2005

Takagi–Sugeno Fuzzy Modeling of Skin Permeability

Deepak R. Keshwani
University of Nebraska-Lincoln, dkeshwani2@unl.edu

David D. Jones
University of Nebraska-Lincoln, david.jones@unl.edu

Rhonda M. Brand
Northwestern University

REVIEW

Takagi–Sugeno Fuzzy Modeling of Skin Permeability

Deepak R. Keshwani

Department of Biological and Agricultural Engineering,
North Carolina State University, Raleigh, North Carolina, USA

David D. Jones

Department of Biological and Agricultural Engineering, North Carolina State University,
Raleigh, North Carolina, USA, and Department of Biological Systems Engineering,
University of Nebraska–Lincoln, Lincoln, Nebraska, USA

Rhonda M. Brand

Division of Emergency Medicine, Evanston Northwestern Healthcare and
Department of Internal Medicine, Feinberg School of Medicine at
Northwestern University, Evanston, Illinois, USA

Corresponding author — David D. Jones, Department of Biological Systems Engineering,
University of Nebraska–Lincoln, 215 L. W. Chase Hall, East Campus,
Lincoln, NE 68583, USA; email djones1@unl.edu

Abstract

The skin is a major exposure route for many potentially toxic chemicals. It is, therefore, important to be able to predict the permeability of compounds through skin under a variety of conditions. Available skin permeability databases are often limited in scope and not conducive to developing effective models. This sparseness and ambiguity of available data prompted the use of fuzzy set theory to model and predict skin permeability. Using a previously published database containing 140 compounds, a rule-based Takagi–Sugeno fuzzy model is shown to predict skin permeability of compounds using octanol-water partition coefficient, molecular weight, and temperature as inputs. Model performance was estimated using a cross-validation approach. In addition, 10 data points were removed prior to model development for additional testing with new data. The fuzzy model is compared to a regression model for the same inputs using both R^2 and root mean square error measures. The quality of the fuzzy model is also compared with previously published models. The statistical analysis demonstrates that the fuzzy model performs better than the regression model with identical data and validation protocols. The prediction quality for this model is similar to others that were published. The fuzzy model provides insights on the relationships between lipophilicity, molecular weight, and temperature on percutaneous penetration. This model can be used as a tool for rapid determination of initial estimates of skin permeability.

Keywords: fuzzy modeling, skin permeability, dermal toxicology, cross-validation

Introduction

The penetration of chemicals through skin is a major source of absorption for numerous classes of compounds encountered occupationally (1), in consumer products (2), and in the water supply (3). The ability to predict skin permeability is important for the assessment of toxicity following topical exposure.

One approach to modeling transdermal absorption is to describe the system by adapting existing mass-transfer equations (4–6). The functional model, however, is very complex; therefore, a multiple regression approach to estimate relationships between physical properties and percutaneous penetration has often been utilized (7–9). Clustering techniques, in which a chemical database is subdivided into subgroups or clusters and equations are developed for each of the groups, have also been used to describe skin permeability (10,11). In these studies, however, the clusters were crisp, i.e., each data point belonged entirely to one group. Fuzzy modeling allows each data point to have varying degrees of membership in different clusters. Pannier et al. (12) concluded that a rule-based fuzzy modeling approach is realistic and promising by comparing results from their model to regression analysis done by Potts and Guy (7,8), Abraham and coworkers (13), and Flynn (5).

The quality of data-driven models is largely dependent on the completeness of the database available for training and subsequent validation. This is a problem in the area of skin permeability, where databases are usually small. A typical approach would be to divide the data into training and testing sets. However, most current empirical models have used all the data in a single training set and based the quality of the model on this analysis. With a small dataset, it would be ideal to use all the data available for training, but at the same time be able to test and validate the model. Cross-validation, a method to test model performance by resampling the available data, can be used for this purpose (14).

The databases used in previously published models often differ in the types and ranges of the parameters included. As such, combining these data sources is problematic because of inherent inconsistencies in the experimental conditions. Vecchia and Bunge (15) addressed this by providing a database that overcomes some of the problems associated with diverse data sources. They compiled information on the permeability of over 140 compounds. Each data point was validated, and information about experimental conditions was recorded. This expanded database provides a means by which to examine the hypothesis that a rule-based fuzzy model, combined with cross-validation sampling, will result in a more effective predictor of skin permeability than the corresponding regression analysis. Molecular weight (MW) and octanol-water partition coefficient (Log K_{ow}) were selected as inputs based on their availability within the database and their demonstrated importance in influencing dermal penetration (7,11). Temperature was also included as a confounding factor, because it was available for each data point, and it is known to influence transdermal absorption (15–18).

The objectives of this study were to develop a Takagi–Sugeno fuzzy model to predict skin permeability from MW, Log K_{ow}, and temperature using a cross- validation approach; compare model performance with a regression model on the same data and previously published models; and interpret the model to obtain information on the system.

Fuzzy Set Theory and Modeling

Systems and phenomena to be modeled are typically complex, arising from uncertainty in the form of ambiguity. This uncertainty is usually assumed to be random in nature, and probability theory is used to describe it. However, it can be argued that the uncertainty related to complex systems and phenomena is not entirely random (19). Hence, a probabilistic approach would fail to capture the nonrandom uncertainty. Fuzzy set theory provides a means by which to represent this ambiguity through possibility rather than probability (20). In the context of skin permeability, potential sources of uncertainty include inadequate understanding of the system, experimental procedures, and data collection protocols.

Unlike classical set theory where objects have complete or zero membership in a set, fuzzy set theory allows for partial memberships in multiple sets. Lofti Zadeh introduced the idea of a fuzzy set and defined it as a class with a continuum of grades of membership (21). For instance, a group of people exactly 6 ft tall would comprise a crisp set, with each person having complete membership in the set. However, a group of people around 6 ft tall would comprise a fuzzy set, with each person having different degrees of membership (complete or partial) in the set. Membership functions are used to describe the ambiguity associated with a fuzzy set. For a crisp set, these functions are constant for all members. For a fuzzy set, membership functions are shaped based on degree of belongingness of each point in the data space. Fuzzy membership functions can be obtained through several means, including intuition, neural networks, genetic algorithms, clustering, and soft partitioning (19).

Clustering partitions data and develops rules for the fuzzy model. This involves dividing data points into homogeneous clusters so that items in the same group are as similar as possible and items not in that group are as dissimilar as possible. Clustering can also be thought of as a form of data compression, where a large number of samples are converted into a small number of representative prototypes. In nonfuzzy systems, the data space is divided into crisp clusters, where each data point belongs to exactly one group. In a fuzzy system, however, the data points can belong to more than one cluster, and associated with each of the points are membership grades that indicate the degree to which the data points belong to the different clusters. Babuska (22) provides a description of various clustering methods that can be used. Once data clustering is accomplished, rules are developed for each group. The output from each rule is a crisp linear function. The final predicted value is a weighted average of the outputs from each rule. Weights for individual rules are determined by the membership of input values to each cluster.

Methods

The methods used in this work are set forth to predict the skin permeabilities of compounds using octanol-water partition coefficient, molecular weight, and temperature as inputs. This work entails the use of the Vecchia and Bunge database to develop fuzzy rule-based models and regression models. As such, the detailed procedures include the use of the data, the development of the fuzzy clusters and subsequent rules (fuzzy model development), the procedures for the regression models, the model validation strategies, the measures of model performance, and the eliciting of information from the fuzzy model.

Database

A dataset of 140 compounds was obtained from the Vecchia and Bunge database (15). Ten compounds were randomly selected and set aside to mimic new data for additional model testing. The database contained information on the permeability of over 140 compounds. Each of the data points was validated, and information about the setting and environmental conditions associated with the evaluations was recorded. The dataset contained compounds with MWs ranging from 32 to 585 Daltons, Log K_{ow} ranging from –3.1 to 4.34, reported temperatures ranging from 22 to 39 °C, and Log K_p ranging from –5.5 to –0.08. The Vecchia and Bunge database consists of compounds from various studies that utilized the following human skin types: isolated stratum corneum, epidermal membranes, and split and full-thickness skin (15). However, the database does not report the location from which skin samples were obtained. It has been reported that there are regional variations in transdermal absorption (23). Although skin type and location were not considered as inputs in the model, their effects on skin permeability are considered a source of ambiguity that prompts the need for fuzzy set theory as a modeling tool.

Fuzzy Model Development

Subtractive clustering was used to develop the fuzzy models. This method iteratively assumes each data point to be a possible cluster center and calculates the potential of each datum to define the midpoint based on the density of surrounding data. Based on optimization of distance measures, the routine first selects the data point with the highest potential to be the first cluster center. It then removes all data points in the vicinity of the first cluster center in order to determine the next data cluster and the location of its center. The routine iterates on this process until predefined stopping criteria are met. A detailed description of the subtractive clustering routine is described by Chiu (24). Rules were developed and optimized for each cluster. The optimization was done using the Adaptive Neural Fuzzy Inference System (25) as implemented in the Fuzzy Logic Toolbox within MATLAB (Version 6.1, Release 12.1, Mathworks, Inc., 2000). The code to produce these models and implement the validation approach is shown in Figure 1.

Regression Model Development

A linear regression model was developed to predict the skin permeability of a compound as a function of octanol-water partition coefficient, molecular weight, and temperature. This was done to enable comparison of two inherently different types of modeling scheme utilizing the same data and validation approach.

Validation Approach

The leave-one-out (LOO) cross-validation approach (14) is an extension of the K-fold approach (26). Both approaches are resampling methods that maximize the use of available data. K-folding involves the division of the dataset into several (K) subsets. Each time a

```
% This matlab m file implements LOO cross-validation
% for the Takagi–Sugeno fuzzy modeling scheme

% Note: infis, genfis2, TRNAOPT, DISOPT, anfis and evalfis are
% functions available in the fuzzy logic toolbox in MATLAB.

% Note: Relevant comments are included below command lines

start_time=clock

A=load('Sugenodata.dat');
% loading dataset into MATLAB

t=130;
% "t" is the number of points in the dataset

results=[0 0 0 0];
% initializing the result matrix

for n=1:1:t;
   B=A(n,1:3);
   % "B" contains the input values of the omitted datum
   act=A(n,4);
   % "act" contains the actual output of the test point
   D1=A(1:n-1,:);
   D2=A(n+1:t,:);
   D=[D1;D2];
   % concatenating to get training dataset the fuzzy model
   Din=D(:,1:3);
   Dout=D(:,4);
   infis=genfis2(Din,Dout,0.5);
   TRNOPT=[100 0 0.01 0.9 1.1];
   DISPOPT=[0 0 0 0];
   [outfis,xerror1,xstepsize]=anfis(D,infis,TRNOPT,DISPOPT,1);
   rules=size(getfield(outfis,'rule'),2);
   out=evalfis(B,outfis);
   ES= (act-out)^2;
   % ES is the square error between actual and predicted
   output= [act out rules ES];
   results=vertcat(results,output);
end

dlmwrite('Sugenoresults',results);
% results are written into a data file

stop_time=clock
```

Figure 1. MATLAB code used to produce Takagi–Sugeno fuzzy models and implement the Leave-one-out (LOO) cross-validation procedure.

model is developed using the dataset, one of the subsets is set aside to test the model. This is repeated until each subset is used for testing. For a dataset of size N, the LOO approach develops N number of models. During each trial, one datum is omitted, and a model is

generated using $N - 1$ data points. The process is repeated until each datum has been used as a testing point once. The prediction errors from test cases are pooled and used to cross-validate the model. The advantage with using the LOO approach is that it utilizes all of the data and avoids difficulties associated with traditional data splitting, such as inadequate homogeneity between training and validation samples and sample size dilution (14). Figure 2 illustrates how the dataset was utilized in this study for model development and validation. This validation approach was employed for both the fuzzy model and the linear regression model.

Measures of Model Performance

Model performance is determined in the same manner for the fuzzy rule-based model and the regression model. Each model was evaluated using three different measures of root mean square error (RMSE) and three different R^2 (square of correlation coefficient) measures. Each of the measures of model performance corresponds to each collection of data shown in Figure 2. The description of each follows:

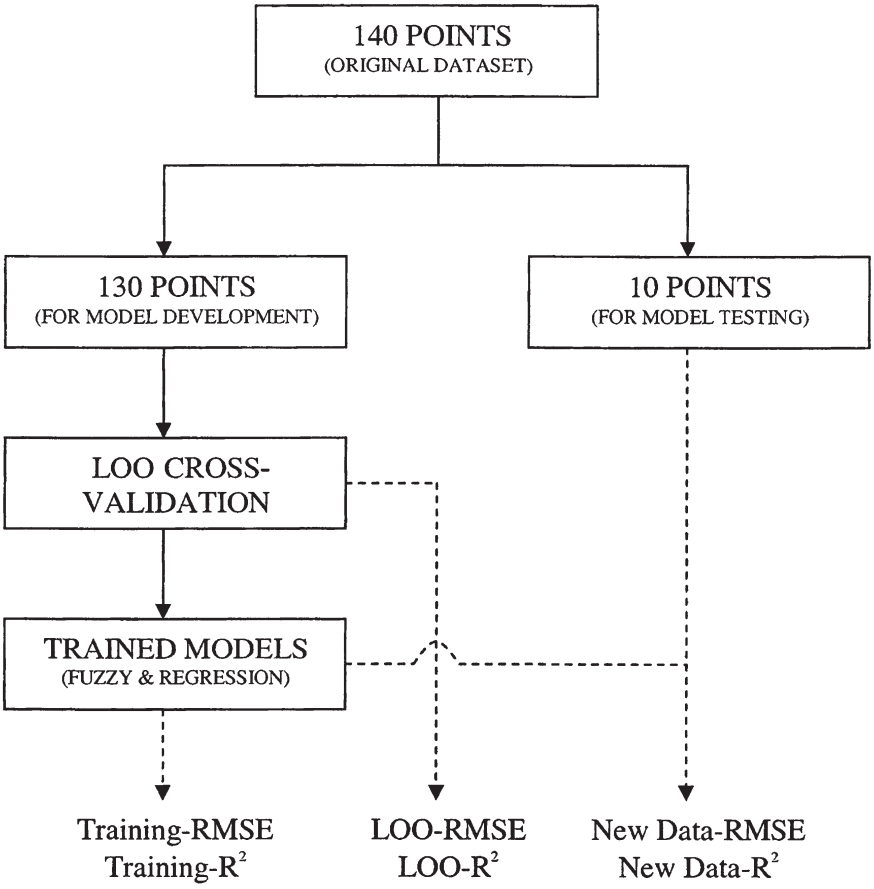

Figure 2. Utilization of data for model development and validation.

- Training-RMSE: This calculates the average square root of the sum of squared errors associated with the training data set.
- LOO-RMSE: This calculates the average square root of the sum of squared errors during the LOO cross-validation procedure. This sum of squares from the LOO procedure is also referred to as a predicted residual sum of squares statistic (14).
- New Data-RMSE: This calculates the average square root of the sum of squared errors for a new dataset.
- Training-R^2: This is a correlation of predicted vs. actual values for the entire data- set.
- LOO-R^2: This is correlation of predicted vs. actual during cross-validation by pooling omitted test cases from all the trials.
- New Data-R^2: This is a correlation of prediction vs. actual values for new test data.

Eliciting Information from the Fuzzy Model

The rules for the fuzzy model were developed based on the clustering of the data space and can be interpreted to reveal the nature of the relationships between the inputs and the output and even some information about the system. The fuzzy model developed using MATLAB includes a graphic user interface (GUI), where Log K_p can be determined by setting the rules at a particular set of inputs. The GUI was used to conduct a sensitivity analysis of the model with respect to the inputs. The data collected refer to model predicted values of Log K_p in the following input scenarios:

- Three temperatures: 25 °C, 32 °C, and 37 °C, representing the most frequently reported temperature values in the original database and the locations of the cluster centers in the temperature domain
- Six Log K_{ow}: –1, 0, 1, 2, 3, and 4, representing a range from hydrophilic to lipophilic compounds
- Six values of molecular weight: 50, 100, 200, 300, 400, and 500 Daltons

Results

Description of Fuzzy Model

The inputs to the model were Log K_{ow}, MW, and temperature. The output is Log K_p. The rule structure of the model is shown in Figure 3. There are seven rules in the model, and each row of membership functions represents a rule that describes the cluster in each of the input dimensions and the corresponding relationship to the output space. Figure 3 shows the model predicting a Log K_p of –2.66 for a compound that has a Log K_{ow} of 1.25 and a MW of 200 at a temperature of 32 °C. Gaussian functions are used to describe the membership profile for each input parameter. There are two sets of parameters associated with each Gaussian membership function: a center and a standard deviation. The output is a linear function localized for each cluster, which uses fuzzy inputs to calculate a crisp output. The weighted average of the outputs from each rule is the final Log K_p predicted by the model. Each localized linear function has four parameters associated with it that are optimized during model development.

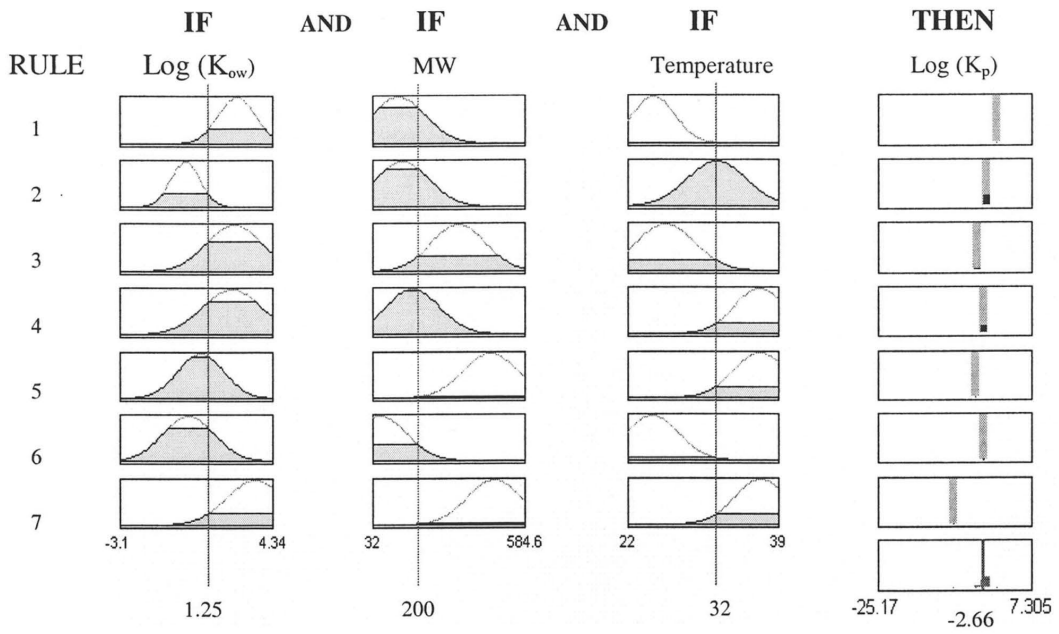

Figure 3. Rule structure of the developed fuzzy model consisting of three input parameters and seven rules.

Description of Regression Model

The regression model was a simple linear model. The coefficients resulted in the following expression:

$$\text{Log } K_p = a \cdot \text{Log } K_{ow} + b \cdot \text{MW} + c \cdot T + d$$

where a = 0.55743; b = –0.00589; c = 0.07434; and d = –4.63779.

Evaluating Model Performance

The six previously described measures of model performance were calculated and are presented in Table 1. A low RMSE and a high R^2 are indicative of a model that performs well with new data. The fuzzy model had a LOO-RMSE of 0.64 and a LOO-R^2 of 0.74. The regression model had a LOO-RMSE of 0.82 and a LOO-R^2 of 0.60. These indicate that the fuzzy model is better for predicting new values. To further validate this claim, a dataset of 10 new compounds was introduced and evaluated using the fuzzy and regression model. New data-RMSE and New data-R^2 measures were calculated for each model. The fuzzy model had a New data-RMSE of 0.50 and a New data-R^2 of 0.79. The corresponding values for the regression model were 0.72 and 0.50. For the fuzzy model, the New data-R^2 was reasonably close to the LOO-R^2 estimate. This indicates that there is no overfitting of the data used to develop the model. Additional comparisons between the two mod-

Table 1. RMSE and R^2 measures to compare fuzzy and regression models

Measure	Fuzzy model	Regression model
Training R^2	0.88	0.64
LOO R^2	0.74	0.60
New data R^2	0.79	0.50
Training RMSE	0.42	0.73
LOO RMSE	0.64	0.82
New data RMSE	0.50	0.72

els were made using the Training-RMSE and the Training-R^2 obtained by evaluating the models using the entire dataset. The results from training statistics once again suggest the superiority of the fuzzy model. These results indicate that the fuzzy model will perform better in predicting skin permeability coefficients for new data.

Interpretation of the Rule-Based Fuzzy Model

A sensibility analysis of the model was performed to assess the influence of the inputs in predicting skin permeability coefficients. The rules were used to collect Log K_p values predicted by the model at different combinations of inputs (shown in Table 2). Figure 4 shows the effects of MW at 25, 32, and 37 °C. Skin permeability is observed to have a neg-

Table 2. Data collected during sensitivity analysis of rules

	Molecular weight						
Log K_{ow}	50	100	200	300	400	500	Temperature
–1	–3.32	–3.58	–4.22	–5.93	–7.02	–6.53	25
0	–2.98	–3.19	–3.89	–5.84	–6.30	–5.73	25
1	–2.60	–2.72	–3.52	–5.47	–5.54	–4.94	25
2	–2.01	–1.85	–2.36	–4.42	–4.69	–4.13	25
3	–1.52	–1.27	–1.69	–3.61	–3.88	–3.33	25
4	–1.26	–1.04	–1.76	–3.24	–3.14	–2.54	25
–1	–4.93	–4.62	–5.00	–5.46	–5.72	–5.29	32
0	–3.37	–3.53	–3.85	–4.32	–4.67	–4.30	32
1	–2.07	–2.26	–2.80	–3.63	–3.91	–3.72	32
2	–0.95	–1.38	–2.53	–3.41	–3.75	–4.16	32
3	–0.44	–0.96	–2.15	–3.05	–4.03	–5.19	32
4	0.02	–0.54	–1.73	–2.82	–4.28	–5.50	32
–1	–4.73	–4.82	–5.17	–6.15	–6.18	–5.63	37
0	–3.60	–3.73	–4.16	–5.07	–5.04	–4.48	37
1	–1.56	–1.91	–2.80	–3.98	–3.92	–3.39	37
2	–0.43	–0.97	–2.10	–3.18	–2.98	–2.53	37
3	0.02	–0.54	–1.66	–2.62	–2.33	–2.00	37
4	0.43	–0.12	–1.24	–2.07	–1.69	–1.42	37

Black values represent the inputs; red numbers show predicted Log K_p values.

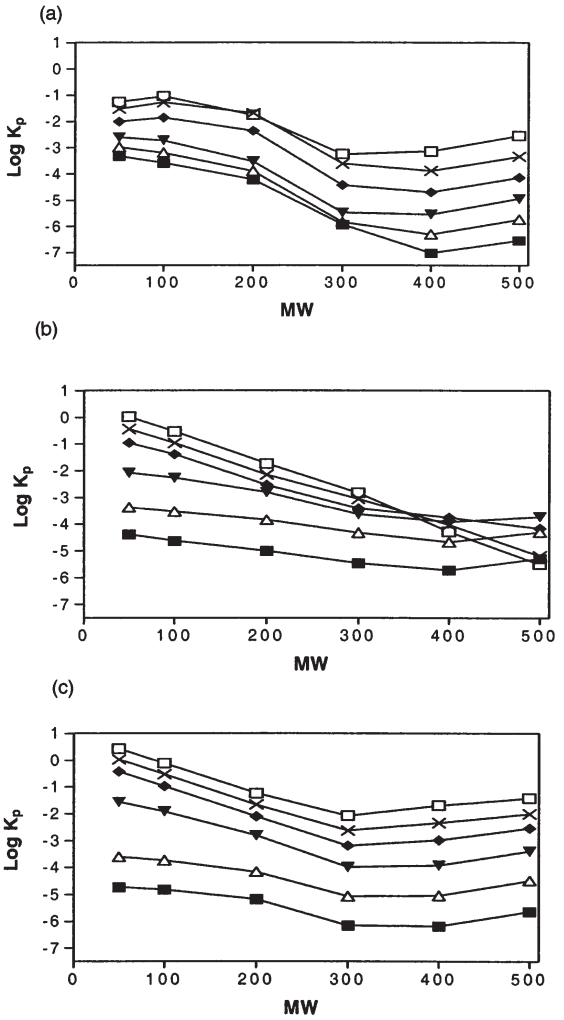

Figure 4. Effect of molecular weight on Log K_p at (**a**) 25 °C, (**b**) 32 °C, and (**c**) 37 °C. Data are divided by log K_{ow}: –1(■), 0 (♦), 1 (Δ), 2 (▼), 3 (×), and 4 (□).

ative correlation with MW in general. The drop in Log K_p seems to be greater between MWs of 100 to 300. For lower and higher MW compounds, the Log K_p seems to level off. The only exception is for compounds with high lipophilicity at 32 °C. In this case, the negative correlation between MW and Log K_{ow} occurs at all molecular weights. Figure 5 shows the effects of K_{ow} at 25, 32, and 37 °C, respectively, for different values of MW. Skin permeability is observed to have a positive correlation with Log K_{ow} in general. The exception is at 32 °C, where the correlation between Log K_{ow} and temperature is negative for high lipophilicity. Data collected during sensitivity analysis show that there is an effect of temperature on Log K_p. For lipophilic compounds, skin permeability has a positive correlation with temperature. For low molecular weight hydrophilic compounds, penetration actually decreases as temperature increases.

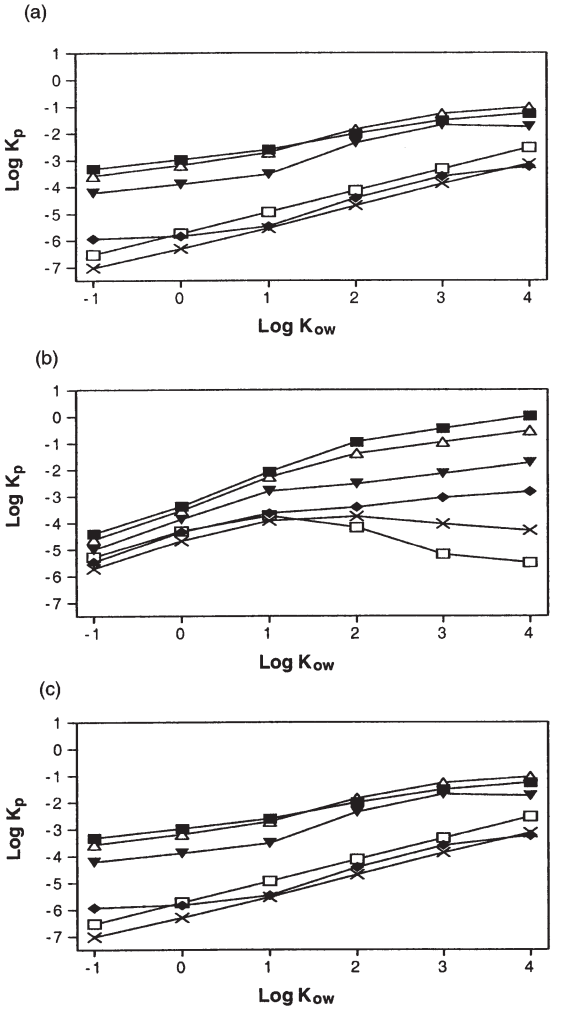

Figure 5. Effect of Log K_{ow} on Log K_p at (a) 25 °C, (b) 32 °C, and (c) 37 °C. Data are divided by MW: 50 (■), 100 (♦), 200(Δ), 300 (▼), 400 (×), and 500 (□).

Discussion

The prevalence of dermal absorption as a source of potentially toxic compounds warrants the developments of methods to predict penetration under a variety of exposure conditions. Efforts in this area have been hampered by limited and inconsistent data. The Vecchia and Bunge database attempts to unify available data by validating each point and providing information on experimental conditions. Fuzzy set theory was used as a means to control for both limited availability of data and variability within the dataset. The quality of the model was determined by correlating predicted outputs to actual values using both RMSE and R^2 statistics. Both statistics indicate that the fuzzy model is more accurate than the regression model. Furthermore, the interactions between molecular weight, lipo-

philicity, and temperature on transdermal penetration could be elucidated using the rules developed by the model.

In order to determine whether a modeling technique is valuable, it is necessary to compare the quality of its output to other techniques using the same data. Model performance was assessed using a cross-validation LOO approach. Cross-validation enables true model performance with new data to be estimated without sacrificing a portion of the data for testing. Results in Table 1 indicate that the fuzzy model is better than the regression model for the same dataset and validation strategy. The superiority of the fuzzy model over the regression model is clear for both cross-validated and completely new data. This differential, therefore, becomes particularly important when using the model to predict the absorption of compounds that are not present in available databases.

It is also important to examine the quality of the model with data that were not used in its development. Most studies using data-based models to predict transdermal absorption do not perform this analysis. Table 3 compares the results of the developed model to previously published models. The training-R^2 estimate of the fuzzy model compares well with those of previously published models. In most published models that predict skin permeability, all of the data are used to create the model, leaving the R^2 estimate to correlate only the training data. While this provides an idea of model fitness, it does not, however, give sufficient information on performance with entirely new data. The LOO-RMSE and LOO-R^2 presented in this study achieve this objective and demonstrate that the model quality is not significantly reduced when new data are applied.

Studying the rules enabled examination of the interactions between molecular weight, lipophilicity, and temperature on skin permeability of compounds. The model predicts an inverse correlation between molecular weight and penetration. This relationship was previously observed by multiple investigators (7,9,12,13). Magnusson et al. propose that molecular weight is actually the single most important factor in determining percutaneous absorption (27). The fuzzy model predicted a plateau in transdermal penetration for

Table 3. Comparison of results with previously published models

Ref.	Dataset used	Size	Type	Inputs	Training R^2	LOO R^2
—	Vecchia and Bunge*	130	F	Log K_{ow}, MW, T	0.88	0.74
11	Flynn	93	F	Log K_{ow}, MW	0.83	—
11	Potts and Guy	37	F	Log K_{ow}, H_d	0.97	—
11	Abraham	53	F	MW, H_d	0.96	—
2	Flynn	93	R	Log K_{ow}, MW	0.67	—
2	Potts and Guy	37	R	Log K_{ow}, MW, Log D_o/p	0.84	—
2	Scheuplein and Blank	23	R	Log K_{ow}, MW, Log D_o/p	0.90	—
4	Kirchner/Flynn	65	R	Log K_{ow}, MW	0.85	—
4	Kirchner/Flynn	39	R	Log K_{ow}, MW	0.91	—
—	Abraham**	53	R	Log K_{ow}, MW, pi, H_a, H_d	0.77	—

* Model presented in this paper.

** Model developed by authors during analysis of results.

F refers to fuzzy model; **R** refers to regression model; $\mathbf{H_d}$ is the hydrogen bond donor activity; $\mathbf{H_a}$ is the hydrogen bond acceptor activity; $\boldsymbol{D_o}$ is the molecular diffusivity; and $\boldsymbol{p}$ is the diffusion path length. Column 1 cites the specific article in the References.

chemicals with molecular weights between 300 and 500. While somewhat surprising, it is consistent with the work of Mitragotri (28), who modeled skin permeability based on the different permeation pathway and found exponential decays as a function of size, which results in a penetration plateau.

The fuzzy model also predicts that increasing lipophilicity correlates with higher transdermal penetration, which is consistent with other works (7,29). Overall, for lipophilic compounds, the model predicted a positive correlation between temperature and Log K_p. The permeability coefficient becomes greater as the temperature increases from room to body temperature for the lipophilic compounds. The relationship between increasing temperature and greater transdermal penetration has been studied for a number of chemicals, including parathion (30), chloroform (31), trichloroethylene and tetrachloroethylene (32), VX (33), methyl paraben and butyl paraben (16), and sodium laurel sulfate (34). Each of these chemicals is lipophilic with a Log K_{ow} ranging from 1.5 for chloroform to 3.7 for parathion. There are two likely mechanisms by which this could occur. The first is that increasing the temperature raises the solubility of the chemical in the donor solution, leading to a higher donor concentration, which according to Fick's law would cause increased transdermal absorption. The second mechanism involves a direct effect on skin lipids. An increase in temperature from 22 °C to 37 °C would be expected to cause an increase in lipid fluidization, even though there is not a differential scanning calorimetry (DSC)-measured lipid transition point until approximately 40 °C (35). This increase in fluidization would, therefore, be expected to impact absorption for hydrophobic chemicals.

The fuzzy model also suggests that the penetration of hydrophilic compounds is not strongly influenced by temperature and actually decreases at low molecular weights. It was reported that increasing the temperature from 30 to 37 °C did not significantly increase the transdermal penetration of caffeine, a hydrophilic chemical with a log K_{ow} of –0.07 (16). This observed independence of temperature on the permeability of hydrophilic species is also consistent with the solubility theory described above in that the concentration of hydrophilic chemicals in the donor solution is less likely to change with increasing temperature than their hydrophobic counterparts. However, it may also reflect the differences in penetration pathways used by hydrophilic and lipophilic compounds. Lipophilic chemicals will partition directly through the lipid bilayers of the stratum corneum during dermal penetration. Hydrophilic molecules, however, are less likely to move directly through these bilayers and have been proposed to penetrate via an aqueous pathway. While there is no experimental evidence to support the presence of this pathway, theoretically, these molecules would penetrate closer to the polar head groups, areas close to the more hydrophilic corneocytes (36) or through imperfections in the lipid bilayer (28).

The data-driven models have a limited range for the inputs being used due to the nature of existing datasets. This highlights the need for a generalized skin permeability model for a greater range of inputs for which data are unavailable. In this regard, a logical step is to move toward non-data-driven models to predict skin permeability. An example of this would be a Mamdani-type fuzzy model. In the Mamdani modeling scheme, membership functions and rules are manually developed using knowledge acquired from different sources. These sources could include previous models, skin permeability literature, and expert opinion.

Nonetheless, this study presents a cross-validated data-driven fuzzy rule-based skin permeability model with measures of model performance. Results indicate the fuzzy model is better than the regression model in predicting skin permeability and is comparable to previously published models. The rules of the model were subsequently interpreted so as to identify trends in the interaction and effect of inputs. Certain trends such as the plateau effect for low-and high-MW compounds need to be further investigated to determine if they are representative of the system. This model can be used as a tool to rapidly determine initial estimates of skin permeability coefficient to determine the potential for transdermal drug delivery and toxicity.

Acknowledgments

The authors would like to acknowledge Dr. Annette Bunge for providing the fully validated skin permeability database. In addition, the authors would like to acknowledge Dr. George E. Meyer and Dr. David Marx for their help in preparing the manuscript.

References

1. Fenske RA, Elkner KE. Multi-route exposure assessment and biological monitoring of urban pesticide applicators during structural control treatments with chlorpyrifos. Toxicol Ind Health 1990; 6:349–371.
2. Stott WT, Bartels MJ, Brzak KA, Mar M, Markham DA, Thornton CM, Zeisel SH. Potential mechanisms of tumorigenic action of diethanolamine in mice. Toxicol Lett 2000; 114:67–75.
3. Xu X, Mariano TM, Laskin JD, Weisel CP. Percutaneous absorption of trihalomethanes, haloacetic acids, and haloketones. Toxicol Appl Pharmacol 2002; 184:19–26.
4. Edwards DA, Langer RA. Linear theory of transdermal transport phenomenon. J Pharm Sci 1994; 83:1315–1334.
5. Flynn GL. Physiochemical 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.
6. Kasting GB. Kinetics of finite dose absorption through skin 1: Vanillylnonanamide. J Pharm Sci 2001; 90:202–212.
7. Potts RO, Guy RH. Predicting skin permeability. Pharm Res 1992; 9:663–669.
8. Potts RO, Guy RH. A predictive algorithm for skin permeability: the effects of molecular size and hydrogen bonding activity. Pharm Res 1995; 12:1628–1633.
9. Moody RP, MacPherson H. Determination of dermal absorption QSAR=QSPRs by brute force regression: multiparameter model development using MOLSUITE 2000. J Toxicol Environ Health A2003; 66:1927–1942.
10. Kirchner LA, Moody RP, Doyle E, Bose R, Jeffrey J, Chu I. The prediction of skin permeability by using physicochemical data. ATLA 1997; 25:359–370.
11. Magnusson BM, Pugh WJ, Roberts MS. Simple rules defining the potential of compounds for transdermal delivery or toxicity. Pharm Res 2004; 21:1047–1054.
12. Pannier AK, Brand RM, Jones DD. Fuzzy modeling of skin permeability coefficients. Pharm Res 2003; 20:143–148.
13. Abraham MH, Martins F, Mitchell RC. Algorithms for skin permeability using hydrogen bond descriptors: the problem of steroids. J Pharm Pharmacol 1997; 49:858–865.
14. Holiday DB, Ballard JE, McKeown BC. PRESS-related statistics: regression tools for cross-validation and case diagnostics. Med Sci Sports Exerc 1995; 27:612–620.

15. Vecchia BE, Bunge AL. Skin absorption databases and predictive equations. In: Hadgraft J, Guy RH, eds. Transdermal Drug Delivery Systems. New York: Marcel Dekker; 2003; 123:57–414.
16. Akomeah F, Nazir T, Martin GP, Brown MB. Effect of heat on the percutaneous absorption and skin retention of three model penetrants. Eur J Pharm Sci 2004; 21:337–345.
17. Johnson ME, Blankschtein D, Langer R. Permeation of steroids through skin. J Pharm Sci 1995; 84:1144–1146.
18. Liron ZVI, Cohen S. Percutaneous absorption of alkanoic acids I: a study of operational conditions. J Pharm Sci 1984; 73:534–537.
19. Ross TJ. Fuzzy Logic with Engineering Applications; New York: McGraw-Hill, Inc., 1995.
20. Zadeh LA. Fuzzy sets as a basis for a theory of possibility. Fuzzy Set Syst 1978; 1:3–28.
21. Zadeh LA. Fuzzy sets. Inform Contr 1965; 8:338–353.
22. Babuska R. Fuzzy Modeling for Control; Boston, MA: Kluwer Academic Publishers, 1998.
23. Maibach HI, Feldman RJ, Milby TH, Serat WF. Regional variation in percutaneous penetration in man—pesticides. Arch Environ Health 1971; 23:208–211.
24. Chiu S. Fuzzy model identification based on cluster estimation. J of Intelligent and Fuzzy Syst 1994; 2(3):267–278.
25. Jang JSR. ANFIS: adaptive-network-based fuzzy inference system. IEEE T Syst Man Cyb 1993; 23:665–684.
26. Blum A, Kalai A, Langford J. Beating the hold-out: bounds for k-fold and progressive cross-validation, 12th Annual Conference on Computational Learning Theory, Santa Cruz, CA, July 6–9, 1999.
27. Magnusson BM, Anissimov YG, Cross SE, Roberts MS. Molecular size as the main determinant of solute maximum flux. J Invest Dermatol 2004; 122:993–999.
28. Mitragotri S. Modeling skin permeability to hydrophilic and hydrophobic solutes based on four permeation pathways. J Control Rel 2002; 86:68–92.
29. Lee CK, Uchida T, Kitagawa K, Yagi A, Kim N, Goto S. Relationship between lipophilicity and skin permeability of various drugs from an ethanol=water=lauric acid system. Biol Pharm Bull 1994; 17:1421–1424.
30. Chang SK, Reviere JE. Percutaneous absorption of parathion in vitro in porcine skin: effects of dose, temperature, humidity, and perfusate. Fund Appl Toxicol 1998; 17: 494–504.
31. Gordon SM, Wallace LA, Callahan PJ, Kenny DV, Brinkman MC. Effect of water temperature on dermal exposure to chloroform. Environ Health Perspect 1998; 106:337–345.
32. Nakai JS, Stathopulos PB, Campbell GL, Chu I, Li-Muller A, Aucoin R. Penetration of chloroform, trichloroethylene, and tetrachloroethylene through human skin. J Toxicol Environ Health A 1999; 58:157–170.
33. Craig FN, Cummings EG, Sim VM. Environmental temperature and the percutaneous absorption of a cholinesterase inhibitor, VX. J Invest Dermatol 1977; 68:357–361.
34. Emilson A, Lindberg M, Forslind B. The temperature effect on in vitro penetration of sodium lauryl sulfate and nickel chloride through human skin. Acta Derm Venereol 1993; 73:203–207.
35. Knutson K, Potts RO, Guzek GB, Golden GM, McKie JW, Lambert WJ, Higuchi WI. Macro-and molecular physical–chemical considerations in understanding drug transport in the stratum corneum. J Control Rel 1985; 2:67–87.
36. Kurihara-Bergstrom T, Knutson K, DeNoble LJ, Goates CY. Percutaneous absorption enhancement of an ionic molecule by ethanol–water systems in human skin. Pharm Res 1990; 7:762–766.