UDC 615.07:544.4

Yaroslava PUSHKAROVA
Candidate of Chemical Sciences, Associate Professor at the Department of Analytical, Physical and Colloid Chemistry, Bogomolets National Medical University, Beresteyskyi Aveneu, 34, Kyiv, Ukraine, 01601 (yaroslava.pushkarova@gmail.com)
ORCID: *0000-0001-9856-7846*
SCOPUS: *55375914700*

Galina ZAITSEVA
Candidate of Chemical Sciences, Head of Department of Analytical, Physical and Colloid Chemistry, Bogomolets National Medical University, Beresteyskyi Aveneu, 34, Kyiv, Ukraine, 01601 (galinazaitseva777@gmail.com)
ORCID: *0000-0003-3138-6324*
SCOPUS: *58554904800*

Sedat ACIK
Master of Pharmacy, Bogomolets National Medical University, Beresteyskyi Aveneu, 34, Kyiv, Ukraine, 01601 (sedat3334@hotmail.com)

To cite this article: Pushkarova Ya., Zaitseva G., Acik S. (2025). Analiz biofarmatsevtychnoi klasyfikatsiinoi systemy iz zastosuvanniam khemometrychnykh metodiv [Analysis of Biopharmaceutics Classification System by means of chemometric methods]. *Fitoterapiia. Chasopys – Phytotherapy. Journal,* 3, 205–209, doi: https://doi.org/10.32782/2522-9680-2025-3-205

ANALYSIS OF BIOPHARMACEUTICS CLASSIFICATION SYSTEM BY MEANS OF CHEMOMETRIC METHODS

Actuality. *The integration of the Biopharmaceutics Classification System with chemometric methods represents a promising direction in pharmaceutical research. Biopharmaceutics Classification System categorizes drugs into four classes based on their solubility and permeability. Predicting these properties using chemometric models can reduce reliance on in vitro and in vivo experiments, thereby streamlining the drug development process. Combining* ***Biopharmaceutics Classification System*** *with chemometric techniques, including statistical analyses and neural network models, creates new opportunities for predicting bioavailability, optimizing drug formulations, and supporting regulatory decision-making. This approach is therefore highly relevant for both scientific inquiry and practical pharmaceutical applications.*

Aim of research. The aim of this study is to develop an accurate drug classification model in accordance with the Biopharmaceutics Classification System using chemometric methods.

Material and methods. *The dataset includes 122 drug compounds characterized by 11 physicochemical and topological molecular descriptors. The research methods employed were the Kruskal – Wallis test and a Probabilistic Neural Network. The software packages MATLAB R2024b and ChemOffice 2020 were used in this study.*

Research results. *It was found that four descriptors (number of HBond donors, partition coefficient, solubility and polar surface area) are sufficient for accurate drug classification according to the Biopharmaceutics Classification System. The Probabilistic Neural Network architecture with a spread value of 0,1 proved to be an effective for this task.*

Conclusion. *The results demonstrate the strong potential of chemometric methods for constructing predictive models that can optimize and accelerate the drug development pipeline.*

Key words: *drug classification, molecular descriptors, pharmaceutical sciences, chemometrics.*

Ярослава ПУШКАРЬОВА
кандидат хімічних наук, доцент кафедри аналітичної, фізичної та колоїдної хімії, Національний медичний університет імені О.О. Богомольця, просп. Берестейський, 34, м. Київ, Україна, 01601 (yaroslava.pushkarova@gmail.com)
ORCID: *0000-0001-9856-7846*
SCOPUS: *55375914700*

Галина ЗАЙЦЕВА
кандидат хімічних наук, завідувач кафедри аналітичної, фізичної та колоїдної хімії, Національний медичний університет імені О.О. Богомольця, просп. Берестейський, 34, м. Київ, Україна, 01601 (galinazaitseva777@gmail.com)
ORCID: *0000-0003-3138-6324*
SCOPUS: *58554904800*

Седат АЧІК
магістр фармації, Національний медичний університет імені О.О. Богомольця, просп. Берестейський, 34, м. Київ, Україна, 01601 (sedat3334@hotmail.com)

Бібліографічний опис статті: Пушкарьова Я., Зайцева Г., Ачік С. (2025). Аналіз біофармацевтичної класифікаційної системи із застосуванням хемометричних методів. *Фітотерапія. Часопис,* 3, 205–209, doi: https://doi.org/10.32782/2522-9680-2025-3-205

АНАЛІЗ БІОФАРМАЦЕВТИЧНОЇ КЛАСИФІКАЦІЙНОЇ СИСТЕМИ ІЗ ЗАСТОСУВАННЯМ ХЕМОМЕТРИЧНИХ МЕТОДІВ

Актуальність. *Інтеграція біофармацевтичної класифікаційної системи з хемометричними методами є перспективним напрямом у фармацевтичній науці. Ця система поділяє лікарські речовини на чотири класи, залежно від їхньої розчинності та проникності. Використання хемометричних моделей для прогнозування цих властивостей дозволяє зменшити потребу в лабораторних (in vitro) та доклінічних (in vivo) експериментах, що сприяє прискоренню та здешевленню розроблення нових лікарських засобів. Поєднання* ***біофармацевтичної класифікаційної системи*** *зі статистичними методами та нейромережевими підходами відкриває нові можливості для точнішого прогнозування біодоступності, удосконалення лікарських форм і підвищення ефективності регуляторних рішень. Такий підхід має високу наукову та практичну цінність.*

Мета дослідження *– розроблення точної моделі класифікації лікарських речовин, відповідно до біофармацевтичної класифікаційної системи, з використанням хемометричних методів.*

Матеріал і методи. *У дослідженні використано набір даних, що включає 122 лікарські речовини, охарактеризовані за допомогою 11 фізико-хімічних і топологічних молекулярних дескрипторів. Основними методами дослідження були критерій Крускала – Уолліса та ймовірнісна нейронна мережа. У цьому дослідженні були використані програмні пакети MATLAB R2024b та ChemOffice 2020.*

Результати дослідження. *Установлено, що чотирьох молекулярних дескрипторів (кількість донорів водневого зв'язку, коефіцієнт розподілу, розчинність і полярна поверхнева площа) досить для точної класифікації лікарських речовин, відповідно до біофармацевтичної класифікаційної системи. Архітектура ймовірнісної нейронної мережі зі значенням параметра розсіювання 0,1 виявилася ефективною для розв'язання поставленого завдання.*

Висновок. *Отримані результати підтверджують високий потенціал хемометричних методів у побудові моделей прогнозування, здатних оптимізувати та прискорити процес розроблення лікарських препаратів.*

Ключові слова: *класифікація лікарських речовин, молекулярні дескриптори, фармацевтична галузь, хемометрія.*

Actuality. The Biopharmaceutics Classification System (BCS) offers a systematic and scientifically grounded approach to categorizing drug substances based on their aqueous solubility and intestinal permeability. By also considering the dissolution of the dosage form, the BCS evaluates three key factors regulating the rate and extent of drug absorption from solid oral formulations: dissolution, solubilization and intestinal permeability (Agnihotri, 2024; Mehta, 2023).

Chemometric techniques, including Quantitative Structure-Activity/Property Relationship (QSAR/QSPR) modeling, Principal Component Analysis and Partial Least Squares, enable both qualitative and quantitative evaluation of drug properties based on chemical structure. These methods can predict the BCS class of new compounds or refine classifications for existing drugs without extensive experimental testing, reducing development time and costs (Pushkarova, 2023).

Moreover, the rise of personalized medicine amplifies the need for precise drug formulations tailored to individual patient profiles. Enhanced BCS classification accuracy via chemometric models supports the design of formulations with optimized bioavailability, improving therapeutic outcomes and safety. The growing volume of pharmaceutical data and advances in machine learning and artificial intelligence integrated into chemometrics further increase the potential to improve prediction accuracy and discover novel drug property patterns. This integration promises to accelerate drug discovery, reduce costs and enhance regulatory compliance by supporting early-stage *in silico* assessments (Niazi, 2023).

Overall, combining the BCS with chemometric methods opens new avenues for predicting bioavailability and developing pharmaceuticals, making this research highly significant for both scientific progress and practical applications.

Aim of research. The aim of this study is to develop an accurate drug classification model in accordance with the BCS using chemometric methods.

Materials and methods. The studied dataset included 122 drug compounds, which were divided into two subsets: a training set (104 drug compounds, 85%) and a testing set (18 drug compounds, 15%). The BCS classes of these compounds were known from scientific literature, which allowed for supervised training and validation of the classification model (Bergström, 2014). The training set is used to develop the classification model, while the testing set is used to evaluate its predictive accuracy on previously unseen data.

Table 1

Results of the Kruskal – Wallis test for 11 molecular descriptors

	Descriptor										
	MW	**HBA**	**HBD**	**MR**	**logP**	**logS**	**BI**	**MTI**	**RotB**	**PSA**	**WI**
χ^2	5,91	4,41	**11,35**	5,75	**29,64**	**28,46**	5,08	6,26	2,11	**13,66**	5,88

For each compound 11 physicochemical and topological molecular descriptors were calculated using ChemOffice 2020 software: molecular weight (MW); number of HBond acceptors (HBA); number of HBond donors (HBD); molar refractivity (MR); partition coefficient (logP); solubility (logS); Balaban index (BI); molecular topological index (MTI); number of rotatable bonds (RotB); polar surface area (PSA); Wiener index (WI).

To identify descriptors with the most significant impact on drug classification, the Kruskal – Wallis test was applied (Aravind Kumar, 2024). For the classification task, a Probabilistic Neural Network (PNN) was used due to its ability to provide high accuracy, robustness to noisy data and fast convergence. The network's architecture was optimized by adjusting the spread parameter to improve classification performance (Mohebali, 2020).

The Kruskal – Wallis test and PNN model were implemented using the software package Matlab R2024b.

Research results and their discussion. The Kruskal – Wallis test was performed for 122 drug compounds characterized by 11 molecular descriptors and the results are presented in table 1. The critical value of the chi-squared (χ^2) statistic at a 5% significance level and 3 degrees of freedom is 7,81 (Miller, 2010, p. 268).

Descriptors for which the calculated χ^2 value exceeds the critical value of 7,81 are considered statistically significant and therefore informative for classification. It was found that the classification of drug compounds according to the BCS is most significantly influenced by the following four molecular descriptors: number of HBond donors, partition coefficient, solubility and polar surface area.

For the remaining descriptors, the calculated χ^2 values were lower than the critical value. Consequently, molecular weight, number of HBond acceptors, molar refractivity, Balaban index, molecular topological index, number of rotatable bonds and Wiener index were determined to be uninformative for the classification of drug compounds within the BCS.

In a Probabilistic Neural Network, the spread parameter plays a crucial role in determining the shape and width of the Gaussian function used in the pattern layer. The spread parameter controls the smoothness of the probability density function estimates. A small spread value results in a narrow Gaussian curve, making the network sensitive to small variations and prone to overfitting. A large spread value results in a wider Gaussian curve, leading to smoother decision boundaries but potentially causing underfitting. Selecting an appropriate spread value is essential for achieving a good balance between model generalization and accuracy (Mohebali, 2020).

The training set of 104 drug compounds, characterized by four descriptors (number of HBond donors, partition coefficient, solubility, polar surface area), was used to determine the optimal architecture of the PNN for accurate drug classification according to the BCS. The proportion of misclassified drug compounds at different spread parameter values is presented in table 2. This proportion was calculated as the number of misclassified compounds divided by the total number of compounds in the training set.

Table 2

Results of Probabilistic Neural Network training

Spread parameter value	Proportion of misclassified drug compounds, %
0,1–0,6	0,0
0,7	0,8
0,8–1,0	1,6

Accurate training of the PNN was achieved at spread parameter values ranging from 0,1 to 0,6, where no classification errors occurred. At a spread value of 0,7, one drug compound was misclassified. At spread values between 0,8 and 1,0, two drug compounds were misclassified.

To verify the correctness and predictive capability of the proposed model, which was based on four informative molecular descriptors and a Probabilistic Neural Network, its performance was evaluated using the testing set. The list of compounds in the testing set and the results of their classification according to the BCS are presented in table 3.

Correct classification of all 18 drug compounds in the testing set was achieved at a spread parameter value of 0,1. One drug compound (Zidovudine) was misclassified at spread values ranging from 0,2 to 1,0. Therefore, a spread parameter value of 0,1 is recommended as the optimal setting for drug classification according to the BCS.

Table 3

Results of class predictions according to the BCS for the testing set

Drug compound	Correct BCS class (Bergström, 2014)	Spread parameter value of 0,1	Spread parameter values ranging from 0,2 to 1,0
		Predicted BCS class	Predicted BCS class
Tramadol	1	1	1
Zolpidem	1	1	1
Bisoprolol	1	1	1
Venlafaxine	1	1	1
Zidovudine	1	1	**2**
Doxepin	1	1	1
Enalapril	1	1	1
Rofecoxib	2	2	2
Simvastatin	2	2	2
Celecoxib	2	2	2
Lorazepam	2	2	2
Naproxen	2	2	2
Cimetidine	3	3	3
Risperidone	2	2	2
Acebutolol	3	3	3
Atenolol	3	3	3
Ergotamine	3	3	3
Nevirapine	2	2	2

Conclusions. It was established that four molecular descriptors (number of HBond donors, partition coefficient, solubility and polar surface area) are sufficient for accurate classification of drug compounds according to the BCS.

The Probabilistic Neural Network architecture with a spread parameter value of 0.1 proved to be an effective tool for this classification task.

The results demonstrate the significant potential of chemometric methods for developing predictive models, that can optimize and accelerate the drug development process.

BIBLIOGRAPHY

Agnihotri T. G., Paradia P. K., Jain A. Biopharmaceutical Classification System: a strategic tool in pharmaceutical formulation. Dosage Forms. *Formulation Developments and Regulations.* 2024. P. 443–469. DOI: 10.1016/B978-0-323-91817-6.00016-4.

Aravind Kumar P., Sivaranjini S., Jyothir Ganesh P., Garlapati C. New solubility model to correlate solubility of anticancer drugs in supercritical carbon dioxide and evaluation with Kruskal – Wallis test. *Fluid Phase Equilibria*. 2024. № 582. P. 114099. DOI: 10.1016/j.fluid.2024.114099.

Bergström C. A., Andersson S. B., Fagerberg J. H., Ragnarsson G., Lindahl A. Is the full potential of the biopharmaceutics classification system reached? *Eur. J. Pharm. Sci.* 2014. № 57. P. 224–231. DOI: 10.1016/j.ejps.2013.09.010.

Mehta M., Polli J. E., Seo P., Bhoopathy S., Berginc K., Kristan K., ... Velagapudi R. Drug permeability-best practices for Biopharmaceutics Classification System (BCS)-based biowaivers: a workshop summary report. *J. Pharm. Sci*. 2023. № 112 (7). P. 1749–1762. DOI: 10.1016/j.xphs.2023.04.016.

Miller R. D., Miller J. C., Miller J. Statistics and Chemometrics for Analytical Chemistry. 6th ed. UK : Prentice Hall, 2010. 278 p.

Mohebali B., Tahmassebi A., Meyer-Baese A., Gandomi A. H. Probabilistic neural networks: a brief overview of theory, implementation, and application. *Handbook of probabilistic models.* 2020. P. 347–367. DOI: 10.1016/B978-0-12-816514-0.00014-X.

Niazi S. K., Mariam Z. Recent advances in machine-learning-based chemoinformatics: a comprehensive review. *Int. J. Mol. Sci.* 2023. № 24 (14). P. 11488. DOI: 10.3390/ijms241411488.

Pushkarova Y., Zaitseva G., Kaliuzhenko A. Classification of Residual Solvents by Risk Assessment Using Chemometric Methods. *13th International Conference on Advanced Computer Information Technologies (ACIT),* Wrocław, Poland, 21–23 September 2023. DOI: 10.1109/ACIT58437.2023.10275405.

REFERENCES

Agnihotri, T.G., Paradia, P.K., & Jain, A. (2024). Biopharmaceutical Classification System: a strategic tool in pharmaceutical formulation. In *Dosage Forms, Formulation Developments and Regulations* (pp. 443–469). Academic Press. DOI: 10.1016/B978-0-323-91817-6.00016-4.

Aravind Kumar, P., Sivaranjini, S., Jyothir Ganesh, P., & Garlapati, C. (2024). New solubility model to correlate solubility of anticancer drugs in supercritical carbon dioxide and evaluation with Kruskal – Wallis test. *Fluid Phase Equilibria*, *582*, 114099. DOI: 10.1016/j.fluid.2024.114099.

Bergström, C.A., Andersson, S.B., Fagerberg, J.H., Ragnarsson, G., & Lindahl, A. (2014). Is the full potential of the biopharmaceutics classification system reached? *Eur. J. Pharm. Sci.*, *57*, 224–231. DOI: 10.1016/j.ejps.2013.09.010.

Mehta, M., Polli, J.E., Seo, P., Bhoopathy, S., Berginc, K., Kristan, K., ... & Velagapudi, R. (2023). Drug permeability-best practices for Biopharmaceutics Classification System (BCS)-based biowaivers: a workshop summary report. *J. Pharm. Sci*. *112* (7), 1749–1762. DOI: 10.1016/j.xphs.2023.04.016.

Miller, R.D., Miller, J.C., Miller, J. (2010). Statistics and Chemometrics for Analytical Chemistry. 6th ed. UK: Prentice Hall, 278 p.

Mohebali, B., Tahmassebi, A., Meyer-Baese, A., & Gandomi, A.H. (2020). Probabilistic neural networks: a brief overview of theory, implementation, and application. *Handbook of probabilistic models*, 347–367. DOI: 10.1016/B978-0-12-816514-0.00014-X.

Niazi, S.K., & Mariam, Z. (2023). Recent advances in machine-learning-based chemoinformatics: a comprehensive review. *Int. J. Mol. Sci. 24* (14), 11488. DOI: 10.3390/ijms241411488.

Pushkarova, Y., Zaitseva, G., & Kaliuzhenko, A. (2023, September). Classification of Residual Solvents by Risk Assessment Using Chemometric Methods. In *2023 13th International Conference on Advanced Computer Information Technologies (ACIT)* (pp. 562–565). IEEE. DOI: 10.1109/ACIT58437.2023.10275405.

Стаття надійшла до редакції 05.05.2025
Стаття прийнята до друку 09.07.2025
Опубліковано: 15.10.2025

Conflict of interests: There is no conflict of interests related to this research study.
Authors' contribution:
Pushkarova Ya.M. – conceptualization and study design, data collection and analysis, manuscript writing;
Zaitseva G.M. – conceptualization and study design, manuscript writing, manuscript revision;
Acik Sedat – data collection and analysis, manuscript writing.
Email address for correspondence with authors: ***yaroslava.pushkarova@gmail.com***