. *In Silico* Predictions of Human Skin Permeability using Nonlinear Quantitative Structure-Property Relationship Models. *Pharm Res* **32**, 2360–2371, <https://doi.org/10.1007/s11095-015-1629-y> (2015).
25. Moss, G. P. *et al.* The application of Gaussian processes in the prediction of percutaneous absorption. *J Pharm Pharmacol* **61**, 1147–1153, <https://doi.org/10.1211/jpp/61.09.0003> (2009).
26. Stepanov, D. huskinDB. *Synapse* <https://doi.org/10.7303/syn21998881> (2020).
27. Stepanov, D., Canipa, S. & Wolber, G. HuskinDB, a database for skin permeation of xenobiotics. *Scientific Data* **7**, 426, <https://doi.org/10.1038/s41597-020-00764-z> (2020).
28. Silva, J., Marques-da-Silva, D. & Lagoa, R. Reassessment of the experimental skin permeability coefficients of polycyclic aromatic hydrocarbons and organophosphorus pesticides. *Environ Toxicol Pharmacol* **86**, 103671, <https://doi.org/10.1016/j.etap.2021.103671> (2021).
29. Bittremieux, W. *et al.* Physicochemical properties determining drug detection in skin. *Clinical and Translational Science* **15**, 761–770, <https://doi.org/10.1111/cts.13198> (2022).
30. Rougier, A., Lotte, C. & Maibach, H. I. *In vivo* percutaneous penetration of some organic compounds related to anatomic site in humans: predictive assessment by the stripping method. *J Pharm Sci* **76**, 451–454, <https://doi.org/10.1002/jps.2600760608> (1987).
31. Abd, E. *et al.* Skin models for the testing of transdermal drugs. *Clin Pharmacol* **8**, 163–176, <https://doi.org/10.2147/CPAA.S64788> (2016).
32. Neupane, R., Boddu, S. H. S., Renukuntla, J., Babu, R. J. & Tiwari, A. K. Alternatives to Biological Skin in Permeation Studies: Current Trends and Possibilities. *Pharmaceutics* **12**, <https://doi.org/10.3390/pharmaceutics12020152> (2020).
33. Liu, Y., Wang, L., Liu, J. & Di, Y. A study of human skin and surface temperatures in stable and unstable thermal environments. *Journal of Thermal Biology* **38**, 440–448, <https://doi.org/10.1016/j.jtherbio.2013.06.006> (2013).
34. Lee, C. M., Jin, S.-P., Doh, E. J., Lee, D. H. & Chung, J. H. Regional Variation of Human Skin Surface Temperature. *Ann Dermatol* **31**, 349–352, <https://doi.org/10.5021/ad.2019.31.3.349> (2019).
35. Alves, V. M. *et al.* Predicting chemically-induced skin reactions. Part II: QSAR models of skin permeability and the relationships between skin permeability and skin sensitization. *Toxicol Appl Pharmacol* **284**, 273–280, <https://doi.org/10.1016/j.taap.2014.12.013> (2015).
36. Naseem, S., Zushi, Y. & Nabi, D. Development and evaluation of two-parameter linear free energy models for the prediction of human skin permeability coefficient of neutral organic chemicals. *Journal of Cheminformatics* **13**, 25, <https://doi.org/10.1186/s13321-021-00503-5> (2021).
37. Guth, K. *et al.* *In silico* models to predict dermal absorption from complex agrochemical formulations. *SAR QSAR Environ Res* **25**, 565–588, <https://doi.org/10.1080/1062936x.2014.919358> (2014).

Acknowledgements

The authors wish to thank the University of Huddersfield for funding this work.

Author contributions

L.J.W. and X.L.Q. jointly conceived the study. X.L.Q. created the QSPR models. L.J.W. supervised the analysis and wrote the manuscript with minor edits from X.L.Q.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to L.J.W.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022