

RESEARCH

Open Access



# In silico analysis for factors affecting anti-malarial penetration into red blood cells

Natapol Pornputtpong<sup>1,2,3</sup> , Bovornpat Suriyapakorn<sup>4</sup>, Anchisa Satayamapakorn<sup>4</sup>, Kanidsorn Larpadisorn<sup>4</sup>, Pariyachut Janviriyakul<sup>4</sup> and Phisit Khemawoot<sup>5,6\*</sup> 

## Abstract

**Background:** Malaria is a parasitic disease that produces significant infection in red blood cells. The objective of this study is to investigate the relationships between factors affecting the penetration of currently available anti-malarials into red blood cells.

**Methods:** Fifteen anti-malarial drugs listed in the third edition of the World Health Organization malaria treatment guidelines were enrolled in the study. Relationship analysis began with the prioritization of the physicochemical properties of the anti-malarials to create a multivariate linear regression model that correlates the red blood cell penetration.

**Results:** It was found that protein binding was significantly correlated with red blood cell penetration, with a negative coefficient. The next step was repeated analysis to find molecular descriptors that influence protein binding. The coefficients of the number of rotating bonds and the number of aliphatic hydrocarbons are negative, as opposed to the positive coefficients of the number of hydrogen bonds and the number of aromatic hydrocarbons. The p-value was less than 0.05.

**Conclusions:** Anti-malarials with a small number of hydrogen bonds and aromatic hydrocarbons, together with a high number of rotatable bonds and aliphatic hydrocarbons, may have a higher tendency to penetrate the red blood cells.

**Keywords:** Antimalarials, Red blood cell, Penetration

## Background

Malaria is an infectious disease generated by *Plasmodium* spp., which continues to be a public health problem in Thailand. The 2018 Thai guidelines for the treatment of malaria recommend artemisinin-based combination therapy as the first-line regimen [1]. Currently, the first-generation artemisinin derivatives, including artemisinin, artemether, arteether, artesunate, and dihydroartemisinin, are still widely used [2]. Each derivative penetrates the red

blood cells differently and has a distinctive ability to kill malaria parasites [3, 4]. This study aims to determine the factors that confer a different capability to enter the red blood cells. Therefore, we selected 15 anti-malarial drugs according to the World Health Organization (WHO) malaria treatment guidelines for this study [5]. These include artemisinin, dihydroartemisinin, artemether, arteether, artesunate, chloroquine, mefloquine, primaquine, amodiaquine, piperaquine, quinine, sulfadoxine, pyrimethamine, doxycycline, and proguanil. The screening procedures here identified the three most influential physicochemical parameters that could affect erythrocyte penetration. Information obtained from this study would be beneficial for the development of new anti-malarial drugs that are more effective in penetrating red blood cells.

\*Correspondence: phisit.khe@mahidol.ac.th

<sup>6</sup> Chakri Naruebodindra Medical Institute, Faculty of Medicine Ramathibodhi Hospital, Mahidol University, Bang Phli, Samut Prakarn 10540, Thailand

Full list of author information is available at the end of the article



© The Author(s) 2020. 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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

**Table 1 Physicochemical parameters of anti-malarial drugs**

Drug	Molecular weight (g/mol)	RBC ratio	LogP [8]	Protein binding (%)	Water solubility (mg/L)
Amodiaquine	355.8 [6]	1.313 [7]	5.1792	92 [9]	2.83 [6]
Arteether	312.41 [6]	0.23 [10]	3.2309	78.7 [11]	1700 [12]
Artemether	298.38 [6]	0.28 [10]	2.8408	95.4 [6]	12.1 [6]
Artemisinin	282.34 [6]	0.49 [10]	2.3949	82.5 [13]	51.9 [14]
Artesunate	384.18 [6]	0.71 [10]	2.6024	88 [15]	56.2 [6]
Chloroquine	319.88 [6]	4.8 [16]	4.8106	52.5 [6]	0.14 [6]
Dihydroartemisinin	284.35 [6]	0.52 [10]	2.1867	82 [17]	3160 [18]
Doxycycline	444.44 [6]	0.37 [19]	0.5458	90 [20]	630 [6]
Mefloquine	378.32 [6]	2.8 [21]	4.4479	98 [6]	6.212 [6]
Piperaquine	535.52 [6]	1.5 [22]	5.4241	97 [23]	0 [24]
Primaquine	259.35 [6]	1 [25]	2.7827	75 [26]	1300 [6]
Proguanil	253.73 [20]	4.9 [27]	2.209	75 [20]	156 [20]
Pyrimethamine	248.71 [6]	0.42 [28]	2.3542	87 [29]	121 [6]
Quinine	324.42 [20]	1.89 [30]	3.1732	70 [20]	334 [20]
Sulfadoxine	310.33 [6]	0.163 [31]	0.8768	90 [32]	296 [20]

RBC, red blood cell

**Table 2 Molecular descriptors of anti-malarial drugs**

Drug	Rot	HBA	HBD	Ali carbo	Ali het	Aro carbo	Aro het	Sat carbo
Amodiaquine	6	4	2	0	0	2	1	0
Arteether	2	5	0	1	4	0	0	1
Artemether	1	5	0	1	4	0	0	1
Artemisinin	0	5	0	1	4	0	0	1
Artesunate	4	7	1	1	4	0	0	1
Chloroquine	8	3	1	0	0	1	1	0
Dihydroartemisinin	0	5	1	1	4	0	0	1
Doxycycline	2	9	7	3	0	1	0	1
Mefloquine	2	3	2	0	1	1	1	0
Piperaquine	6	6	0	0	2	2	2	0
Primaquine	6	4	2	0	0	1	1	0
Proguanil	2	2	5	0	0	1	0	0
Pyrimethamine	2	3	3	0	0	1	1	0
Quinine	4	4	1	0	3	1	1	0
Sulfadoxine	5	7	2	0	0	1	1	0

Rot, number of rotatable bonds; HBA, number of hydrogen bond acceptors; HBD, number of hydrogen bond donors; AliCarbo, number of aliphatic carbocycles; AliHet, number of aliphatic heterocycles; AroCarbo, number of aromatic carbocycles; AroHet, number of aromatic heterocycles; and SatCarbo, number of saturated carbocycles

## Methods

### Data collection

The WHO's malaria treatment guidelines recommend 15 anti-malarial drugs for first-line malaria treatment. Structures and molecular weights were mostly retrieved from the PubChem database [6, 7]. The structures were downloaded in InChI and SMILES formats, which were

more convenient for molecular descriptor calculation. Protein binding, water solubility, and red blood cell to plasma drug concentration ratio were gathered from various sources as shown in Table 1. These parameters were converted into numeric data for statistical analysis. The charge state properties of drugs and the acid–base characteristics were not explicitly described in the learning

model, but implicated in the hypothesis as LogP and protein binding, which were included in the model and subject to the charge and acid–base characteristics of drugs. There is also a possibility of active drug transporters, but the information is sparse and limited; therefore, this feature was not included in the model.

### Molecular descriptor calculation

Chemical structure, which is a graphical notation of the compound, is complicated to use in mathematical calculations. To make computation feasible, a molecular descriptor is created. A molecular descriptor is a numerical notation associated with the chemical constitution; it is used in machine learning for correlation calculations for a compound regarding physical properties and biological activities. LogP, Number of rotatable bonds (Rot), number of hydrogen bond acceptors (HBA), number of hydrogen bond donors (HBD), number of aliphatic carbocycles (AliCarbo), number of aliphatic heterocycles (AliHet), number of aromatic carbocycles (AroCarbo), number of aromatic heterocycles (AroHet), and number of saturated carbocycles (SatCarbo) were calculated using a python package RDKit [8] as shown in Table 2.

### Data analysis

Several independent variables, including physicochemical properties and chemical descriptors, were retrieved. Using all variables to fit the model might lead to overfitting. Variables were selected based on relative importance, which was calculated using the *relaimpo* package in R programming language [33] for each independent variable. The relative importance is a comparative score among independent variables themselves to rank the effect of changing the variables to the right prediction.

An extreme gradient boosting tree regression, a non-linear regression method from Extreme Gradient Boosting (XGBoost) library [34], was used to describe the relationship a drug to its red blood cell to plasma drug concentration ratio. A non-linear regression method was chosen to challenge with a non-linear property of pharmacokinetic distribution processes, which might be a part of the relationship between physicochemical properties of the drug and red blood cell distribution. The first model was computed using relative-importance-selected physicochemical properties as independent variables and red blood cells to plasma drug concentration ratio as the dependent variable. The model was optimized using a black-box optimization method implemented in the Optuna framework [35] objected to mean squared error (MSE). The optimized model was evaluated using a permutation test [36] with five k-fold for 1000 rounds before fitting. The fitted extreme gradient boosting tree regression model was further analysed using a SHapley

Additive exPlanations (SHAP) algorithm [37] for unveiling relationship of each independent variable to the dependent variable.

A linear regression model was used to describe the relationship of molecular descriptors of anti-malarial drugs to its protein binding. Before fitting the model, a boxplot and a correlation plot were constructed, and independent variables were tested for normality using the Shapiro–Wilks test. The model was fitted into multiple linear regression models, as shown in Eq. 1 by *lm* function. The statistical analyses were performed using R version 3.4.0 [38]. The second model was trained using relative-importance-selected molecular descriptors of anti-malarial drugs as independent variables and protein binding as dependent variables to elucidate the structural features and relationships among them. Multiple R-squared, adjusted R-squared, and F-statistics were calculated by *lm* function after the model was fitted.

$$Y = \beta_0 + \sum_{i=1}^n \beta_i x_i + \varepsilon \quad (1)$$

Equation 1, The multiple linear regression model. *Y* represents an independent variable.  $\beta_0$  represents an intercept.  $\beta_i$  represents a regression coefficient.  $\varepsilon$  represents error.

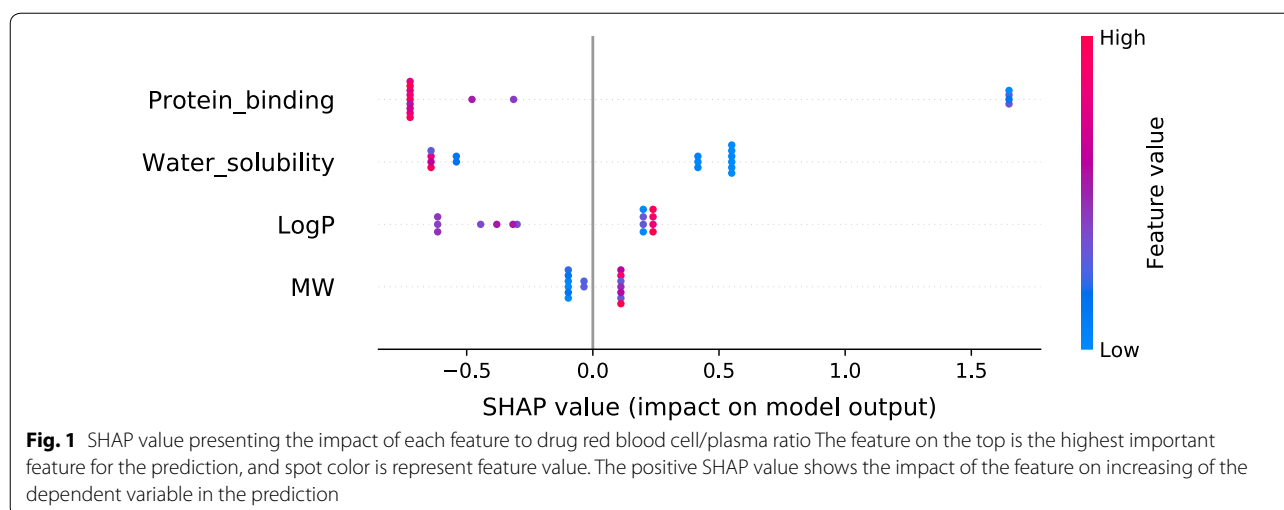
## Results

### Study on the relationships between factors influencing the red blood cell penetration of anti-malarial drugs

From the important factors analysis, protein binding, logP, water solubility, and molecular weight were considered as the most important factors and used as a feature set for extreme gradient boosting model construction. The model was optimized, then the permutation test with five-fold validation was performed. The average MSE and R-squared were 1.90 and 0.27, with *p*-0.037 and 0.009, respectively, showed a statistically significant model. From a summary plot from the SHAP algorithm, protein binding was shown to be the most important factor for red blood cell penetration properties prediction. The higher value of protein binding impacts lowering the drug red blood cell/plasma ratio, as shown in Fig. 1.

### Study of the relationships between the molecular descriptors affecting protein binding

From the important factors analysis, the number of rotatable bonds was the most important, followed by the number of hydrogen bond acceptors. Then, testing for the cross-validation and statistical significance of the correlation coefficients showed that the number of rotatable bonds, hydrogen bond acceptors, aliphatic hydrocarbons,



**Table 3** Multiple linear regression of anti-malarial drugs and their abilities in protein binding

Parameters	Estimate	SE	t-value	p-value
(Intercept)	0.599	0.098	6.123	0.000174 *
Rotatable bond	-0.047	0.014	-3.381	0.00811 *
Hydrogen bond acceptor	0.052	0.020	2.584	0.029 *
Aliphatic hydrocarbon	-0.148	0.063	-2.337	0.044 *
Aromatic hydrocarbon	0.183	0.072	2.517	0.033 *
Saturated hydrocarbon	0.185	0.142	1.297	0.227

Protein binding =  $-0.047 \times \text{Rotatable bond} + 0.052 \times \text{Hydrogen bond acceptor} - 0.148 \times \text{Aliphatic hydrocarbon} + 0.183 \times \text{Aromatic hydrocarbon} + 0.185 \times \text{Saturated hydrocarbon} + 0.599$

Residual standard error: 0.083 on 9 degrees of freedom

Multiple R-squared: 0.692, Adjusted R-squared: 0.521

F-statistic: 4.05 on 5 and 9 DF, p-value: 0.033 \*

\* Significance value at 0.05

and aromatic hydrocarbons were significant factors correlated with protein binding as shown in Table 3.

## Discussion

Machine learning is a powerful approach that widely used in many fields in the sciences for finding valuable information from data. The aims of a machine learning model development can be both to build a robust predictive model and to explain a relationship of features to outcomes. To create a predictive model needs a vast dataset to be learned by the model. While the anti-malarial drug is orphan, so the data of the drug is limited. Thus, the objective of this analysis was to investigate features that could involve drug-red blood cell partition, not to build a robust predictive model due to a limitation of data.

From the extreme gradient boosting regression of anti-malarial drugs and their abilities in red blood cell penetration, the R-squared was 0.27. Also, the multiple

linear regression of anti-malarial drugs, and their abilities in protein binding Adjusted R-squared was 0.521. These could illustrate that the predictive power of the model is incompetent. However, we can find the essential feature protein binding and some statistically significant chemical descriptors from the model, which demonstrate the relationship of them to the drug-red blood cell partition. This conclusion might lead to new potential substances that can protect against malaria in the future.

According to the analysis of factors affecting penetration of 15 anti-malarial drugs into red blood cells, we found that protein binding dominantly affects the penetration. Low protein binding causes an increased level of free drug in plasma, allowing the drug to distribute and penetrate into red blood cells. This finding is consistent with the hypotheses in previous studies of different drugs. A study of cyclosporin A revealed that the level of free drug was directly related to the concentration of the drug in red blood cells, in similar manner to another study of phenytoin [39–41]. Moreover, analysis of molecular descriptors affecting protein-binding property showed that the number of rotatable bonds, hydrogen bond acceptors, aliphatic hydrocarbons and aromatic hydrocarbons was significantly related to the protein-binding property of the drug. This property decreased with a lack of hydrogen bond acceptors and aromatic hydrocarbons; on the other hand, it increased with a lack of rotatable bonds and aliphatic hydrocarbons.

Approximately 50% of the protein in plasma is albumin. This protein plays an important role in binding to unbound drugs in plasma. There are two major binding sites in the albumin structure. The first site tends to fit with large drug molecules, while the other one is less flexible and stereo specifically bound to the drug [41]. It is implied that the drug with large size and less

flexibility has higher ability to bind to a protein. In this study, a molecule containing a higher number of rotatable bonds had less ability to bind to plasma proteins, as the molecule was flexible. The number of hydrogen bond acceptors is directly related to protein-binding property; thus, the fewer hydrogen bond acceptors, the higher the red blood cell penetration. The study of Samari et al. found that Van der Waals forces and hydrogen bonds were dominant in the binding between amodiaquine and albumin in plasma [42]. The results presented here are also consistent with a previous study which found that drugs with a low tendency to create hydrogen bonds had increased penetration into red blood cells [43]. As for the number of aliphatic and aromatic hydrocarbons, molecules with a high number of aliphatic hydrocarbons and a low number of aromatic hydrocarbons would have decreased protein-binding property, facilitating penetration into red blood cells. This concept was mentioned in a previous study; a drug containing not more than two aromatic hydrocarbons will have more unbound drug in plasma than a drug containing more than two aromatic hydrocarbons. It will also tend to bypass metabolism in the liver, leading to high concentration of the drug in plasma [44].

In terms of pharmacokinetics and pharmacodynamics, the efficacy of an antimicrobial drug generally depends on its concentration and duration of exposure. Likewise, the efficacy of artemisinin derivatives was most related to its maximum concentration in plasma [23]. More unbound drug in the plasma would be a factor that could lead to a higher concentration of the drug at the targeted site, which for an anti-malarial drug is the red blood cell. Accumulation of the drug in red blood cells increased its half-life and consequently increased the efficacy of the drug actions. In a practical aspect, anti-malarial drugs containing higher numbers of rotatable bonds and aliphatic hydrocarbons, and lower numbers of hydrogen bond acceptors and aromatic hydrocarbons, would have less protein-binding property. Therefore, more drug will penetrate through the red blood cells, facilitating its pharmacodynamic activities.

## Conclusions

The most influential physicochemical factor for the penetration of anti-malarial drugs into red blood cells is protein binding. The less a drug is bound to protein, the more it is available in free form, which can penetrate into the red blood cell. For molecular descriptors affecting protein binding, drugs with a small number of hydrogen bond acceptors and aromatic hydrocarbons, together with a high number of rotatable bonds and aliphatic hydrocarbons, may have a higher amount of free drug in the plasma available to penetrate into the red blood cell.

## Abbreviations

AliCarbo: Number of aliphatic carbocycles; AliHet: Number of aliphatic heterocycles; AroCarbo: Number of aromatic carbocycles; AroHet: Number of aromatic heterocycles; HBA: Number of hydrogen bond acceptors; HBD: Number of hydrogen bond donors; InChI: IUPAC international chemical identifier; LogP: Partition coefficient; Pb: Protein binding; RBC: Red blood cells; RMSE: Root mean square error; Rot: Number of rotatable bonds; SatCarbo: Number of saturated carbocycles; SE: Standard error; SMILES: Simplified molecular input line-entry system.

## Acknowledgements

Not applicable.

## Authors' contributions

PK, BS and NP designed the study. AS, KL and PJ conducted the analysis. PK, BS, AS, KL, PJ, and NP wrote the initial draft of the manuscript. PK, BS and NP reviewed and revised the manuscript. All authors read and approved the final manuscript.

## Funding

The authors express their gratitude to the Faculty of Pharmaceutical Sciences, Chulalongkorn University for providing research fund (Grant number Phar2562-RG003) to Natapol Pornputtpong.

## Availability of data and materials

The data that support the findings of this study are available in the manuscript.

## Ethics approval and consent to participate

Not applicable.

## Consent for publication

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

## Author details

<sup>1</sup> Department of Biochemistry and Microbiology, Faculty of Pharmaceutical Sciences, Chulalongkorn University, Bangkok, Thailand. <sup>2</sup> Vaccine and Therapeutic Protein, The Special Task Force for Activating Research, Faculty of Pharmaceutical Sciences, Chulalongkorn University, Bangkok, Thailand. <sup>3</sup> Center of Excellence in Systems Biology, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand. <sup>4</sup> Department of Pharmacy Practice, Faculty of Pharmaceutical Sciences, Chulalongkorn University, Bangkok, Thailand. <sup>5</sup> Preclinical Pharmacokinetics and Interspecies Scaling for Drug Development Research Unit, Chulalongkorn University, Bangkok, Thailand. <sup>6</sup> Chakri Naruebodindra Medical Institute, Faculty of Medicine Ramathibodhi Hospital, Mahidol University, Bang Phli, Samut Prakarn 10540, Thailand.

Received: 3 November 2019 Accepted: 30 May 2020

Published online: 23 June 2020

## References

- Wilairatana P, Tangpukdee N, Krudsod S. Update management of malaria and drug resistant malaria. Bangkok: Faculty of Tropical Medicine Mahidol University; 2016. p. 1–12.
- WHO. Antimalarial drug combination therapy: report of a technical consultation. Geneva: World Health Organization; 2001.
- Li Q, Xie L, Haeberle A, Zhang J, Weina P. The evaluation of radiolabeled artesunate on tissue distribution in rats and protein binding in humans. *Am J Trop Med Hyg*. 2006;75:817–26.
- Skinner T, Manning L, Johnston W, Davis T. In vitro stage-specific-sensitivity of *Plasmodium falciparum* to quinine and artemisinin drugs. *Int J Parasit*. 1996;26:519–25.
- WHO. Guidelines for the treatment of malaria. 3rd ed. Geneva: World Health Organization; 2015.
- Kim S, Thiessen P, Bolton E, Chen J, Fu G, Gindulyte A, et al. PubChem substance and compound databases. *Nucleic Acids Res*. 2015;44:D1202–13.

7. Winstanley P, Edwards G, Orme M, Breckenridge A. The disposition of amodiaquine in man after oral administration. *Br J Clin Pharmacol*. 1987;23:1–7.
8. Landrum G. RDKit: Open-Source Cheminformatics Software. 2019.
9. Friberg Hietala S. Clinical pharmacokinetics and pharmacodynamics of antimalarial combination therapy. Thesis: The Sahlgrenska Academy, University of Gothenburg; 2009.
10. Li Q, Milhous W, Weina P. Artemisinins in malaria therapy. New York: Nova Biomedical Books; 2007.
11. Wanwimolruk S, Edwards G, Ward S, Breckenridge A. The binding of the antimalarial arteether to human plasma proteins in-vitro. *J Pharm Pharmacol*. 1992;44:940–2.
12. Ali Z, Mishra N, Baldi A. Development and characterization of arteether-loaded nanostructured lipid carriers for the treatment of malaria. *Artif Cells Nanomed Biotechnol*. 2014;44:545–9.
13. Gordi T. Clinical pharmacokinetics of the antimalarial artemisinin based if saliva sampling. PhD Dissertation, Acta Universitatis Upsaliensis, Uppsala; 2001.
14. Chambers M. ChemDplus - 63968-64-9-BLUA FEHZUWYNDE-NNWCW-BAJSA-N-Artemisinin [INN]-Similar structures search, synonyms, formulas, resource links, and other chemical information. Chem.nlm.nih.gov. 2018 (cited 27 December 2017). <https://chem.nlm.nih.gov/chemidplus/rm/63968-64-9>.
15. Pharmaceutical Guilin. Summary of product characteristics: artesunate. Guangxi: Guilin Pharmaceutical; 2013.
16. Gustafsson L, Walker O, Alvan G, Beermann B, Estevez F, Gleisner L, et al. Disposition of chloroquine in man after single intravenous and oral doses. *Br J Clin Pharmacol*. 1983;15:471–9.
17. Morris C, Duparc S, Borghini-Fuhrer I, Jung D, Shin C, Fleckenstein L. Review of the clinical pharmacokinetics of artesunate and its active metabolite dihydroartemisinin following intravenous, intramuscular, oral or rectal administration. *Malar J*. 2011;10:263.
18. Wishart DS, Feunang YD, Marcu A, Guo AC, Liang K, et al. HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res*. 2018;46:D608–17.
19. Deshmukh P, Badgujar P, Gatne M. In-vitro red blood cell partitioning of doxycycline. *Indian J Pharmacol*. 2009;41:173.
20. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant J, et al. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res*. 2017;46:D1074–82.
21. Vieira J, Borges L, Ferreira M, Rivera J, Gomes M. Patient age does not affect mefloquine concentrations in erythrocytes and plasma during the acute phase of falciparum malaria. *Braz J Infect Dis*. 2016;20:482–6.
22. Hung T, Davis T, Ilett K. Measurement of piperazine in plasma by liquid chromatography with ultraviolet absorbance detection. *J Chromatogr B Analyt Technol Biomed Life Sci*. 2003;791:93–101.
23. Li Q, Hickman M. Pharmacokinetics and pharmacodynamics of antimalarial drugs used in combination therapy. Sharjah: Bentham Science Publisher; 2015.
24. Piperazine. U.S. Environmental Protection Agency Chemistry Dashboard. (cited 26 December 2017). <https://comptox.epa.gov/dashboard/DTXSID00193825>.
25. World Health Organization. Methods and techniques for assessing exposure to antimalarial drugs in clinical field studies. Geneva: World Health Organization; 2011.
26. Primaquine phosphate. International Programme on Chemical Safety. 1994 (cited 28 December 2017). <http://www.inchem.org/documents/pims/pharm/primaqui.htm>.
27. Edstein M, Veenendaal J, Scott H, Rieckmann K. Steady-state kinetics of proguanil and its active metabolite, cycloguanil, in man. *Chemotherapy*. 1988;34:385–92.
28. Rudy A, Poynor W. Binding of pyrimethamine to human plasma proteins and erythrocytes. *Pharm Res*. 1990;7:1055–60.
29. Drug, OTCs & Herbals. Medscape Reference. 2017 (cited 26 December 2017). <https://reference.medscape.com/drugs>.
30. Vieira J, Gomes A, Borges L, Guimarães E. Relationship between plasma and red blood cell concentrations of quinine in Brazilian children with uncomplicated *Plasmodium falciparum* malaria on oral therapy. *Rev Inst Med Tropical de São Paulo*. 2009;51:109–10.
31. Berneis K, Boguth W. Distribution of sulfonamides and sulfonamide potentiators between red blood cells, proteins and aqueous phases of the blood of different species. *Chemotherapy*. 1976;22:390–409.
32. Florey K. Analytical profiles of drug substances. San Diego: Academic Press; 1988.
33. Grömping U. Relative importance for linear regression in R: the Package relaimpo. *J Stat Softw*. 2006;17:1–27.
34. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016; 13–17-Aug. pp. 785–94.
35. Akiba T, Sano S, Yanase T, Ohta T, Koyama M. Optuna: a next-generation hyperparameter optimization framework. Proceedings of the ACM SIGKDD International Conference on knowledge discovery and data mining. 2019. pp. 2623–31.
36. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. *J Mach Learning Res*. 2011;12:2825–30.
37. Lundberg S, Lee S-I. A unified approach to interpreting model predictions. *Adv Neural Inform Processing Syst*. 2017;2017:4766–75.
38. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2014. <http://www.R-project.org/>.
39. Shibata N, Shimakawa H, Minouchi T, Yamaji A. Erythrocyte uptake and protein binding of cyclosporin A (CyA) in human blood: factors affecting CyA concentration in erythrocytes. *Biol Pharm Bull*. 1993;16:702–7.
40. Kurata D, Wilkinson G. Erythrocyte uptake and plasma binding of di phenyl hydantoin. *Clin Pharm Ther*. 1974;16:355–62.
41. Quinlan G, Martin G, Evans T. Albumin: biochemical properties and therapeutic potential. *Hepatology*. 2005;41:1211–9.
42. Samari F, Shamsipur M, Hemmateenejad B, Khayamian T, Gharaghani S. Investigation of the interaction between amodiaquine and human serum albumin by fluorescence spectroscopy and molecular modeling. *Eur J Med Chem*. 2012;54:255–63.
43. Fagerholm U. Prediction of human pharmacokinetics—evaluation of methods for prediction of volume of distribution. *J Pharm Pharmacol*. 2007;59:1181–90.
44. Ritchie T, Macdonald S, Young R, Pickett S. The impact of aromatic ring count on compound developability: further insights by examining carbo- and hetero-aromatic and -aliphatic ring types. *Drug Discov Today*. 2011;16:164–71.

## Publisher's Note

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