The 8th International Electronic Conference on Medicinal Chemistry (ECMC 2022)
01–30 NOVEMBER 2022 | ONLINE

In silico prediction of skin permeability by various models

Chaired by DR. ALFREDO BERZAL-HERRANZ;
Co-Chaired by PROF. DR. MARIA EMÍLIA SOUSA

pharmaceuticals MDPI

Giang Huong Ta[1], Max K. Leong[1]*

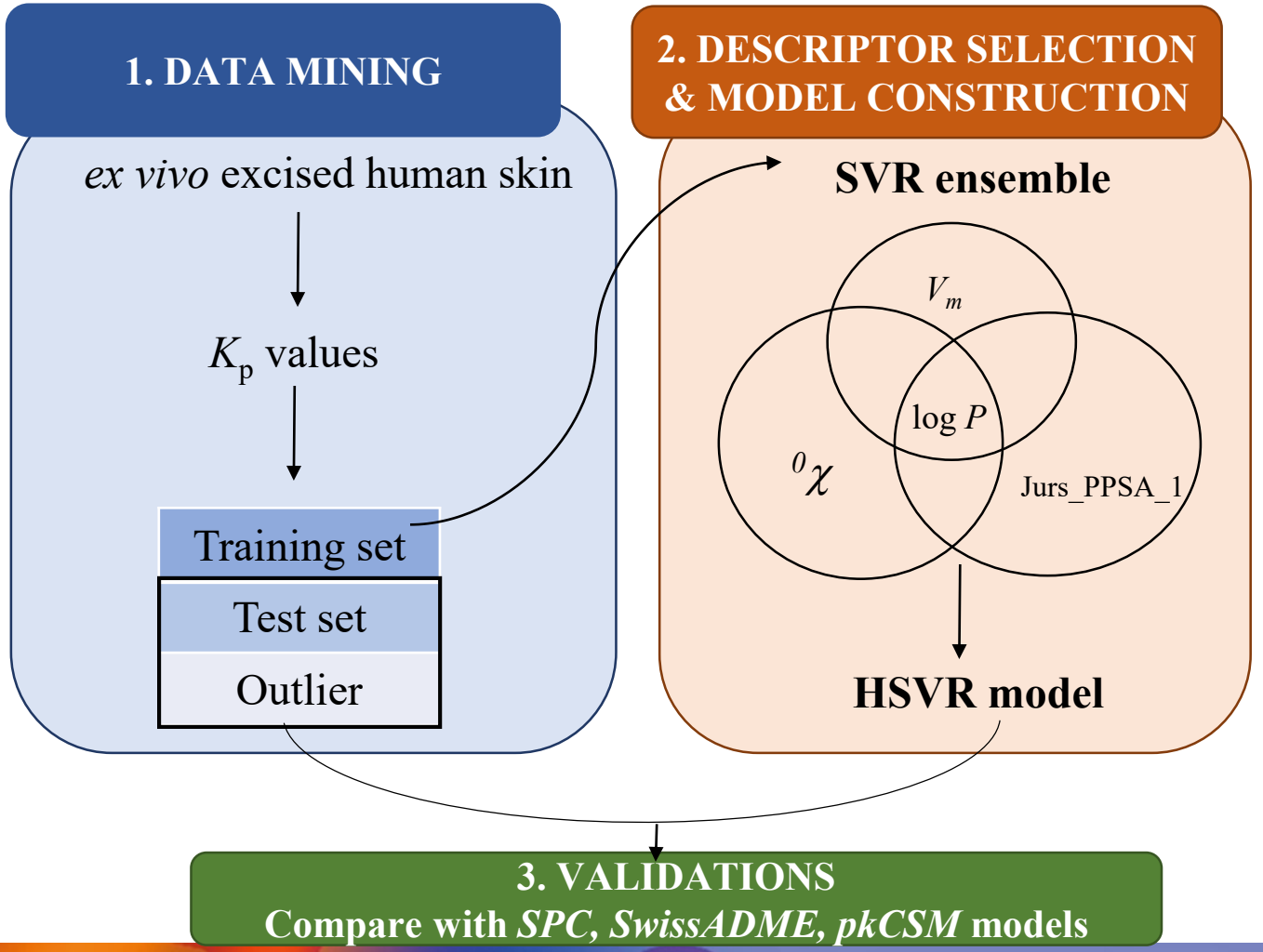
[1] *Department of Chemistry, National Dong Hwa University, Shoufeng, Hualien 974301, Taiwan.*

**Corresponding author: leong@gms.ndhu.edu.tw*

國立東華大學
NATIONAL DONG HWA UNIVERSITY

In silico prediction of skin permeability by various models

Graphical abstract



Abstract

Skin is the largest organ in the human body and works as the natural barrier against the external environment. Furthermore, topical and transdermal drug delivery has been emerged as the new effective and safer administration. A variety of *in vitro*, *in vivo*, and *ex vivo* assays have been adopted to evaluate the retention of the drug in the skin layers and skin permeability, in which the *ex vivo* excised human skin has been considered as the gold standard to assess the skin penetration despite its potential of ethical issues. In this study, the novel machine learning-based hierarchical support vector regression (HSVR) was adopted to generate a nonlinear quantitative structure-activity relationship (QSAR) model, which can predict the K_p values based on the *ex vivo* human skin permeability data. The HSVR model showed consistent performance with the experimental data and mock test, which was designated to mimic the real challenges. In addition, HSVR exhibited better prediction performance than *SwissADME*, *SPC*, and *pkCSM*. Thus, it can be concluded that this novel HSVR model can be utilized to facilitate the assessment of skin permeability of the novel compounds in drug discovery.

Keywords

skin permeability; *ex vivo* excised human skin; hierarchical support vector regression (HSVR); quantitative structure-activity relationship (QSAR);

1. Introduction

- The skin has the largest surface and accounts 15% of adult body weight.
- Topical and transdermal drug delivery become an attractive and preferred route of therapeutic delivery due to its noninvasive nature and more desirable safety profiles.
- *Ex vivo* excised human skin is still considered as the gold standard for evaluating the skin permeability.
- *In vitro* skin permeability is normally defined by the permeability coefficient or constant (K_p) as follows:

$$K_p = \frac{J_{ss}}{\Delta C_v}$$

J_{ss}: the steady state flux
ΔC_v: the chemical concentration difference

In silico modeling

- Advantages:

- Less time-consuming, economically efficient, no ethical issues.
- Applicable to the virtual compounds.

- Model:

- To date, most of the quantitative *in silico* models are constructed by the linear regression or machine learning (ML) schemes.

The purpose:
A novel skin permeability model based on hierarchical support vector regression (HSVR) scheme was evaluated and compared with the other available models such as *SPC*, *SwissADME*, and *pkCSM*.

QSAR and HSVR scheme

- **SVM, SVR**

- Vapnik *et al.* proposed a support vector machine (SVM) for classification and further modified SVM for regression, termed as support vector regression (SVR).
- SVR functions by nonlinearly transferring the input into a higher-dimension space where linear regression is conducted.

- **HSVR**

- A SVR ensemble (SVRE) is compiled by assembling a pool of SVR models, which are generated based on different descriptor combinations to represent various local models with distinct application domains (ADs).

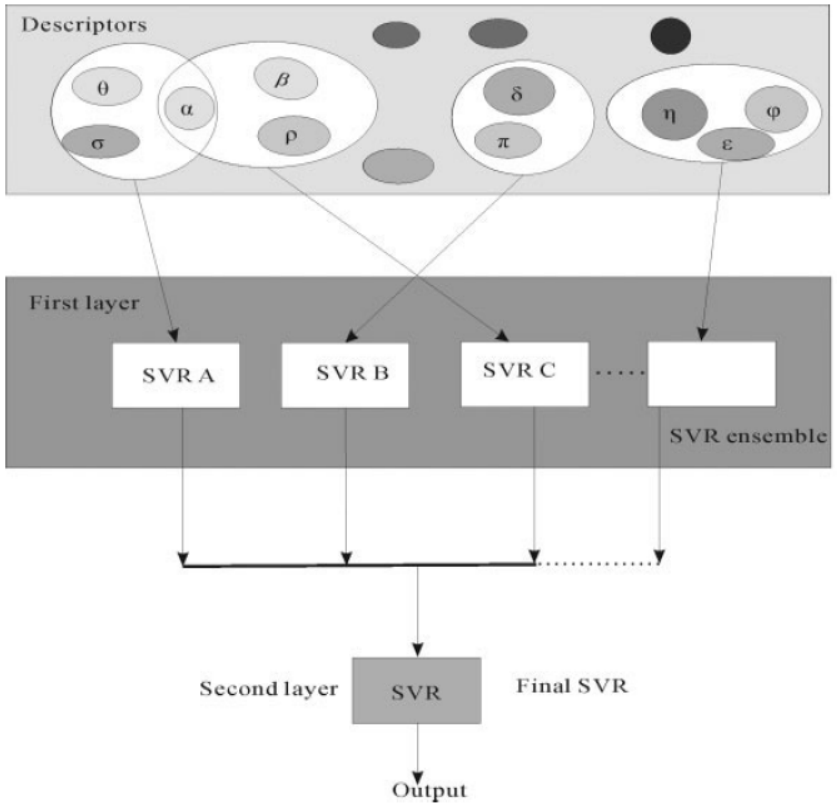


Figure 1. Schematic presentation of HSVR architecture.

2. Materials and Methods

- HSVR model was built and evaluated based on experimental K_p values of 96 compounds were compiled from the public domain.
- Descriptors were calculated using *Discovery Studio*, *SwissADME*, *E-dragon*; structures were optimized by DFT.
- The genetic function approximation (GFA) was used to select the descriptor combinations.
- The other predictors:
  1. *Skin Permeation Calculator* (*SPC*) (https://www.cdc.gov/niosh/topics/skin/skinpermcalc.html)
  2. *SwissADME* (http://www.swissadme.ch/)
  3. *pkCSM* (https://biosig.lab.uq.edu.au/pkcsm/prediction)
- Mock test: The data were collected from Soriano-Meseguer *et al*.

3. Results and Discussion

HSVR model

- HSVR model were generated based on 3 SVR models in ensemble.
- Each model adopted two descriptors with different selection.

Table 1. Descriptor selected as the input of SVR models in the ensemble, the correlation coefficient (r) with log K_p, and their descriptions.

Descriptor	SVR A	SVR B	SVR C	r	Description
V_m	x			–0.43	Molecular volume
${}^0\chi$		x		–0.48	Molecular connectivity index of order zero
Jurs_PPSA_1			x	–0.43	Partial positive surface area
log P	x	x	x	0.42	Logarithm of the n-octanol–water partition coefficient

SPC, SwissADME, pkCSM and HSVR prediction for compounds in dataset

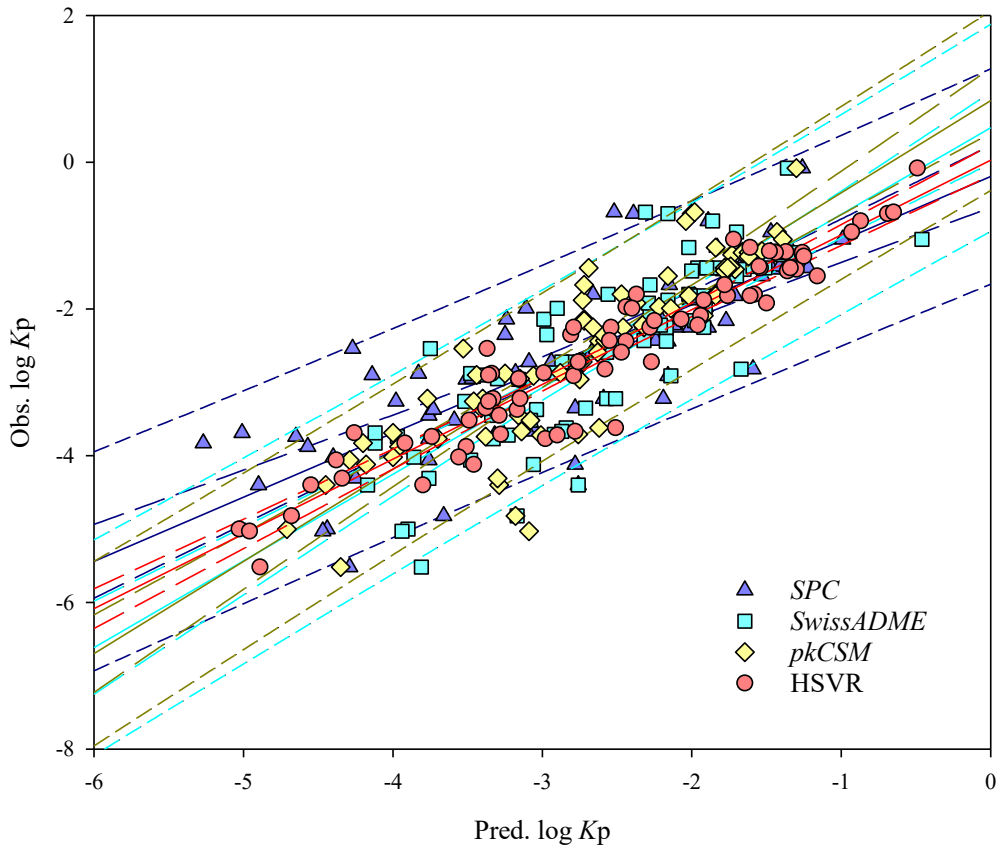


Figure 2. Observed log K_p *vs*. the predicted log K_p for the molecules in the dataset. The purple, cyan, yellow and red solid lines, dashed lines, and dotted lines correspond to *SPC, SwissADME, pkCSM*, and HSVR regressions of the data, 95% confidence intervals for *SPC*, *SwissADME*, *pkCSM*, and HSVR regressions, and 95% confidence intervals for the prediction, respectively. (p <0.01)

SPC, *SwissADME*, *pkCSM*, and HSVR comparisons

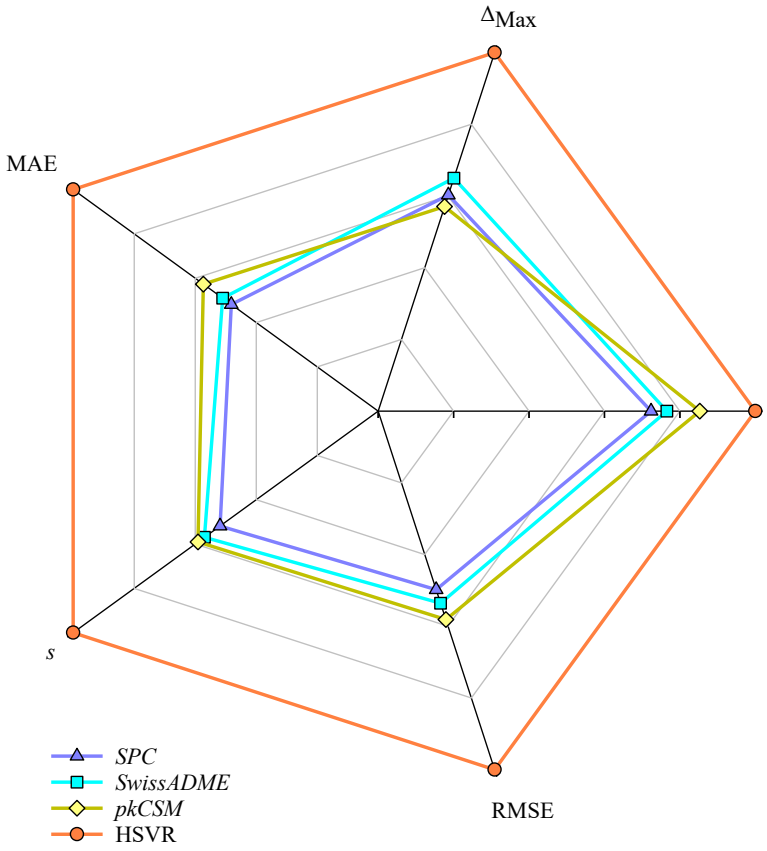


Figure 3. Radar chart of statistical evaluations of *SPC*, *SwissADME*, *pkCSM*, and HSVR.

Table 2. Statistical evaluations squared correlation coefficient (r^2), maximum error (Δ_{Max}), mean absolute error (MAE), standard deviation (s), and root mean square (RMSE), evaluated by *SPC*, *SwissADME*, *pkCSM*, and HSVR.

	SPC	*SwissADME*	*pkCSM*	*HSVR*
r^2	0.66	0.70	0.77	0.91
Δ_{Max}	1.84	1.71	1.95	1.11
MAE	0.54	0.51	0.46	0.26
s	0.48	0.44	0.42	0.25
RMSE	0.73	0.67	0.62	0.36

The HSVR model showed the best performance as compared with the other prediction models.

Prediction for the mock test - common compounds

- Mock test data: Soriano-Meseguer *et al*. measured the effective permeability coefficient (P_e) based on the skin-PAMPA system.
- Mock test: 23 common compounds + 23 novel compounds.

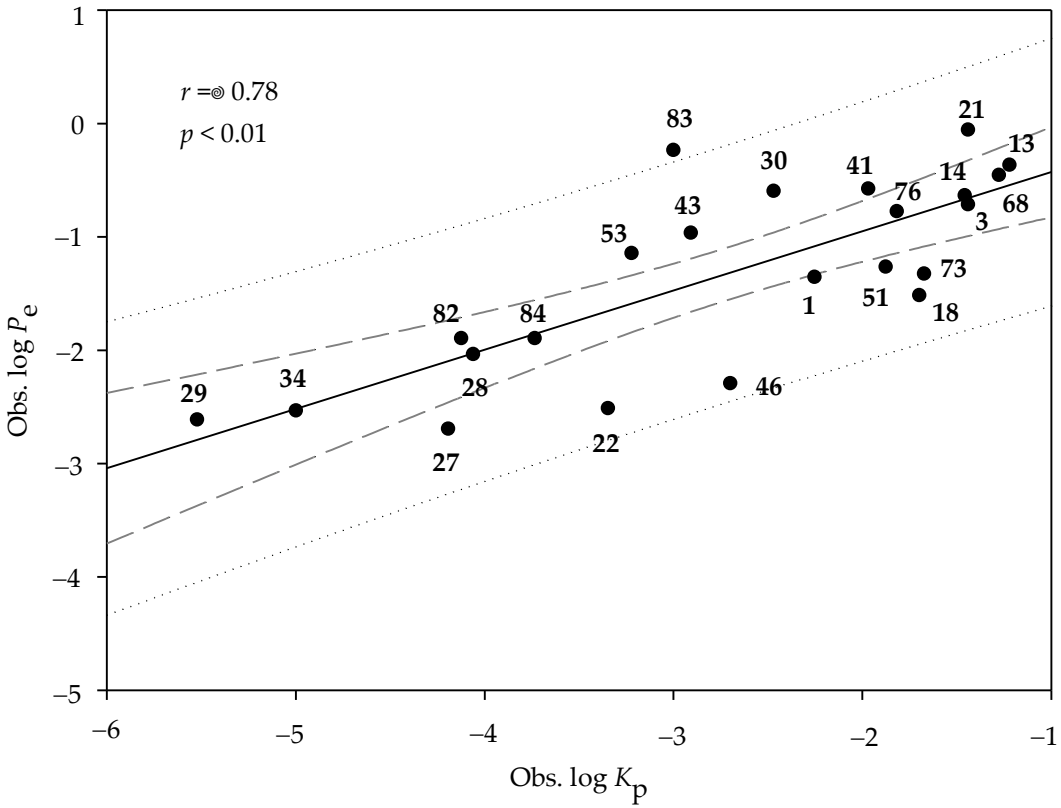


Figure 4. Observed log P_e *vs*. observed log K_p for the common compounds in the mock test.

Prediction for the mock test - novel compounds

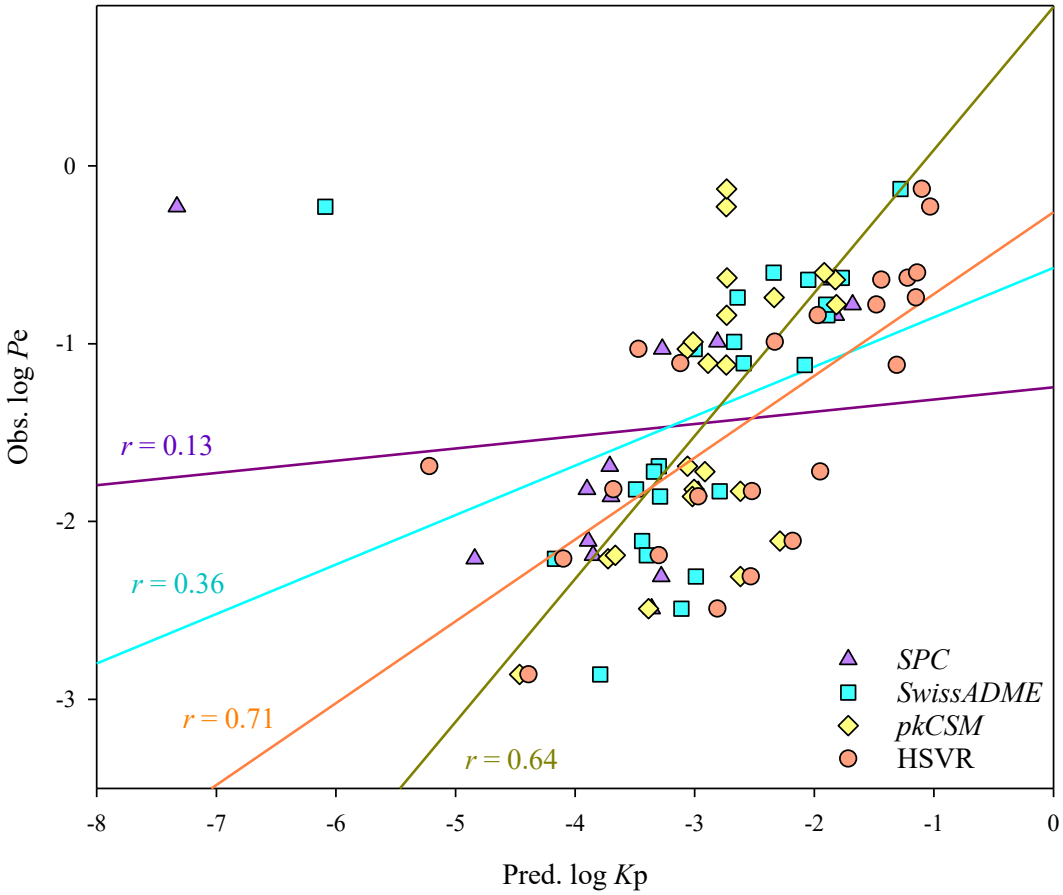


Figure 5. Observed log K_p *vs*. the predicted log K_p for the molecules in the mock test. The purple, cyan, yellow and red solid lines correspond to *SPC*, *SwissADME*, *pkCSM*, and HSVR regressions of the data, respectively. ($p < 0.01$)

4. Conclusion

- The unique architectures of HSVR retain the advantageous features of a local model and a global model, namely broader applicability domain. Therefore, the built HSVR model exhibited exceptional performance every aspects compared with the other available models.
- HSVR model presented the consistent results with different datasets, suggesting that the HSVR is the potential models to quantitatively predict skin permeability.
- Of various online predictors, *pkCSM* executed better than the others as manifested by those statistic parameters as well as the mock test, where *SPC* failed to predict some of compounds, rendering its limited applicability.

Acknowledgments

Ministry of Science and Technology, Taiwan.

References

Cortes, C. and V. Vapnik (1995). "Support-vector networks." Mach. Learn. 20(3): 273-297.

Gramatica, P., N. Chirico, E. Papa, S. Cassani and S. Kovarich (2013). "QSARINS: A new software for the development, analysis, and validation of QSAR MLR models." **34**(24): 2121-2132.

Leong, M. K., Y.-M. Chen and T.-H. Chen (2009). "Prediction of human cytochrome P450 2B6-substrate interactions using hierarchical support vector regression approach." J. Comput. Chem. **30**(12): 1899-1909

Wu, Y.-W., G. H. Ta, Y.-C. Lung, C.-F. Weng and M. K. Leong (2022). "In Silico Prediction of Skin Permeability Using a Two-QSAR Approach." Pharmaceutics **14**(5): 961.