GEORGIAN MEDICAL NEWS

ISSN 1512-0112

NO 11 (332) ноябрь 2022

ТБИЛИСИ - NEW YORK

ЕЖЕМЕСЯЧНЫЙ НАУЧНЫЙ ЖУРНАЛ

Медицинские новости Грузии

საქართველოს სამედიცინო სიახლენი

GEORGIAN MEDICAL NEWS

Monthly Georgia-US joint scientific journal published both in electronic and paper formats of the Agency of Medical Information of the Georgian Association of Business Press. Published since 1994. Distributed in NIS, EU and USA.

GMN: Georgian Medical News is peer-reviewed, published monthly journal committed to promoting the science and art of medicine and the betterment of public health, published by the GMN Editorial Board since 1994. GMN carries original scientific articles on medicine, biology and pharmacy, which are of experimental, theoretical and practical character; publishes original research, reviews, commentaries, editorials, essays, medical news, and correspondence in English and Russian.

GMN is indexed in MEDLINE, SCOPUS, PubMed and VINITI Russian Academy of Sciences. The full text content is available through EBSCO databases.

GMN: Медицинские новости Грузии - ежемесячный рецензируемый научный журнал, издаётся Редакционной коллегией с 1994 года на русском и английском языках в целях поддержки медицинской науки и улучшения здравоохранения. В журнале публикуются оригинальные научные статьи в области медицины, биологии и фармации, статьи обзорного характера, научные сообщения, новости медицины и здравоохранения. Журнал индексируется в MEDLINE, отражён в базе данных SCOPUS, PubMed и ВИНИТИ РАН. Полнотекстовые статьи журнала доступны через БД EBSCO.

GMN: Georgian Medical News – საქართველოს სამედიცინო სიახლენი – არის ყოველთვიური სამეცნიერო სამედიცინო რეცენზირებადი ჟურნალი, გამოიცემა 1994 წლიდან, წარმოადგენს სარედაქციო კოლეგიისა და აშშ-ის მეცნიერების, განათლების, ინდუსტრიის, ხელოვნებისა და ბუნებისმეტყველების საერთაშორისო აკადემიის ერთობლივ გამოცემას. GMN-ში რუსულ და ინგლისურ ენებზე ქვეყნდება ექსპერიმენტული, თეორიული და პრაქტიკული ხასიათის ორიგინალური სამეცნიერო სტატიები მედიცინის, ბიოლოგიისა და ფარმაციის სფეროში, მიმოხილვითი ხასიათის სტატიები.

ჟურნალი ინდექსირებულია MEDLINE-ის საერთაშორისო სისტემაში, ასახულია SCOPUS-ის, PubMed-ის და ВИНИТИ РАН-ის მონაცემთა ბაზებში. სტატიების სრული ტექსტი ხელმისაწვდომია EBSCO-ს მონაცემთა ბაზებიდან.

WEBSITE

www.geomednews.com

К СВЕДЕНИЮ АВТОРОВ!

При направлении статьи в редакцию необходимо соблюдать следующие правила:

1. Статья должна быть представлена в двух экземплярах, на русском или английском языках, напечатанная через **полтора интервала на одной стороне стандартного листа с шириной левого поля в три сантиметра**. Используемый компьютерный шрифт для текста на русском и английском языках - **Times New Roman** (**Кириллица**), для текста на грузинском языке следует использовать **AcadNusx.** Размер шрифта - **12**. К рукописи, напечатанной на компьютере, должен быть приложен CD со статьей.

2. Размер статьи должен быть не менее десяти и не более двадцати страниц машинописи, включая указатель литературы и резюме на английском, русском и грузинском языках.

3. В статье должны быть освещены актуальность данного материала, методы и результаты исследования и их обсуждение.

При представлении в печать научных экспериментальных работ авторы должны указывать вид и количество экспериментальных животных, применявшиеся методы обезболивания и усыпления (в ходе острых опытов).

4. К статье должны быть приложены краткое (на полстраницы) резюме на английском, русском и грузинском языках (включающее следующие разделы: цель исследования, материал и методы, результаты и заключение) и список ключевых слов (key words).

5. Таблицы необходимо представлять в печатной форме. Фотокопии не принимаются. **Все цифровые, итоговые и процентные данные в таблицах должны соответствовать таковым в тексте статьи**. Таблицы и графики должны быть озаглавлены.

6. Фотографии должны быть контрастными, фотокопии с рентгенограмм - в позитивном изображении. Рисунки, чертежи и диаграммы следует озаглавить, пронумеровать и вставить в соответствующее место текста **в tiff формате**.

В подписях к микрофотографиям следует указывать степень увеличения через окуляр или объектив и метод окраски или импрегнации срезов.

7. Фамилии отечественных авторов приводятся в оригинальной транскрипции.

8. При оформлении и направлении статей в журнал МНГ просим авторов соблюдать правила, изложенные в «Единых требованиях к рукописям, представляемым в биомедицинские журналы», принятых Международным комитетом редакторов медицинских журналов - http://www.spinesurgery.ru/files/publish.pdf и http://www.nlm.nih.gov/bsd/uniform_requirements.html
В конце каждой оригинальной статьи приводится библиографический список. В список литературы включаются все материалы, на которые имеются ссылки в тексте. Список составляется в алфавитном порядке и нумеруется. Литературный источник приводится на языке оригинала. В списке литературы сначала приводятся работы, написанные знаками грузинского алфавита, затем кириллицей и латиницей. Ссылки на цитируемые работы в тексте статьи даются в квадратных скобках в виде номера, соответствующего номеру данной работы в списке литературы. Большинство цитированных источников должны быть за последние 5-7 лет.

9. Для получения права на публикацию статья должна иметь от руководителя работы или учреждения визу и сопроводительное отношение, написанные или напечатанные на бланке и заверенные подписью и печатью.

10. В конце статьи должны быть подписи всех авторов, полностью приведены их фамилии, имена и отчества, указаны служебный и домашний номера телефонов и адреса или иные координаты. Количество авторов (соавторов) не должно превышать пяти человек.

11. Редакция оставляет за собой право сокращать и исправлять статьи. Корректура авторам не высылается, вся работа и сверка проводится по авторскому оригиналу.

12. Недопустимо направление в редакцию работ, представленных к печати в иных издательствах или опубликованных в других изданиях.

При нарушении указанных правил статьи не рассматриваются.

REQUIREMENTS

Please note, materials submitted to the Editorial Office Staff are supposed to meet the following requirements:

1. Articles must be provided with a double copy, in English or Russian languages and typed or compu-ter-printed on a single side of standard typing paper, with the left margin of **3** centimeters width, and **1.5** spacing between the lines, typeface - **Times New Roman** (**Cyrillic**), print size - **12** (referring to Georgian and Russian materials). With computer-printed texts please enclose a CD carrying the same file titled with Latin symbols.

2. Size of the article, including index and resume in English, Russian and Georgian languages must be at least 10 pages and not exceed the limit of 20 pages of typed or computer-printed text.

3. Submitted material must include a coverage of a topical subject, research methods, results, and review.

Authors of the scientific-research works must indicate the number of experimental biological species drawn in, list the employed methods of anesthetization and soporific means used during acute tests.

4. Articles must have a short (half page) abstract in English, Russian and Georgian (including the following sections: aim of study, material and methods, results and conclusions) and a list of key words.

5. Tables must be presented in an original typed or computer-printed form, instead of a photocopied version. **Numbers, totals, percentile data on the tables must coincide with those in the texts of the articles.** Tables and graphs must be headed.

6. Photographs are required to be contrasted and must be submitted with doubles. Please number each photograph with a pencil on its back, indicate author's name, title of the article (short version), and mark out its top and bottom parts. Drawings must be accurate, drafts and diagrams drawn in Indian ink (or black ink). Photocopies of the X-ray photographs must be presented in a positive image in **tiff format**.

Accurately numbered subtitles for each illustration must be listed on a separate sheet of paper. In the subtitles for the microphotographs please indicate the ocular and objective lens magnification power, method of coloring or impregnation of the microscopic sections (preparations).

7. Please indicate last names, first and middle initials of the native authors, present names and initials of the foreign authors in the transcription of the original language, enclose in parenthesis corresponding number under which the author is listed in the reference materials.

8. Please follow guidance offered to authors by The International Committee of Medical Journal Editors guidance in its Uniform Requirements for Manuscripts Submitted to Biomedical Journals publication available online at: http://www.nlm.nih.gov/bsd/uniform_requirements.html
http://www.icmje.org/urm_full.pdf
In GMN style for each work cited in the text, a bibliographic reference is given, and this is located at the end of the article under the title "References". All references cited in the text must be listed. The list of references should be arranged alphabetically and then numbered. References are numbered in the text [numbers in square brackets] and in the reference list and numbers are repeated throughout the text as needed. The bibliographic description is given in the language of publication (citations in Georgian script are followed by Cyrillic and Latin).

9. To obtain the rights of publication articles must be accompanied by a visa from the project instructor or the establishment, where the work has been performed, and a reference letter, both written or typed on a special signed form, certified by a stamp or a seal.

10. Articles must be signed by all of the authors at the end, and they must be provided with a list of full names, office and home phone numbers and addresses or other non-office locations where the authors could be reached. The number of the authors (co-authors) must not exceed the limit of 5 people.

11. Editorial Staff reserves the rights to cut down in size and correct the articles. Proof-sheets are not sent out to the authors. The entire editorial and collation work is performed according to the author's original text.

12. Sending in the works that have already been assigned to the press by other Editorial Staffs or have been printed by other publishers is not permissible.

Articles that Fail to Meet the Aforementioned Requirements are not Assigned to be Reviewed.

ავტორთა საყურადღებოდ!

რედაქციაში სტატიის წარმოდგენისას საჭიროა დავიცვათ შემდეგი წესები:

1. სტატია უნდა წარმოადგინოთ 2 ცალად, რუსულ ან ინგლისურ ენებზე,დაბეჭდილი სტანდარტული ფურცლის 1 გვერდზე, 3 სმ სიგანის მარცხენა ველისა და სტრიქონებს შორის 1,5 ინტერვალის დაცვით. გამოყენებული კომპიუტერული შრიფტი რუსულ და ინგლისურენოვან ტექსტებში - **Times New Roman (Кириллица)**, ხოლო ქართულენოვან ტექსტში საჭიროა გამოვიყენოთ **AcadNusx.** შრიფტის ზომა – 12. სტატიას თან უნდა ახლდეს CD სტატიით.

2. სტატიის მოცულობა არ უნდა შეადგენდეს 10 გვერდზე ნაკლებს და 20 გვერდზე მეტს ლიტერატურის სიის და რეზიუმეების (ინგლისურ, რუსულ და ქართულ ენებზე) ჩათვლით.

3. სტატიაში საჭიროა გაშუქდეს: საკითხის აქტუალობა; კვლევის მიზანი; საკვლევი მასალა და გამოყენებული მეთოდები; მიღებული შედეგები და მათი განსჯა. ექსპერიმენტული ხასიათის სტატიების წარმოდგენისას ავტორებმა უნდა მიუთითონ საექსპერიმენტო ცხოველების სახეობა და რაოდენობა; გაუტკივარებისა და დაძინების მეთოდები (მწვავე ცდების პირობებში).

4. სტატიას თან უნდა ახლდეს რეზიუმე ინგლისურ, რუსულ და ქართულ ენებზე არანაკლებ ნახევარი გვერდის მოცულობისა (სათაურის, ავტორების, დაწესებულების მითითებით და უნდა შეიცავდეს შემდეგ განყოფილებებს: მიზანი, მასალა და მეთოდები, შედეგები და დასკვნები; ტექსტუალური ნაწილი არ უნდა იყოს 15 სტრიქონზე ნაკლები) და საკვანძო სიტყვების ჩამონათვალი (key words).

5. ცხრილები საჭიროა წარმოადგინოთ ნაბეჭდი სახით. ყველა ციფრული, შემაჯამებელი და პროცენტული მონაცემები უნდა შეესაბამებოდეს ტექსტში მოყვანილს.

6. ფოტოსურათები უნდა იყოს კონტრასტული; სურათები, ნახაზები, დიაგრამები - დასათაურებული, დანომრილი და სათანადო ადგილას ჩასმული. რენტგენოგრამების ფოტოასლები წარმოადგინეთ პოზიტიური გამოსახულებით **tiff** ფორმატში. მიკროფოტოსურათების წარწერებში საჭიროა მიუთითოთ ოკულარის ან ობიექტივის საშუალებით გადიდების ხარისხი, ანათალების შეღებვის ან იმპრეგნაციის მეთოდი და აღნიშნოთ სურათის ზედა და ქვედა ნაწილები.

7. სამამულო ავტორების გვარები სტატიაში აღინიშნება ინიციალების თანდართვით, უცხოურისა – უცხოური ტრანსკრიპციით.

8. სტატიას თან უნდა ახლდეს ავტორის მიერ გამოყენებული სამამულო და უცხოური შრომების ბიბლიოგრაფიული სია (ბოლო 5-8 წლის სიღრმით). ანბანური წყობით წარმოდგენილ ბიბლიოგრაფიულ სიაში მიუთითეთ ჯერ სამამულო, შემდეგ უცხოელი ავტორები (გვარი, ინიციალები, სტატიის სათაური, ჟურნალის დასახელება, გამოცემის ადგილი, წელი, ჟურნალის №, პირველი და ბოლო გვერდები). მონოგრაფიის შემთხვევაში მიუთითეთ გამოცემის წელი, ადგილი და გვერდების საერთო რაოდენობა. ტექსტში კვადრატულ ფჩხილებში უნდა მიუთითოთ ავტორის შესაბამისი N ლიტერატურის სიის მიხედვით. მიზანშეწონილია, რომ ციტირებული წყაროების უმეტესი ნაწილი იყოს 5-6 წლის სიღრმის.

9. სტატიას თან უნდა ახლდეს: ა) დაწესებულების ან სამეცნიერო ხელმძღვანელის წარდგინება, დამოწმებული ხელმოწერითა და ბეჭდით; ბ) დარგის სპეციალისტის დამოწმებული რეცენზია, რომელშიც მითითებული იქნება საკითხის აქტუალობა, მასალის საკმაობა, მეთოდის სანდოობა, შედეგების სამეცნიერო-პრაქტიკული მნიშვნელობა.

10. სტატიის ბოლოს საჭიროა ყველა ავტორის ხელმოწერა, რომელთა რაოდენობა არ უნდა აღემატებოდეს 5-ს.

11. რედაქცია იტოვებს უფლებას შეასწოროს სტატია. ტექსტზე მუშაობა და შეჯერება ხდება საავტორო ორიგინალის მიხედვით.

12. დაუშვებელია რედაქციაში ისეთი სტატიის წარდგენა, რომელიც დასაბეჭდად წარდგენილი იყო სხვა რედაქციაში ან გამოქვეყნებული იყო სხვა გამოცემებში.

აღნიშნული წესების დარღვევის შემთხვევაში სტატიები არ განიხილება.

GEORGIAN MEDICAL NEWS
No 11 (332) 2022

Содержание:

GEORGIAN MEDICAL NEWS
No 11 (332) 2022

DRUG DEVELOPMENT BY *IN SILICO* METHODS

Lenskaya K[1,2], Bagaturiya G[3], Buinov L[4], Lebedev A[5], Grishin V[6,7], Proshin S[4].

[1]*Saint Petersburg State University, Russia.*

[2]*Saint Petersburg State Chemical and Pharmaceutical University, Russia.*

[3]*Saint Petersburg State Pediatric Medical University, Russia.*

[4]*Herzen University, St. Petersburg, Russia.*

[5]*Institute of Experimental Medicine, St. Petersburg, Russia.*

[6]*Lesgaft National State University of Physical Education, Sport and Health, St. Petersburg, Russia.*

[7]*Research Institute of Dentistry and ChLH, I.P. Pavlov St. Petersburg State Medical University, Saint Petersburg, Russia.*

Abstract.

Multicenter studies of the last decade have confirmed the direct relationship between the spatial structure of drug compounds and their biological activity. **Aim:** The relevance of this topic, on the one hand, is explained by the rapid development of new technologies, on the other hand, by a large number of researches works devoted to the study of the influence of stereoisomeric factors on the therapeutic effect and side effects of drugs.

Methods: In recent decades, computational approaches have been used to reveal the relationship between the biological properties of substances and their structure and to quantitatively describe this relationship. Quantitative Structure - Activity Relationships (QSAR) is widely used in modern drug chemistry. In combination with molecular modeling techniques, this trend has been commonly referred to as in silico methods, which implies a preliminary search for drugs by computer only, prior to experimental screening.

Results: Hundreds of QSAR and molecular modeling programs have been developed around the world: dozens of specialized scientific journals are published dedicated to computer-based methods of searching for biologically active substances. Thousands of articles have been published demonstrating the results of successful application of QSAR methods in the search for new biologically active substances. There are currently more than 101 million known organic compounds in the world, and the known number of biological activities exceeds 25,000, so global experimental testing of all substances is impossible.

Conclusion: Thus, the use of in silico methods to search for drugs, in particular, for the preliminary selection of unpromising compounds is super relevant.

Key words. *In silico* methods, Quantitative Structure - Activity Relationships, stereoisomeric factors.

Introduction.

The traditional way of searching for drugs in the 20th century was to synthesize a wide variety of organic compounds and test them for different types of activity. Hundreds of thousands of substances were synthesized and tested. The laboriousness and duration of this process forced scientists to think about the development of the theoretical foundations of the search for drugs and, mainly, the problem of limiting the number of synthesized compounds [1].

Gradually, computational approaches emerged to identify the relationship between the biological properties of substances and their structure and to quantitatively describe this relationship. Such studies led to the birth of a scientific trend called QSAR ("Quantitative Structure - Activity Relationships" or "quantitative structure-activity relationships") in modern drug chemistry. In combination with molecular modeling methods, this direction received the common name of *in silico* methods, which implies preliminarily searching for drugs only in a computerized way, before the experimental screening.

The history of the development of *in silico* methods coincides with the development of methods for statistical analysis of empirical patterns. Corwin Hansch found that a regression equation describing the relationship between biological activity and the lipophilic, electronic, and steric parameters of compounds is a very convenient tool for the researcher [2]. The regression equation is a statistical analogue of the exact functional dependence including several independent variables X1, X2... Xn.

$$Y = a_0 + a_1X_1 + a_2X_2 + a_3X_3 + \ldots + a_iX_i$$

As the mathematical apparatus in QSAR is mainly used multiple regression analysis, and the regression equation in the form of a polynomial is most often used in practice [3]. The early 1970s are characterized by the development of pattern (object) recognition theory (TRO). In this case, the prediction consists in the assignment of an unstudied compound to one of two classes of substances - active or inactive (highly active/lowly active). The basis of the prediction technique is the regularities obtained by TRO methods linking between the presence or absence of activity and the structure of compounds represented in the form of vector (matrix) of descriptors. The descriptor description of a compound is substituted into the previously calculated dividing function and the metric of new chemical structure belonging to each of two classes is calculated.

Since the early 80s of the last century QSAR began to actively use neural network modeling methodology actively. Currently, the most popular are multilayer neural networks [4]. Simple neural networks have certain limitations, which do not allow to solve a number of important scientific and practical problems. To overcome existing limitations the concept of self-organizing networks was proposed. It allows to construct learning algorithms for neural networks with search behavior, capable of training complex hierarchical systems, operating on large time intervals, maintaining homeostasis, etc. [5].

The history of computer modeling development reveals a number of problems arising from the need to perceive and

analyze large volumes of data generated by modern high-performance computer complexes. Visualization is an essential part of the computer modeling process, providing analysis and proper interpretation of computational results, as well as further work with the computational model. It was the emergence of graphics processors and programs for visualizing chemical structures in the early 90's that led to a leap in the field of 3D molecular modeling [6].

The mid 90's are characterized by the development of methods for 3D modeling of proteins. Currently, the main method for modeling the spatial structure of pharmacologically important target proteins is homology modeling [7]. It consists in the prediction of the three-dimensional structure of a protein molecule from the amino acid sequence using data on the spatial structure of one or more experimentally studied homologues. Molecular modeling and construction are used in the final stages of the process of computer search for new drugs. This process can be viewed as the creation of a new supramolecular system: a "macromolecule-ligand" complex [8].

A rapidly developing and very fruitful approach to the development of new substances with a certain biological activity is 3D modeling of the target protein molecule with the simultaneous selection of organic molecules that bind optimally to it optimally, using molecular docking methods. Currently, docking is the main tool for predicting the binding mode of a biological macromolecule to a ligand, which allows finding the most energetically favorable location of the ligand in the binding site of the biological target [9]. Docking is very important in the search and synthesis of new drugs.

It should be noted that out of more than 40 thousand proteins of the human proteome only for about 2.5 thousand proteins their functions have been established, among which only 518 pharmacologically relevant biotargets have been identified so far [10], for 324 of which over 1.2 thousand drug substances are known to act on them and are produced worldwide. These targets comprise 75 subfamilies grouped into seven major classes: G-protein coupled receptors, ion channels, nuclear receptors, receptor kinases and proteases, enzymes, transporters/ translocators, and nucleic acids [11]. Various methods can be used to qualitatively and quantitatively evaluate the action of compounds at the stage of receptor-ligand complex formation, but 3D-QSAR methods have gained popularity in recent years. Among computational procedures in QSAR, methods of linear, nonlinear, and nonparametric regression, TRO, methods of similarity analysis and clustering, artificial neural networks, statistical analysis methods, decision-making methods, logical inference based on classical, inductive, probabilistic, fuzzy, and other logics are used [12]. In fact, all 2D- and 3D-QSAR methods are based on describing the structure of a chemical compound using a set of numerical characteristics - descriptors and constructing "correlations" between the property (activity) value and descriptor values [13]. Empirical QSAR models allow predicting the properties of new substances by the structural formula, as well as the directed design of new compounds with a given set of properties.

Currently, the use of *in silico* methods to search for new drugs is a fairly common practice (DRUG Design 2019; Recent Advances in QSAR 2019). Hundreds of QSAR and molecular modeling programs have been developed around the world: dozens of specialized scientific journals are published dedicated to computer-based methods of searching for biologically active substances: only two leading publishers issue publish 34 journals related to this topic. Thousands of articles have been published to show the results of successful application of QSAR methods in the search for new biologically active substances.

In particular, computer methods are widely used in the search for compounds with antidiabetic activity [14].

For example, jointly used discriminant and regression analysis to identify the relationship between hypoglycemic activity and the molecular topology of antidiabetic drugs. This model has been used to design new hypoglycemic agents.

The use of *in silico* methods to search for new drugs is a common practice [15]. The power of the chemical-biological universe (the space in which all biologically active substances by their known types of activity are located) is very large, so it is impossible to experimentally test all compounds for at least one type of pharmacological activity.

The drug discovery process consists of several steps [16].

Finding a promising basic structure (base structure).

virtual screening and planning of experimental screening.

experimental screening and selection of several leader compounds.

specification of pharmacodynamic and pharmacokinetic characteristics of compounds-leaders by *in silico* methods in order to select a compound for further in-depth study.

in-depth study of the leader compound.

in silico analysis of the possible mechanism of action of the leader compound [17].

There are currently more than 101 million known organic compounds in the world [18]. The known number of biological activities exceeds 25 thousand, so global experimental testing of all substances is impossible. Thus, the use of *in silico* methods to search for drugs, in particular to pre-screen unpromising compounds, is now generally accepted.

The main computational methods for searching quantitative relations between the structure and biological activity of chemical compounds (QSAR).

The parabolic Hansch model.

It is known that the logarithm of the distribution coefficient in the octanol-water system is the main characteristic of the lipophilicity of a molecule used in correlation equations. One of Hench's ideas was that this value can be represented as the sum of the values he introduced, which characterize the contribution to the lipophilicity of individual atoms or fragments of the structure. The parameter has negative values for hydrophilic groups and positive values for hydrophobic groups. The importance of this idea lies in the fact that, by summing up certain experimental values for the structural fragments of a given compound, we can obtain the calculated value of its lipophilicity logP and, in some cases, estimate the biological activity of substances prior to their synthesis.

Hansch's equation was first calculated in and in general form is the following equation

$$\log\frac{1}{C} = a_0 + a_1\pi^2 + a_2\pi + a_3\sigma + a_4 E_S,$$

where is the concentration of the substance that causes a certain level of biological effect (e.g., EC50).

π – Hanche's constant (lipophilicity).

σ – Hammett's constant (electronic parameter).

E_S – Taft's constant (steric parameter)

a_0, a_1, a_2, a_3, a_4 – regression coefficients.

To date, Hanche's approach has been developed and is widely used to search for correlations between biological activity, lipophilicity, electronic and steric characteristics of compounds.

Multiple linear regression.

This approach allows us to determine the degree of variation of the dependent variable by independent variables; to predict the values of the dependent variable on the basis of the values of independent variables; and to determine the contributions of individual independent variables to the variation of the dependent variable [19].

The general form of the regression equation is given by the following formula:

$$Y = a_0 + \sum_{i=1}^{M} a_i X_i,$$

where Y – dependent variable.

X_i – independent variables, $i = 1, \ldots \dot{I}$.

M - number of independent variables.

a_0 – free term of regression equation.

a_i – regression coefficients in front of independent variables, $i = 1, \ldots \dot{I}$.

In particular, the Hanche equation is a linear regression equation with fixed variables. Linear structure-activity relationships can include other variables, such as topological indices or quantum-chemical parameters of molecules.

Free-Wilson's model.

This model assumes that the value of the property of this substituent, which is in the base structure in this position is always the same, regardless of which the compound is present in question [20]. The values of the substituents are calculated using multiple linear regression analysis. This requires only information about the molecular structure and biological activity of the compounds (physico-chemical parameters are not used).

When analyzing the data by the Free - Wilson method, a linear equation is drawn up for each compound, the regression coefficients are calculated by the smallest square method. An important result is that with the help of calculated parameter values, you can predict the activity of compounds formed by all sorts of combinations and permutations of the initial substituents in the main structure of this chemical series.

Linear discriminant analysis.

This method is one of the first approaches to the classification of objects, proposed by R. Fisher in 1936. It is based on the assumption that classes have a normal distribution; the boundaries between classes can be approximated by linear functions; the problem is reduced to estimating the parameters of these functions. The independent variables (descriptors) that should be related to the property in question are selected so as to provide the maximum possible separation of positive and negative properties: activity and inactivity [21]. Then the quality of separation is evaluated, and the results are analyzed.

General view of the system of discriminant functions:

$$\begin{cases} g_1 = a_{0} + a_1 X_1 + \ldots + a_{M1} X_M \\ g_2 = a_{0} + a_2 X_1 + \ldots + a_{M2} X_M \end{cases}$$

где g_1, g_2 – discriminant functions, reflect the probability density of the object belonging to class 1 or 2.

X_i – the independent variables, $i = 1, \ldots \dot{I}$.

a_0 , a_0 , a_{i1}, a_{i2} – discriminant coefficients, $i = 1, \ldots \dot{I}$.

When classifying a compound according to the values of its chemical structure parameters, the values of discriminant functions are calculated. If g1>g2, the compound is assigned to class 1 (e.g., highly active); if g1<g2 - to class 2 (Franke M., 1976).

Bayesian classifier.

This classifier is based on Bayes' rule for conditional [22].

Let $\{A_1, A_2\}$ - a complete group of events, consisting of the fact that compound C belongs to class 1 (active compounds) or class 2 (inactive compounds) with a priori probabilities $P(A_1)$ and $P(A_2)$. Let B_i an event with probability $P(B_i) > 0$, which consists in the fact that the descriptor «i» is present in the structure of connection C. Then the «a» posteriori probability that compound C belongs to the class of active A_1 , provided that the event B_i, is

$$P(A_1 \mid B_i) = \frac{P(B_i \mid A_1) \cdot P(A_1)}{P(B_i \mid A_1) \cdot P(A_1) + P(B_i \mid A_2) \cdot P(A_2)}$$

where $P(B_i \mid A_k)$ – is the conditional probability of event B_i, calculated under the assumption that the event occurred $A_k, k = 1, 2$.

The classifier is based on two assumptions:

1) the descriptors in the training sample compounds are equally important.

2) the descriptors are independent of each other.

Prior to the start of the classification $P(A_1) = P(A_2) = 0.5$.

With this in mind, if a connection C contains L descriptors, then the total probability of its belonging to the class of active connections A_1 is determined by the multiplication of conditional probabilities.

$$P(C \in A_1) = \frac{\prod_{i=1}^{L} P(B_i \mid A_1)}{\prod_{i=1}^{L} P(B_i \mid A_1) + \prod_{i=1}^{L} P(B_i \mid A_2)}$$

Bayesian method is used in PASS activity prediction (Merz, K.M., 2015).

Artificial neural networks.

Artificial neural networks are mathematical models, as well as their software or hardware implementations, built on the principle of organization and functioning of biological neural networks, that is, networks of nerve cells of a living organism [23]. General architecture of the neural network includes input neurons (whose responses are usually independent variables Xi),

several layers of hidden neurons (number of layers and number of hidden neurons in each layer are determined in the process of iterative learning) and output neuron (usually a common response function Y - for example, activity level) (Figure 1).

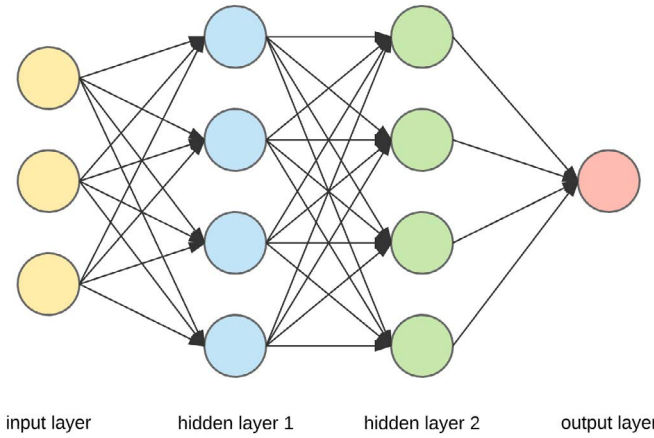

Figure 1. *An example of a three-layer artificial neural network architecture.*

The essence of the method is to iteratively determine weights which connect output neuron's response with signals from hidden neurons, and responses from hidden neurons - with signals from hidden neurons of previous level or from input neurons. Kind of function $Y=f(X)$ can be any, but most often we use linear form of dependence, because of easy implementation of learning algorithm. Initially the values of all weights are set equal. At each step of iteration increase or decrease the value of each weight and conduct classification of training and test samples. If a change in a particular weight improves the classification, it is rewarded, and if it worsens results, it is penalized. The procedure is repeated until minimum error is achieved in predicting test sample objects.

Artificial neural network is one of the most promising developing methods, which has the following advantages:

multilayer neural networks allow you to build any dependencies - linear, nonlinear, discrete, mixed.

the method does not impose strict restrictions on the quantitative composition of the sample.

As an illustration of successful application of neural networks for the search of biologically active substances a review can be given.

Support vector machine, SVM.

The Support Vector Machine (SVM) method is based on the transformation of raw data using so-called kernel functions [24]. The kernel function is chosen so that the classes of objects in the transformed data space are linearly separable. The search for the optimal dividing line is performed by constructing two parallel hyperplanes on both sides of the hyperplane separating the classes in question. The optimal dividing hyperplane is the one for which the distance to the two parallel hyperplanes is maximal. The error is minimized by using an iterative learning algorithm. An example of using SVM for biological activity prediction is a study.

Comparative molecular field analysis.

One of the most well-known methods of 3D-QSAR is Comparative Molecular Field Analysis (CoMFA- Comparative Molecular Field Analysis). CoMFA allows to calculate relationships between three-dimensional structure of molecules and their biological activity [25].

The essence of the method is as follows. 3D models are constructed and optimized for all compounds of the training sample. For each molecular model, the center of a molecule is determined (or a common basic structure is used). All models of the training sample are aligned with each other by this center (common structure), minimizing the sum of squares of atomic deviations. Around the resulting overlay, a parallelepiped is formed, inside which all the matched molecules are located. This parallelepiped is divided into cells of constant size (usually 2x2x2 A). Each molecule placed inside such a lattice determines the parameter values of all lattice cells. As a rule, these are electronic, steric, and lipophilic increments in the total index of the properties of a given compound. The parameters of all cells are used as independent variables in a stepwise multiple regression analysis, by which the final QSAR equation is calculated. An example of the use of CoMFA is given in.

Conclusions.

Based on the above data, it is clear that the rational approaches used for any paradigm of creating new antidiabetic drugs can be divided onto two categories - the search for new targets of action and exploration of their ligands or their analogues. This points to the applicability of molecular studies to analyze receptor-drug interactions, studying the pharmacokinetic features of potential antidiabetic drugs. Target identification and structural analysis are expected to facilitate the search for new drugs.

Thus, a directed search among new compounds with high pharmacotherapeutic activity, including the use of *in silico* methods, is very relevant and in demand. This research will make it possible to identify an innovative compound substance based on which a highly effective competitive drug for the treatment of a certain disease will be created.

The above data testify to the importance of *in silico* methods for successful search of new pharmacologically active substances.

REFERENCES

1. Sosonyuk SE, Peshich A, Tutushkina AV, et al. Synthesis and cytotoxicity of novel simplified eleutherobin analogues as potential antitumour agents. Org Biomol Chem. 2019;17:2792-2797.

2. Hansch C, Fujita T. p-σ-π Analysis. A method for the correlation of biological activity and chemical structure. J Am Chem Soc. 1964;86:1616-1626.

3. Gajewicz-Skretna A, Kar S, Piotrowska M, et al. The kernel-weighted local polynomial regression (KwLPR) approach: an efficient, novel tool for development of QSAR/QSAAR toxicity extrapolation models. J Cheminform. 2021;13:9.

4. Yakovleva O, Bogatova K, Mukhtarova R, et al. Hydrogen sulfide alleviates anxiety, motor, and cognitive dysfunctions in rats with maternal hyperhomocysteinemia via mitigation of oxidative stress. Biomolecules. 2020;10:995.

5. Moreira LS, Chagas BC, Viana Pacheco CS, et al. Development of procedure for sample preparation of cashew nuts using mixture design and evaluation of nutrient profiles by Kohonen neural network. Food Chem. 2019;273:136-143.
6. Lo Y-C, Rensi SE, Torng W, et al. Machine learning in chemoinformatics and drug discovery. Drug discovery today. 2018;23:1538-1546.
7. Baskin II. Practical constraints with machine learning in drug discovery. Expert Opinion on Drug Discovery. 2021;16:929-931.
8. Majewski M, Lis B, Juśkiewicz J, et al. Phenolic fractions from dandelion leaves and petals as modulators of the antioxidant status and lipid profile in an in vivo study. Antioxidants. 2020;9:131.
9. Friesner J. Labor in the food system: A view from INFAS. Journal of Agriculture, Food Systems, and Community Development. 2016;6:25-27.
10. Swindells MB, Porter CT, Couch M, et al. abYsis: Integrated antibody sequence and structure—management, analysis, and prediction. J Mol Biol. 2017;429:356-364.
11. Gaulton A, Hersey A, Nowotka M, et al. The ChEMBL database in 2017. Nucleic Acids Res. 2017;45:D945-D954.
12. Dinparast L, Valizadeh H, Bahadori MB, et al. Design, synthesis, α-glucosidase inhibitory activity, molecular docking and QSAR studies of benzimidazole derivatives. Journal of Molecular Structure. 2016;1114:84-94.
13. Палюлин АМ. Анализ данных из изображений с применением нейронной сети. Россия-Азия-Африка-Латинская Америка: экономика взаимного доверия. 2019:130-131.
14. Combettes PL, Pesquet JC. Deep neural network structures solving variational inequalities. Set-Valued and Variational Analysis. 2020:1-28.
15. Rybińska-Fryca A, Sosnowska A, Puzyn T. Representation of the Structure—A Key Point of Building QSAR/QSPR Models for Ionic Liquids. Materials. 2020;13:2500.
16. Li P, Merz Jr KM. Metal ion modeling using classical mechanics. Chemical reviews. 2017;117:1564-1686.
17. Mekenyan O. In silico toxicology: principles and applications. Royal Society of Chemistry. 2010.
18. Dei Cas A, Khan SS, Butler J, et al. Impact of diabetes on epidemiology, treatment, and outcomes of patients with heart failure. JACC: Heart Fail. 2015;3:136-145.
19. Uyanık GK, Güler N. A study on multiple linear regression analysis. Procedia-Social and Behavioral Sciences. 2013;106:234-240.
20. Chen H, Carlsson L, Eriksson M, et al. Beyond the scope of Free-Wilson analysis: building interpretable QSAR models with machine learning algorithms. Journal of chemical information and modeling. 2013;53:1324-1336.
21. Wu L, Shen C, Van Den Hengel A. Deep linear discriminant analysis on fisher networks: A hybrid architecture for person re-identification. Pattern Recognition. 2017;65:238-250.
22. Mohajerani H, Kholghi M, Mosaedi A, et al. Application of Bayesian Decision Networks for groundwater resources management under the conditions of high uncertainty and data scarcity. Water Resources Management. 2017;31:1859-1879.
23. Davoudi Kakhki F, Freeman SA, Mosher GA. Use of neural networks to identify safety prevention priorities in agro-manufacturing operations within commercial grain elevators. Applied Sciences. 2019;9:4690.
24. Huang S, Cai N, Pacheco PP, et al. Applications of support vector machine (SVM) learning in cancer genomics. Cancer genomics & proteomics. 2018;15:41-51.
25. Ansari SM, Palmer DS. Comparative molecular field analysis using molecular integral equation theory. Journal of chemical information and modeling. 2018;58:1253-1265.

ნარკოტიკების განვითარება სილიკოს მეთოდით

ლენა კ[1,2], ბაღათურია გ[3], ბუინოვი ლ[4], ლებედევი ა[5], გრიშინი იმ[6,7],

პროშინი ერთად[4]

[1]პეტერბურგის სახელმწიფო უნივერსიტეტი, რუსეთი.

[2]სანქტ-პეტერბურგის სახელმწიფო ქიმიისა და ფარმაციის უნივერსიტეტი, რუსეთი.

[3]პეტერბურგის სახელმწიფო პედიატრიული სამედიცინო უნივერსიტეტი, რუსეთი.

[4]ა. ი. ჰერცენის რუსეთის სახელმწიფო პედაგოგიური უნივერსიტეტი, პეტერბურგი, რუსეთი.

[5]ექსპერიმენტული მედიცინის ინსტიტუტი, პეტერბურგი, რუსეთი.

[6]ეროვნული სახელმწიფო უნივერსიტეტის ფიზიკური კულტურის, სპორტისა და ჯანმრთელობა სახელობის P. F. Lesgaft, პეტერბურგი, რუსეთი.

[7]ი. პ. პავლოვის სახელობის სტომატოლოგიის სამეცნიერო-კვლევითი ინსტიტუტი და ChLH PSPBSMU, პეტერბურგი, რუსეთი.

რეზიუმე.

ბოლო ათწლეულის მრავალცენტრულმა კვლევებმა დაადასტურა პირდაპირი კავშირი სამკურნალო ნაერთების სივრცულ სტრუქტურისა და მათ ბიოლოგიურ აქტიურობის შორის.

თემის აქტუალობა განპირობებულია ერთი მხრივ ახალი ტექნოლოგიების სწრაფი განვითარებით, მეორეს მხრივ, დიდ რაოდენობითი კვლევის სამუშაოებით, რომლებიც მიძღვნილია სტერეოიზომერული ფაქტორების გავლენების, თერაპიულ და სამკურნალო ნივთიერებების გვერდით ეფექტების შესწავლას. ბოლო რამოდენიმე ათწლეულის განმავლობაში, გამოთვლითი მიდგომები იქნა გამოყენებული ნივთიერებების ბიოლოგიურ თვისებების და მათ სტრუქტურის შორის კავშირის რაოდენობრივების დასადგენად. თანამედროვე მედიცინაში ფართოდ გამოიყენება მიმართულება - „სტრუქტურა-აქტივობის რაოდენობრივი ურთიერთობები“ (QSAR). ამ მიმართულებამ მოლეკულური მოდელირების მეთოდებთან ერთად მიიღო ზოგადი სახელწოდება in silico მეთოდი, რაც იგულისხმებს წამლების წინასწარ ძიებას მხოლოდ კომპიუტერული მეთოდით ექსპერიმენტულისკრინინგისჩატარებამდე.მსოფლიოში შემუშავებულია ასობით QSAR და მოლეკულური მოდელირების პროგრამა. გამოქვეყნებულია ათობით სპეციალიზებული სამეცნიერო ჟურნალები, რომლებიც შეიცავს ბიოლოგიურად აქტიური ნივთიერებების

ძიების კომპიუტერულ მეთოდებს. გამოქვეყნებულია ათასობით სტატიები, რომლებიც აჩვენებს QSAR მეთოდის წარმატებული გამოყენების შედეგებს ახალი ბიოლოგიურად აქტიური ნივთიერებების საძიებლად. ამჟამად მსოფლიოში ცნობილია 101 მილიონზე მეტი ორგანული ნაერთი, ხოლო ბიოლოგიური აქტივობების ცნობილი რაოდენობა 25 ათასს აჭარბებს, ამიტომ ყველა ნივთიერების გლობალური ექსპერიმენტული შემოწმება შეუძლებელია.

ამრიგად, in silico მეთოდების გამოყენება წამლების გამომჟღავნებისათვის, კერძოდ, არაპერსპექტიული ნაერთების წინასწარ შერჩევისთვის, არის ძალზე მნიშვნელოვანი.

საძიებო სიტყვები. In silico მეთოდებში, «შედარებითი სტრუქტურა - საქმიანობის ურთიერთობა», სტერეომეტრიული ფაქტორები.

РАЗРАБОТКА ЛЕКАРСТВ МЕТОДОМ IN SILICO

Ленская К.В[1,2], Багатурия Г.О[3], Буйнов Л.Г[4], Лебедев А.А[5],

Гришин В.В[6,7], Прошин С.Н[4].

[1]Санкт-Петербургский государственный университет, Россия.

[2]Санкт-Петербургский государственный химико-фармацевтический, университет, Россия.

[3]Санкт-Петербургский государственный педиатрический медицинский, университет, Россия.

[4]Российский государственный педагогический университет им. А.И. Герцена, Санкт-Петербург, Россия.

[5]Институт экспериментальной медицины, Санкт-Петербург, Россия.

[6]Национальный государственный университет физической культуры, спорта и здоровья имени П. Ф. Лесгафта, Санкт-Петербург, Россия.

[7]Научно-исследовательский институт стоматологии и ЧЛХ ПСПБГМУ им. И.П. Павлова, Санкт-Петербург, Россия.

Резюме

Многоцентровые исследования последнего десятилетия подтвердили непосредственную взаимосвязь между пространственным строением лекарственных соединений и их биологической активностью.

Цель. Актуальность данной тематики, с одной стороны, объясняется быстрым развитием новых технологий, с другой – большим количеством научно-исследовательских работ, посвященных изучению влияния стереоизомерных факторов на терапевтическое действие и побочные эффекты лекарственных веществ.

Методы. Последние десятилетия применяются вычислительные подходы, позволяющие выявить связь между биологическими свойствами веществ и их структурой и количественно описать эту связь. В современной химии лекарств широко применяется направление - «Quantitative Structure - Activity Relationships» (QSAR). В сочетании с методами молекулярного моделирования это направление получило общее название методов in silico, что подразумевает предварительный поиск лекарств только компьютерным способом, до проведения экспериментального скрининга.

Результаты. В мире разработаны сотни программ QSAR и молекулярного моделирования: издаются десятки специализированных научных журналов, посвященных компьютерным методам поиска биологически активных веществ. Опубликованы тысячи статей, демонстрирующих результаты успешного применения методов QSAR в поиске новых биологически активных веществ. В настоящее время в мире известно более 101 миллиона органических соединений, а известное количество биологических активностей превышает 25 тысяч, поэтому глобальное экспериментальное тестирование всех веществ невозможно.

Выводы. Таким образом, использование методов in silico для поиска лекарств, в частности, для предварительного отбора бесперспективных соединений является сверхактуальным.

Ключевые слова. In silico методы, «Сравнительная структура – взаимосвязь активности», стереометрические факторы.