DOI 10.18699/SBB-2022-203

Prediction of adverse effects of drug-drug interactions on the cardiovascular system based on the analysis of structure-activity relationships

Sukhachev V.*, Ivanov S., Dmitriev A.
Institute of Biomedical Chemistry, Moscow, Russia
** withstanding@yandex.ru*

Key words: drug-drug interactions, SAR, adverse drug reactions, polypharmacy

Motivation and Aim: The combination of several drugs can lead to a number of diverse side effects, which may be followed by certain risks to life and health. The drug effects on the cardiovascular system are the most serious ones, since cardiovascular diseases are the leading cause of death worldwide [1]. Experimental determination of all possible drug interactions seems impracticable, given the many probable combinations of drugs. The aim of our study is to build a structure-activity relationships (SAR) models for predicting cardiovascular drug interactions.
Methods and Algorithms: Cardiovascular-related and non-cardiovascular data were extracted from the DrugBank [2] and TwoSides [3] databases. Overall, 283,439 drug-pairs were obtained, including 2,313 specific drugs causing 113 side effects, as well as 62,351 additional drug-pairs for negative examples which were opposed to pairs that cause a certain undesirable effect. Thus, we unambiguously categorize data into those that cause an effect and those that do not. The Pearson correlation coefficients (PCC) for each pair of side effects were calculated. Based on PCC values, building a model for blood pressure disorders, we excluded from negative examples drug-pairs related to blood pressure: hypotension, hypertension, vasodilation, vasoconstriction, etc.
The description of drug-pairs was performed using two types of descriptors:
PoSMNA structural descriptors [4, 5]. It is a linear combination of structure fragments for two compounds pivoted in table when each row is a drug-pairs and each column is an appearance of concrete structures fragments in this pair. Totally, the data contains 345,790 rows of drug-pairs and 15,000 columns of structural elements present in each pair.
Linear combinations of estimates of the probability of biological activity obtained with PASS [6]. Totally, data contains 345,790 rows and 14032 columns for each activity.
Also, the most optimal parameters of these descriptors were determined empirically, providing the greatest accuracy of prediction.
Classification models were built by Random Forest followed by 5-fold cross-validation for six most major side effects: arrhythmia, bradycardia, QT-interval prolongation, tachycardia, hypotension and hypertension. PASS descriptors corresponding to the activities that contributed most to the accuracy of the model were extracted to build a network of proteins that are involved in the mechanism of drug side effects.
Results: Accuracy parameters obtained for 12 models is given in the table below.

PoSMNA						
	Bradycardia	arrhythmia	QT	Tachycardia	Hypertension	Hypotension
AUC	0,91	0,88	0,88	0,93	0,94	0,89
SENS	0,71	0,69	0,69	0,73	0,75	0,67
SPEC	0,97	0,99	0,91	0,98	0,95	0,91

PASS						
	Bradycardia	arrhythmia	QT	Tachycardia	Hypertension	Hypotension
AUC	0,94	0,90	0,90	0,96	0,91	0,89
SENS	0,65	0,55	0,76	0,70	0,85	0,77
SPEC	0,97	0,96	0,86	0,97	0,82	0,85

The average estimates of the accuracy parameters are AUC: 91 %, sensibility: 71 %, specificy: 95 %. A high AUC value indicates that the models can produce a binary classification with high accuracy. Close values of sensitivity and specificity indicate that the model makes it possible to distinguish with almost equal accuracy a pair of drugs that cause this effect and those that do not.

Also, protein-protein interactions networks were built based on the descriptors that have the greatest contribution to the accuracy of the model.

Conclusion: The constructed SAR models make it possible to predict a number of undesirable drug interactions with high accuracy. The characteristics of the models are interpretable and may be used to uncover the mechanisms of drug-drug interactions. The accuracy parameters obtained by two different descriptors are very close. This justifies that the chosen mode of describing a pair of drugs make it possible to make a prediction.

Acknowledgements: The authors are grateful for the support by the Russian Science Foundation Grant (Grant No. 17-75-20250) of this study.

References

1. The top 10 causes of death URL [https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death]. Accessed on 24.04.2022.
2. Wishart D. et al. *Nucleic Acids Res*. 2018;46(1):D1074-D1082.
3. Offsides and Twosides (NSIDES v0.1) – TLab URL [http://tatonettilab.org/offsides]. Accessed: 24.04.2022.
4. Filimonov D.A. et al. *Chem Heterocycl Compd*. 2014;50(3):444-457.
5. Filimonov D.A. et al. *J Chem Inf Comput Sci.* 1999;39(4):666-670.
6. Filimonov D.A. et al. *Biomed Chem Res Methods.* 2018;1(1):e00004.