UDC 543.33 + 004.032.26
DOI https://doi.org/10.32689/2663-0672-2024-4-16

Yaroslava PUSHKAROVA
Candidate of Chemical Sciences, Associate Professor, Associate Professor at the Department of Analytical, Physical and Colloid Chemistry, Bogomolets National Medical University, yaroslava.pushkarova@gmail.com
ORCID: 0000-0001-9856-7846

Inara TYMCHENKO
5th Year Student at the Pharmaceutical Faculty, Bogomolets National Medical University, inaraakperova5@gmail.com

PREDICTION OF ORAL DRUG BIOAVAILABILITY BASED ON CHEMICAL STRUCTURE

Prediction of bioavailability is a crucial aspect of drug development and formulation design. ***Aim.*** *To investigate the possibility of predicting the bioavailability of oral drug compounds based on molecular descriptors by means of feedforward neural network.*

Materials and methods. *Software ChemOffice 2020 was used for calculation molecular descriptors from the drug structures. Software Matlab R2022b was used for building the multivariate regression between molecular descriptors and bioavailability as well as feedforward neural network. For each of the 145 drug molecules 16 molecular descriptors were calculated. Of the 145 compounds, 10 were randomly selected for use as a test subset, 5 compounds – as a validation subset. The remaining 130 compounds were used to train feedforward neural network.*

Results and discussion. *It is established that there are 7 the most informativeness molecular descriptors for oral drug bioavailability prediction: molecular weight, hydrogen bond acceptors, hydrogen bond donors, partition coefficient, Balaban index, molecular topological index, Wiener index (multiple coefficient of determination is equal to 0,3701). Optimal number of hidden neurons for effective realization feedforward neural network was found by experimental way and its equal to 16. Trained feedforward neural network with 16 hidden neurons was used for predicting bioavailability values for test and validation subsets.*

Conclusions. *Feedforward neural network is the effective tool for prediction of oral drug bioavailability based on chemical structure. This is evidenced by high values of determination coefficient between predicted bioavailability values and experimental bioavailability values for test and validation subsets (0,7084 and 0,8432, correspondingly). Obtained results can be useful at the stage of experiment planning or drug design.*

Key words: *chemometrics, drug design, molecular descriptor, multivariate regression, neural network, pharmacy.*

Ярослава Пушкарьова, Інара Тимченко. ПРОГНОЗУВАННЯ БІОДОСТУПНОСТІ ЛІКАРСЬКИХ РЕЧОВИН ДЛЯ ПЕРОРАЛЬНОГО ЗАСТОСУВАННЯ НА ОСНОВІ ЇХ ХІМІЧНОЇ СТРУКТУРИ

Прогноз біодоступності є важливим аспектом розробки лікарських засобів. ***Мета.*** *Дослідити можливість прогнозування біодоступності лікарських речовин для перорального застосування на основі молекулярних дескрипторів за допомогою нейронної мережі прямого поширення сигналу.*

Матеріали та методи. *Розрахунок молекулярних дескрипторів був проведений із застосуванням програмного пакету ChemOffice 2020. Побудова множинної регресії між молекулярними дескрипторами та біодоступністю, а також реалізація нейронної мережі прямого поширення сигналу були здійснені із застосуванням програмного пакету Matlab R2022b. Для кожної з 145 молекул лікарських речовин було розраховано 16 молекулярних дескрипторів. Зі 145 досліджених сполук 10 були випадковим чином відібрані для використання в якості тестової вибірки, 5 сполук – в якості валідаційної вибірки, решта 130 сполук були використані для навчання нейронної мережі прямого поширення сигналу.*

Результати та обговорення*. Встановлено, що для прогнозування біодоступності лікарських речовин для перорального застосування найбільш інформативними та доцільними є 7 молекулярних дескрипторів: молекулярна маса, кількість акцепторів водневого зв'язку, кількість донорів водневого зв'язку, логарифмічний коефіцієнт розподілу між октанолом і водою, індекс Балабана, молекулярний топологічний індекс, індекс Вінера (множинний коефіцієнт детермінації дорівнює 0,3701). Оптимальна кількість прихованих нейронів для ефективної реалізації нейронної мережі прямого поширення сигналу була знайдена експериментальним шляхом і складає 16. Навчена нейронна мережа з 16 прихованими нейронами була використана для прогнозування значень біодоступності молекул лікарських речовин тестової та валідаційної вибірок.*

Висновки. *Нейронна мережа прямого поширення сигналу є ефективним інструментом для прогнозування біодоступності пероральних лікарських речовин на основі їх хімічної структури. Про це свідчать високі значення коефіцієнта детермінації між прогнозованими значеннями біодоступності та експериментальними значеннями біодоступності для тестової та валідаційної вибірок (0,7084 та 0,8432, відповідно). Отримані результати можуть бути корисними на етапі планування експерименту або розробки лікарських засобів.*

Ключові слова: *хемометрія, дизайн лікарських засобів, молекулярний дескриптор, множинна регресія, нейронна мережа, фармація.*

Introduction. Prediction of bioavailability is a crucial aspect of drug development and formulation design. Bioavailability refers to the proportion of a drug that enters the systemic circulation in an active form after administration and is available to exert its therapeutic effect. Accurate prediction can help optimize drug formulations, reduce development costs, and ensure therapeutic efficacy. Experimental methods for predicting bioavailability are often expensive, time-consuming, and resource-intensive. While they are critical

for validation and regulatory approval, their limitations have driven the adoption of complementary tools like machine learning models [6, 11].

Machine learning is revolutionizing the prediction of bioavailability by enabling faster, more accurate, and cost-effective assessments. Its ability to integrate and analyze complex datasets ensures better decision-making in drug development. As machine learning methods evolve, they promise to further enhance drug discovery, formulation design, and personalized medicine, making them indispensable tools in modern pharmaceutical science. Artificial neural networks are one specific type of machine learning algorithm inspired by the structure of the human brain. Neural networks are one of the most powerful tools for predictive modeling, pattern recognition, and complex data analysis. In the context of bioavailability prediction, neural networks are particularly effective at capturing nonlinear relationships and interactions between variables, which are common in pharmaceutical data [1, 2, 9].

Oral dosing is the most common method of drug administration because of its convenience, safety, and cost-effectiveness, that is why the authors pay attention on oral bioavailability.

Aim. To investigate the possibility of predicting the bioavailability of oral drug compounds based on molecular descriptors by means of feedforward neural network.

Materials and methods. Software ChemOffice 2020 was used for calculation molecular descriptors from the drug structures. Software Matlab R2022b was used for building the multivariate regression between molecular descriptors and bioavailability as well as feedforward neural network.

For each of the 145 drug molecules 16 molecular descriptors were calculated: molecular weight, hydrogen bond acceptors, hydrogen bond donors, ovality, molar refractivity, partition coefficient logP, solubility logS, Balaban index, cluster count, molecular topological index, polar surface area, shape attribute, shape coefficient, sum of degrees, sum of valence degrees and Wiener index. Experimental values of bioavailability for all 145 compounds was taken from the literature [10]. Informativeness and effectiveness of different sets of molecular descriptors for drug bioavailability prediction was estimate by multiple regression.

Of the 145 compounds, 10 were randomly selected for use as a test subset, 5 compounds – as a validation subset. The remaining 130 compounds were used to train feedforward neural network.

Training parameters for creating the feedforward neural network: initialization method of the weights – Nguyen-Widrow algorithm, training method – Levenberg-Marquardt, performance – mean squared error, transfer function for hidden layer of neurons – tansig, transfer function for output layer of neurons – linear [4, 7].

Results and discussion. *Determination of required set of molecular descriptors for drug bioavailability prediction.*

Starting point for determination of required set of molecular descriptors for drug bioavailability prediction were four criteria of Lipinski`s rule of five: molecular weight, partition coefficient, hydrogen bond acceptors and hydrogen bond donors [3]. To the specified starting set of descriptors, other descriptors were added one by one and their impact on the predictive value was evaluated.

Table 1 presents values of multiple coefficients of determination for regression between bioavailability and different sets of descriptors. One can see, that addition of Balaban index, molecular topological index and Wiener index to starting set of descriptors genuinely improves the regression model. It is shown that set of seven descriptors (molecular weight, hydrogen bond acceptors, hydrogen bond donors, logP, Balaban index, molecular topological index, Wiener index) is best described the variability of the bioavailability (R = 0,3701). The equation for a multiple regression model using 7 above mentioned parameters is: bioavailability = 48,8405 + 0,05167·molecular weight + 0,08941·hydrogen bond acceptors – 1,8777·hydrogen bond donors + 4,5299·logP – 0,00001634·Balaban index – 0,0145·molecular topological index + 0,09957·Wiener index.

Added variables as ovality, molar refractivity, logS, cluster count, polar surface area, shape attribute, shape coefficient, sum of degrees, sum of valence degrees to starting set of descriptors did not contribute meaningful predictive value. Addition of logS as eighth descriptor slightly increased multiple coefficient of determination, but not significant.

List of 145 drug molecules, their contribution between training, test and validation subsets, calculated values of molecular weight (1), hydrogen bond acceptors (2), hydrogen bond donors (3), logP (4), Balaban index (5), molecular topological index (6), Wiener index (7) are presented in Table 2.

Realization of feedforward neural network

A **feedforward neural network** is one of the simplest types of artificial neural networks. It consists of an input layer, one or more hidden layers, and an output layer, with connections that feed data forward from one layer to the next without cycles or loops [5, 8].

The **number of hidden neurons** in a neural network is a critical parameter that directly influences the model's ability to learn patterns and generalize from data. Choosing the right number of hidden neurons is essential for achieving a good balance between underfitting and overfitting. In this work the optimal number of hidden neurons was found by experimental way. Different number of hidden neurons (from 7 till 18) were checked to find their optimal number. The optimal

Table 1

Values of multiple coefficient of determination for regression between bioavailability and different sets of descriptors

Set of descriptors	Multiple coefficient of determination (R)
molecular weight, hydrogen bond acceptors, hydrogen bond donors, logP (starting set of descriptors)	0,2306
(molecular weight, hydrogen bond acceptors, hydrogen bond donors, logP) + ovality	0,2312
(molecular weight, hydrogen bond acceptors, hydrogen bond donors, logP) + molar refractivity	0,2502
(molecular weight, hydrogen bond acceptors, hydrogen bond donors, logP) + logS	0,2664
(molecular weight, hydrogen bond acceptors, hydrogen bond donors, logP) + Balaban index	0,3195
(molecular weight, hydrogen bond acceptors, hydrogen bond donors, logP) + cluster count	0,2401
(molecular weight, hydrogen bond acceptors, hydrogen bond donors, logP) + molecular topological index	0,3312
(molecular weight, hydrogen bond acceptors, hydrogen bond donors, logP) + polar surface area	0,2399
(molecular weight, hydrogen bond acceptors, hydrogen bond donors, logP) + shape attribute	0,2403
(molecular weight, hydrogen bond acceptors, hydrogen bond donors, logP) + shape coefficient	0,2315
(molecular weight, hydrogen bond acceptors, hydrogen bond donors, logP) + sum of degrees	0,2508
(molecular weight, hydrogen bond acceptors, hydrogen bond donors, logP) + sum of valence degrees	0,2336
(molecular weight, hydrogen bond acceptors, hydrogen bond donors, logP) + Wiener index	0,2967
(molecular weight, hydrogen bond acceptors, hydrogen bond donors, logP) + Balaban index + molecular topological index	0,3321
(molecular weight, hydrogen bond acceptors, hydrogen bond donors, logP) + Balaban index + molecular topological index + Wiener index	0,3701
(molecular weight, hydrogen bond acceptors, hydrogen bond donors, logP) + Balaban index + molecular topological index + Wiener index + logS	0,3729

Table 2

List of drug molecules and calculated values of seven molecular descriptors

N	Drug molecule	Subset	Descriptor						
			1	2	3	4	5	6	7
1	2	3	4	5	6	7	8	9	10
1	Acetaminophen	Training	151,17	2	2	0,441	19887	1227	166
2	Acyclovir	Training	225,21	7	3	-1,622	81102	3053	453
3	Alendronate	Training	249,10	4	6	-2,802	101097	1756	281
4	Allopurinol	Training	136,11	4	2	-0,485	7659	777	105
5	Amiloride	Training	229,63	7	4	1,085	81960	2405	368
6	Aminoglutethimide	Training	232,28	3	2	1,308	97697	3596	484
7	Amlodipine	Training	408,88	5	2	1,840	981585	12808	1836
8	Amoxicillin	Training	365,40	5	4	0,045	518482	10602	1551
9	Ampicillin	Training	349,41	4	3	0,454	421931	9609	1365
10	Aspirin	Training	180,16	2	1	1,443	41162	1722	246
11	Atenolol	Training	266,34	4	3	0,460	317560	6458	890
12	Atropine	Training	289,38	3	1	1,802	229709	7230	959
13	Azathioprine	Training	277,26	7	1	0,720	133848	4562	676
14	Baclofen	Training	213,66	2	2	1,343	61670	2176	318
15	Bendrofluazide	Training	421,41	8	3	1,787	678665	11903	1752
16	Betamethasone	Training	392,47	6	3	1,143	573894	11755	1670
17	Bromocriptine	Training	654,61	7	3	3,464	3108999	42905	5950
18	Bumetanide	Training	364,42	5	3	2,594	619100	10536	1446
19	Captopril	Training	217,28	2	2	0,581	57629	2036	297
20	Carbamazepine	Training	236,27	1	1	2,321	93062	4074	521
21	Cephalexin	Training	347,39	4	3	0,008	427499	9735	1383

Continuation of table 2

1	2	3	4	5	6	7	8	9	10
22	Chlorothiazide	Training	295,71	5	2	-0,506	93124	3142	460
23	Chlortalidone	Training	338,76	4	3	1,318	249216	6678	952
24	Cimetidine	Training	252,34	6	3	1,140	190251	4813	664
25	Ciprofloxacin	Training	331,35	6	2	2,310	316562	8681	1234
26	Cladribine	Training	285,69	8	3	-0,065	132604	4399	671
27	Clindamycin	Training	424,98	6	4	1,611	919724	12601	1843
28	Clonazepam	Training	315,71	3	1	2,675	250800	6593	959
29	Clonidine	Training	230,09	3	2	2,796	41807	2106	301
30	Cloxacillin	Training	434,89	4	2	2,262	814301	15164	2212
31	Cyclophosphamide	Training	261,08	1	1	1,305	58317	1960	301
32	Cytarabine	Training	243,22	6	4	-2,977	100180	3266	496
33	Dexamethasone	Training	392,47	6	3	1,143	573894	11755	1670
34	Didanosine	Training	236,23	6	2	-0,103	80271	3497	502
35	Disopyramide	Training	339,48	3	1	3,496	588748	10646	1372
36	Doxapram	Training	378,52	3	0	3,598	632487	14095	1843
37	Doxepin	Training	279,38	2	0	3,983	210631	6973	882
38	Doxorubicin	Training	543,53	12	6	-0,381	2430223	29879	4392
39	Doxycycline	Training	444,44	9	6	-1,055	1037402	16076	2336
40	Enalapril	Training	376,45	4	2	2,046	989295	14260	1982
41	Ethinyloestradiol	Training	296,41	2	2	4,192	204957	7264	941
42	Etoposide	Training	588,56	12	3	1,102	2725244	36722	5464
43	Etidronate	Training	206,03	3	5	-2,297	29666	814	136
44	Famotidine	Training	337,44	7	4	-1,791	414469	7046	1046
45	Fluconazole	Training	306,28	9	1	0,290	261313	6772	1000
46	Furosemide	Training	330,74	5	3	1,220	286199	6372	939
47	Gabapentin	Training	171,24	2	2	1,327	27510	1432	193
48	Gemfibrozil	Training	250,34	2	1	3,305	226682	5262	705
49	Hydralazine	Training	160,18	4	2	0,671	18782	1417	182
50	Hydrochlorthiazide	Training	297,73	5	3	-0,429	93124	3142	460
51	Ibuprofen	Training	206,29	1	1	3,646	89861	3076	404
52	Idarubicin	Training	497,50	10	5	0,320	1684951	24694	3545
53	Indapamide	Training	365,83	4	2	2,188	416392	9719	1345
54	Irbesartan	Training	428,54	6	1	4,730	1188964	23576	3125
55	Isosorbide dinitrate	Training	236,14	4	0	0,130	80397	2542	449
56	Isosorbide mononitrate	Training	191,14	4	1	-0,840	28699	1461	239
57	Ketamine	Training	237,73	2	1	2,875	69755	2913	388
58	Ketoprofen	Training	254,29	2	1	3,224	181720	5402	724
59	Labetalol	Training	328,41	4	4	2,524	636957	11893	1607
60	Lansoprazole	Training	369,36	8	1	1,493	545169	10846	1631
61	Levodopa	Training	197,19	4	4	-1,757	62340	2160	321
62	Lisinopril	Training	405,50	5	4	1,141	1355521	16912	2362
63	Lorazepam	Training	321,16	3	2	3,479	196116	5637	819
64	Losartan	Training	422,92	7	2	4,846	1045260	19332	2665
65	Meperidine	Training	247,34	2	0	2,609	128001	4369	568
66	Mercaptopurine	Training	152,18	4	2	0,156	7868	798	108
67	Metformin	Training	129,17	5	4	0,607	13687	716	96
68	Methotrexate	Training	454,45	10	5	1,405	2190888	26126	3832
69	Methyldopa	Training	211,22	4	4	-1,632	83432	2552	374
70	Methylprednisolone	Training	374,48	5	3	1,507	498115	11202	1553
71	Metoclopramide	Training	299,80	4	2	1,852	357018	6445	902

Continuation of table 2

1	2	3	4	5	6	7	8	9	10
72	Metoprolol	Training	267,37	4	2	1,643	322987	6707	906
73	Metronidazole	Training	171,16	4	1	-0,123	27473	1288	193
74	Mexilitine	Training	179,26	2	1	2,081	42838	2034	256
75	Misoprostol	Training	382,54	4	2	3,449	1677329	16886	2331
76	Naloxone	Training	327,38	5	2	0,612	232022	7720	1047
77	Naltrexone	Training	341,41	5	2	0,705	252019	8803	1189
78	Nifedipine	Training	346,34	4	1	1,953	549472	8792	1280
79	Nimodipine	Training	418,45	5	1	2,667	1401983	15835	2290
80	Nitrazepam	Training	281,27	3	1	2,053	204937	6142	857
81	Nizatidine	Training	331,45	5	2	1,156	535958	8353	1227
82	Norfloxacin	Training	319,34	6	2	2,211	317490	7795	1116
83	Nortriptyline	Training	263,38	1	1	4,683	165063	6143	759
84	Omeprazole	Training	345,42	6	1	0,194	438619	10452	1419
85	Ondansetron	Training	293,37	4	0	2,508	217302	7706	997
86	Orciprenaline	Training	211,26	4	4	0,999	87750	2864	394
87	Oxprenolol	Training	265,35	4	2	2,036	297610	6263	834
88	Oxycodone	Training	315,37	5	1	0,306	186078	6673	907
89	Pamidronate	Training	235,07	4	6	-3,141	67926	1345	220
90	Pantoprazole	Training	383,37	9	1	0,117	650686	11929	1806
91	Pethidine	Training	247,34	2	0	2,609	128001	4369	568
92	Phenobarbitone	Training	232,24	3	2	1,327	92531	3322	458
93	Phenytoin	Training	252,27	2	2	2,252	121875	4679	617
94	Pindolol	Training	248,33	4	3	1,992	155129	5206	687
95	Pravastatin	Training	424,53	5	4	2,133	1592942	18188	2600
96	Prazosin	Training	383,41	8	1	1,312	754065	15647	2195
97	Primidone	Training	218,26	2	2	1,526	69194	2898	385
98	Procainamide	Training	235,33	3	2	1,328	176834	4715	618
99	Prochlorperazine	Training	373,94	3	0	4,544	427920	11513	1544
100	Promethazine	Training	284,42	2	0	4,371	162136	5850	744
101	Propylthiouracil	Training	170,23	1	1	0,477	18909	1161	158
102	Quinapril	Training	438,52	4	2	3,139	1650793	22541	3064
103	Quinine	Training	324,42	4	1	2,584	330149	9833	1286
104	Ranitidine	Training	314,40	5	2	1,301	535958	8479	1227
105	Ribavirin	Training	244,21	8	4	-2,875	103957	3311	515
106	Salicylic acid	Training	138,12	2	2	2,208	11308	822	114
107	Selegiline	Training	187,29	1	0	2,687	65396	2766	337
108	Simvastatin	Training	418,57	3	1	4,649	1159529	17942	2440
109	Sotalol	Training	272,36	4	3	1,079	226065	5254	706
110	Spironolactone	Training	416,58	3	0	3,390	601303	13884	1905
111	Sulfamethoxazole	Training	253,28	5	2	0,690	108079	3926	535
112	Sulfisoxazole	Training	267,30	5	2	0,913	138704	4543	614
113	Sumatriptan	Training	295,40	4	2	0,908	237792	6516	859
114	Tacrine	Training	198,27	2	1	2,205	41269	2657	326
115	Temazepam	Training	300,74	3	1	2,997	194836	5895	814
116	Terbutaline	Training	225,29	4	4	1,414	118746	3420	468
117	Testosterone	Training	288,43	2	1	3,535	160228	6172	802
118	Tetracycline	Training	444,44	9	6	-0,762	1043893	16144	2350
119	Theophyline	Training	180,17	4	1	-0,746	25429	1509	211
120	Timolol	Training	316,42	7	2	1,309	324259	7477	1063
121	Tolbutamide	Training	270,35	3	2	2,141	219163	5180	685

Continuation of table 2

1	2	3	4	5	6	7	8	9	10
122	Triamterene	Training	253,27	7	3	1,693	132417	5012	670
123	Trimipramine	Training	294,44	2	0	4,265	255824	7852	979
124	Trandolaprilat	Training	402,49	4	3	2,610	1041851	17000	2340
125	Tropisetron	Training	284,36	3	1	3,047	188872	7235	944
126	Valproic acid	Training	144,21	1	1	2,661	23285	1000	131
127	Verapamil	Training	454,61	6	0	4,665	2740315	27284	3698
128	Warfarin	Training	308,33	3	1	2,985	310959	8146	1092
129	Zalcitabine	Training	211,22	4	2	-1,394	58431	2583	369
130	Zidovudine	Training	267,25	6	2	-0,884	175676	4676	701
131	Amantadine	Test	151,25	1	1	2,236	8844	1067	124
132	Amiodarone	Test	645,32	4	0	7,861	1463147	21156	2897
133	Amitriptiline	Test	277,41	1	0	4,932	210631	7133	882
134	Chloramphenicol	Test	323,13	4	3	0,669	348207	5269	880
135	Chlordiazepoxide	Test	299,76	3	1	2,729	198023	6091	828
136	Diazepam	Test	284,74	2	0	2,921	158292	5393	726
137	Dolasetron	Test	324,38	4	1	2,217	303937	10213	1368
138	Ethambutol	Test	204,31	4	4	-0,360	138073	2798	383
139	Lamivudine	Test	229,25	4	2	-1,451	58431	2479	369
140	Lamotrigine	Test	256,09	5	2	3,876	74968	2813	417
141	Flunitrazepam	Validation	313,29	4	0	2,354	302738	7365	1063
142	Lidocaine	Validation	234,34	2	1	2,110	159216	4344	556
143	Ramipril	Validation	416,52	4	2	2,797	1210619	18583	2546
144	Morphine	Validation	285,34	4	2	0,775	123644	5332	712
145	Minocycline	Validation	457,48	9	5	0,110	1258222	18638	2672

number of hidden neurons provide the properly training of the artificial neural networks.

We have shown the graph of the network training performance versus the number of hidden neurons (Fig. 1). By concept of network training performance we mean determination coefficient between experimental and predicted bioavailability values for training subset. So, the optimal number of hidden neurons for feedforward neural network is 16, because the maximum value of determination coefficient (R = 0,3160) is observed at this number of hidden neurons.

Trained feedforward neural network with 16 hidden neurons was used for predicting bioavailability values for test and validation subsets. Dependences of experimental bioavailability values on predicted bioavailability values for test and validation subsets

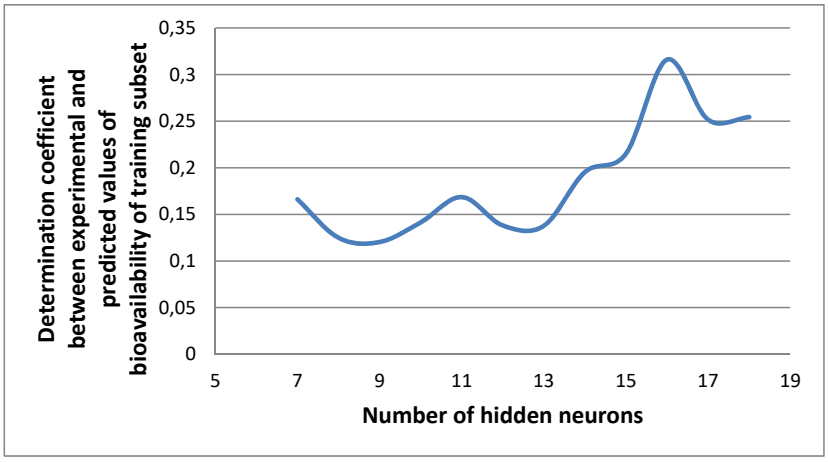

Fig. 1. Dependence of determination coefficient between experimental and predicted bioavailability values for training subset on the number of hidden neurons for feedforward neural network

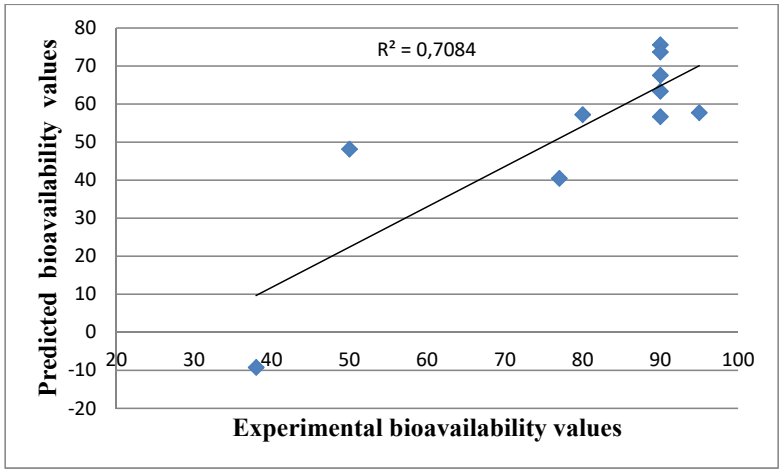

Fig. 2. Dependence of predicted bioavailability values on experimental bioavailability values for test subset

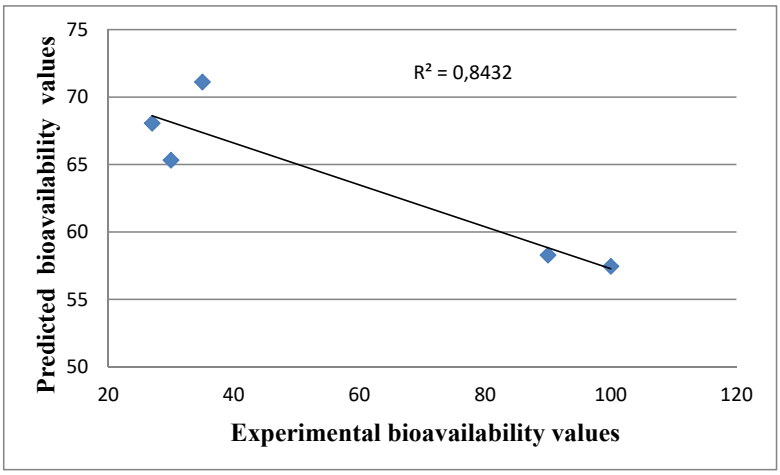

Fig. 3. Dependence of predicted bioavailability values on experimental bioavailability values for validation subset

are characterized by high values of determination coefficient – 0,7084 and 0,8432, correspondingly (Fig. 2, 3).

Conclusion. It is established that from 16 studied molecular descriptors just 7 are informativeness for oral drug bioavailability prediction. They are molecular weight, hydrogen bond acceptors, hydrogen bond donors, logP, Balaban index, molecular topological index, Wiener index. So, four criteria of Lipinski's rule were added with topological descriptors. We can conclude that topological descriptors are critical in understanding and predicting adsorption and permeation behaviors. They bridge molecular structure with functional properties, enabling more efficient design and optimization in drug development.

Feedforward neural network is the effective tool for prediction of oral drug bioavailability based on chemical structure. This is evidenced by high values of determination coefficient between predicted bioavailability values and experimental bioavailability values for test and validation subsets (0,7084 and 0,8432, correspondingly). The created and trained neural network has been saved as software Matlab R2022b file (mat-file) and can be used for bioavailability prediction of new drug molecules. Obtained results can be useful at the stage of experiment planning or drug design.

Bibliography:

1. Пушкарьова Я. М., Калюженко А. В. Прогнозування ступеня небезпечності / ризику залишкових розчинників у лікарських засобах методами хемометрії. *Фармацевтичний часопис*. 2023. № 3. С. 16–25. DOI: 10.11603/2312-0967.2023.3.13985.

2. Феденько С., Волошенюк Т., Різак Г. Аналіз впливу інноваційних технологій на розвиток фармацевтичного ринку в Україні. *Сучасна медицина, фармація та психологічне здоров'я*. 2024. Випуск 2 (16). С. 106–110. DOI: 10.32689/2663-0672-2024-2-17.

3. Chagas C. M., Moss S., Alisaraie L. Drug metabolites and their effects on the development of adverse reactions: Revisiting Lipinski's Rule of Five. *International Journal of Pharmaceutics*. 2018. Vol. 549, no. 1–2. P. 133–149. DOI: 10.1016/j.ijpharm.2018.07.046.

4. Guha S., Jana R. K., Sanyal M. K. Artificial Neural Network approaches for disaster management: A literature review (2010–2021). *International Journal of Disaster Risk Reduction*. 2022. Vol. 81. P. 103276. DOI: 10.1016/j.ijdrr.2022.103276.

5. Houssein E. H., Hosney M. E., Emam M. M., Oliva D., Younis E. M., Ali A. A., Mohamed W. M. Optimizing feedforward neural networks using a modified weighted mean of vectors: Case study chemical datasets. *Swarm and Evolutionary Computation*. 2024. Vol. 89. P. 101656. DOI: 10.1016/j.swevo.2024.101656.

6. Ma L., Yan Y., Dai S., Shao D., Yi S., Wang J., ... Yan, J. Research on prediction of human oral bioavailability of drugs based on improved deep forest. *Journal of Molecular Graphics and Modelling*. 2024. Vol. 133. P. 108851. DOI: 10.1016/j.jmgm.2024.108851.

7. Mumali F. Artificial neural network-based decision support systems in manufacturing processes: A systematic literature review. *Computers & Industrial Engineering.* 2022. Vol. 165. P. 107964. DOI: 10.1016/j.cie.2022.107964.

8. Mutihac L., Mutihac R. Mining in chemometrics. *Analytica Chimica Acta*. 2008. Vol. 612, no. 1. P. 1–18. DOI: 10.1016/j.aca.2008.02.025.

9. Pushkarova Y., Zaitseva G., Saker M. A. Prediction of Toxicity of Phenols Using Artificial Neural Networks. *12th International Conference on Advanced Computer Information Technologies (ACIT)*, Slovakia, 26–28 September 2022. 2022. DOI: 10.1109/acit54803.2022.9913174.

10. Turner J. V., Glass B. D., Agatonovic-Kustrin S. Prediction of drug bioavailability based on molecular structure. *Analytica Chimica Acta*. 2003. Vol. 485, no. 1. P. 89–102. DOI: 10.1016/s0003-2670(03)00406-9.

11. Yang Q., Fan L., Hao E., Hou X., Deng J., Xia Z., Du Z. Construction of An Oral Bioavailability Prediction Model Based on Machine Learning for Evaluating Molecular Modifications. *Journal of Pharmaceutical Sciences*. 2024. Vol. 113. P. 1155–1167. DOI: 10.1016/j.xphs.2024.02.026.