

ПРОГНОЗ СПЕКТРОВ БИОЛОГИЧЕСКОЙ АКТИВНОСТИ И ВИРТУАЛЬНЫЙ СКРИНИНГ

Д. А. Филимонов, В. В. Поройков

Научно-исследовательский институт биомедицинской химии им. В. Н. Ореховича, Москва

E-mail: dmitry.filimonov@ibmc.msk.ru <http://www.ibmc.msk.ru>



Российский
научный
фонд

Исследование выполнено при поддержке Российского научного фонда грант No. 19-15-00396

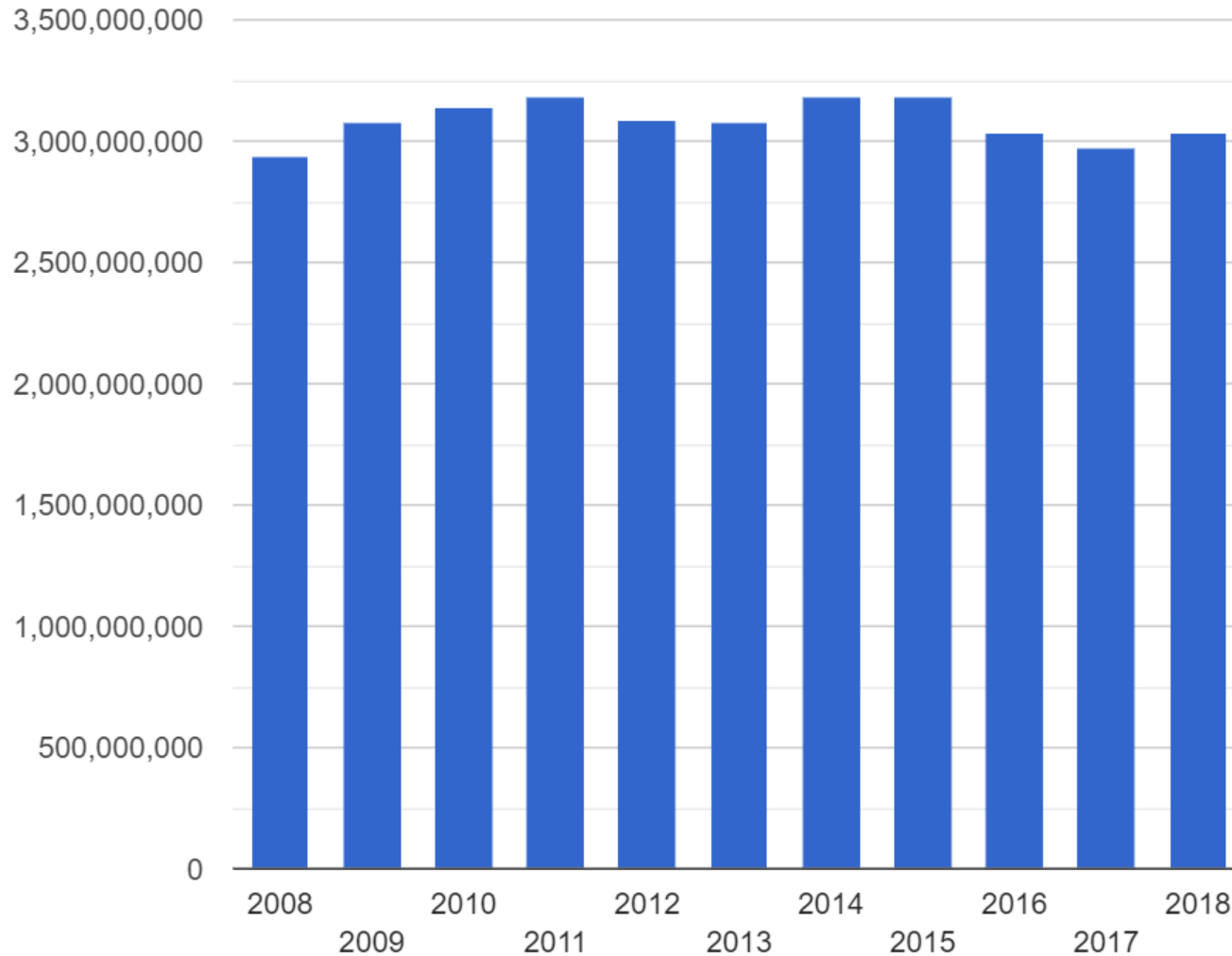
- ▶ About 3,000,000,000 total prescriptions
- ▶ About 500 unique active ingredients

ClinCalc DrugStats Database

Free U.S. Outpatient Drug Usage Statistics

<http://clincalc.com/DrugStats>

Total Prescriptions in DrugStats Database by Year



Drug Rank	Percent of all Prescriptions
1	4%
25	41%
50	58%
75	68%
100	76%
125	81%
150	85%
175	88%
200	91%
225	93%
250	95%
275	96%
300	97%



Количество зарегистрированных лекарств (1899-2018 г.г.)

В среднем: **~80 препаратов ежегодно (1999-2018 г.г.)**

Источник: *Clarivate Analytics Integrity*

Примеры отзыва лекарств с рынка из-за побочных эффектов

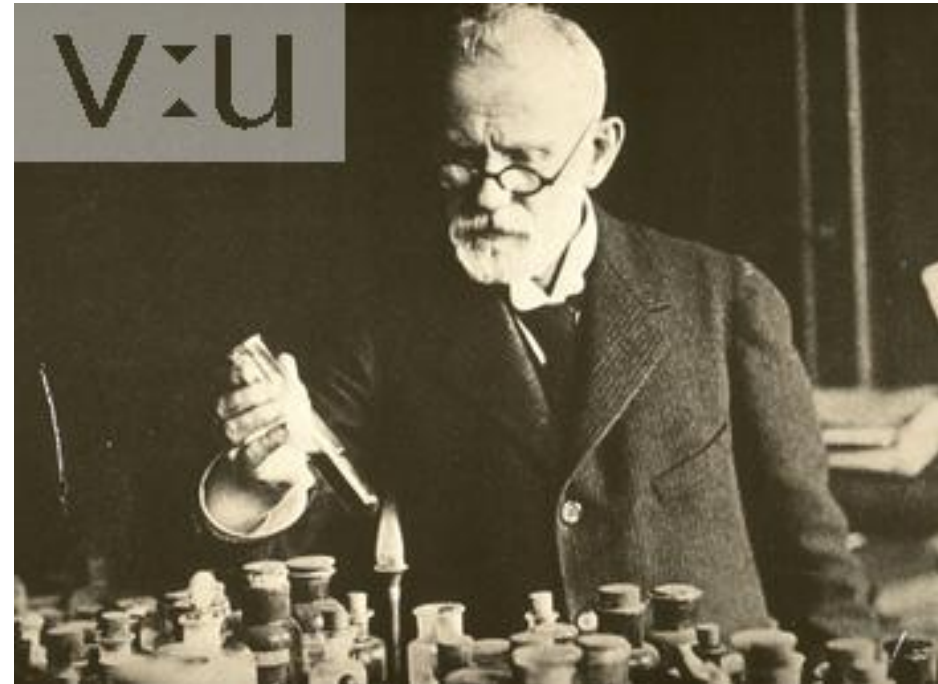
Препарат	Показание	Побочные эффекты	Появление на рынке	Отзыв с рынка
Bromfenac (Duract)	NSAID for ophthalmic use	Hepatotoxicity	1997	1998
Alosetron (Lotronex)	Irritable bowel syndrome	Ischemic colitis	2000	2000
Troglitazone (Rezulin)	Type II diabetes	Hepatotoxicity	1997	2000
Cisapride (Propulsid)	Gastric reflux	Cardiotoxicity	1993	2000
Cerivastatin (Baycol)	Cholesterol lowering	Rhabdomyolysis	1997	2001
Rofecoxib (Vioxx)	NSAID	Heart attack and stroke	1999	2004

Побочное действие лекарств является пятой ведущей причиной смерти в США.

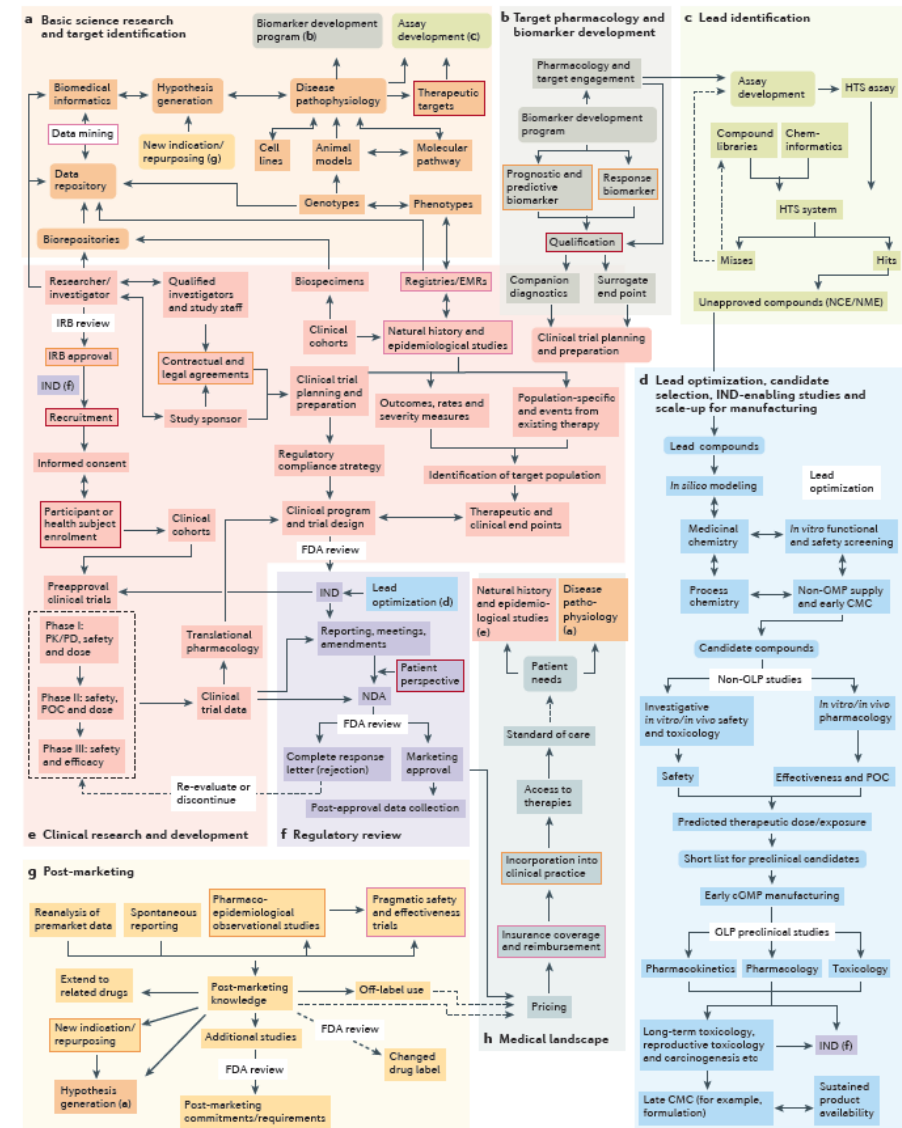
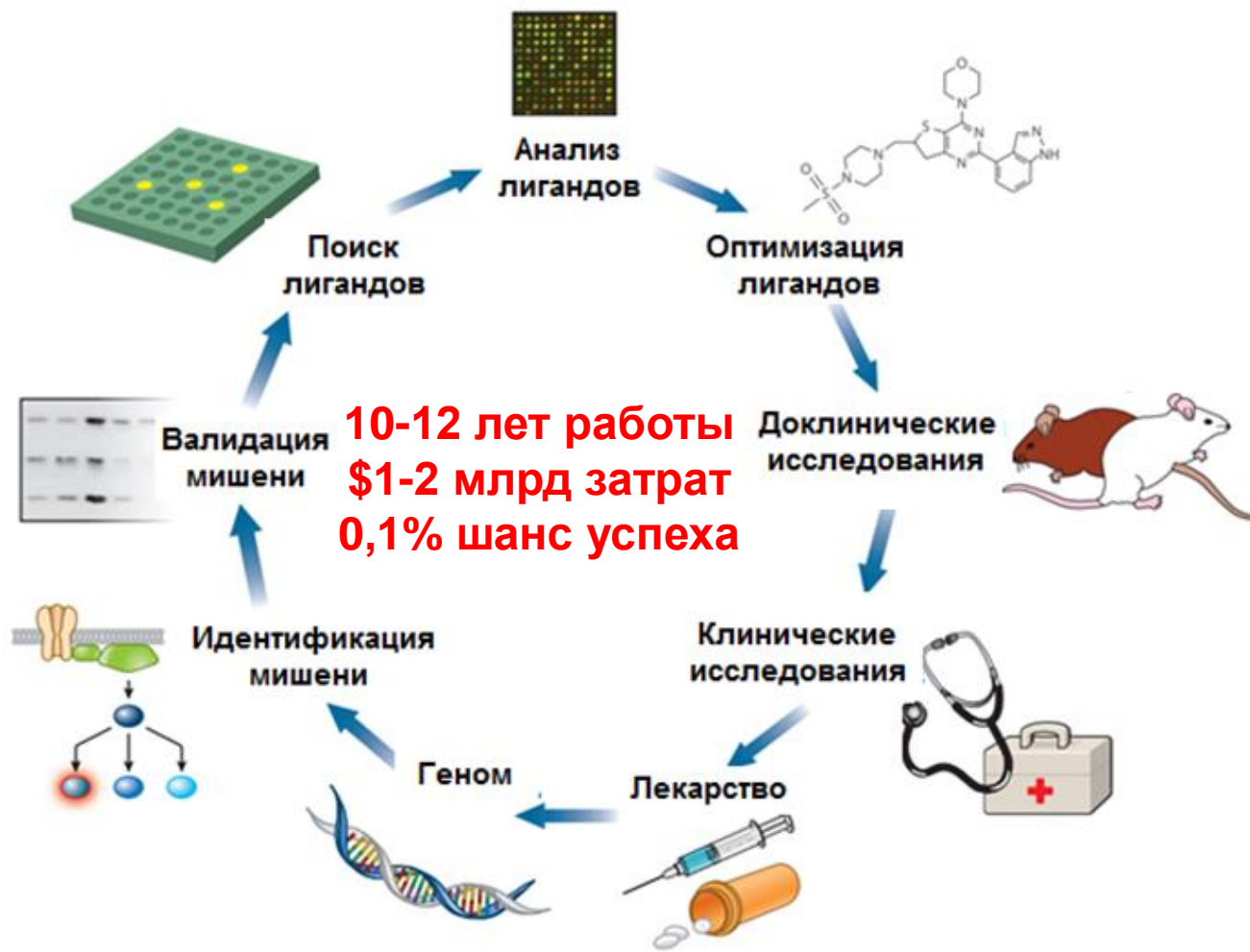
<http://www.fda.gov/medwatch>

Начало XX века: представления о «магической пуле»

В XX веке ведущая парадигма, на которой основывалось создание новых лекарств, базировалась на предположении о том, что селективно воздействуя на определенную фармакологическую мишень, можно достичь нормализации патологического процесса.



Пауль Эрлих
1854-1915

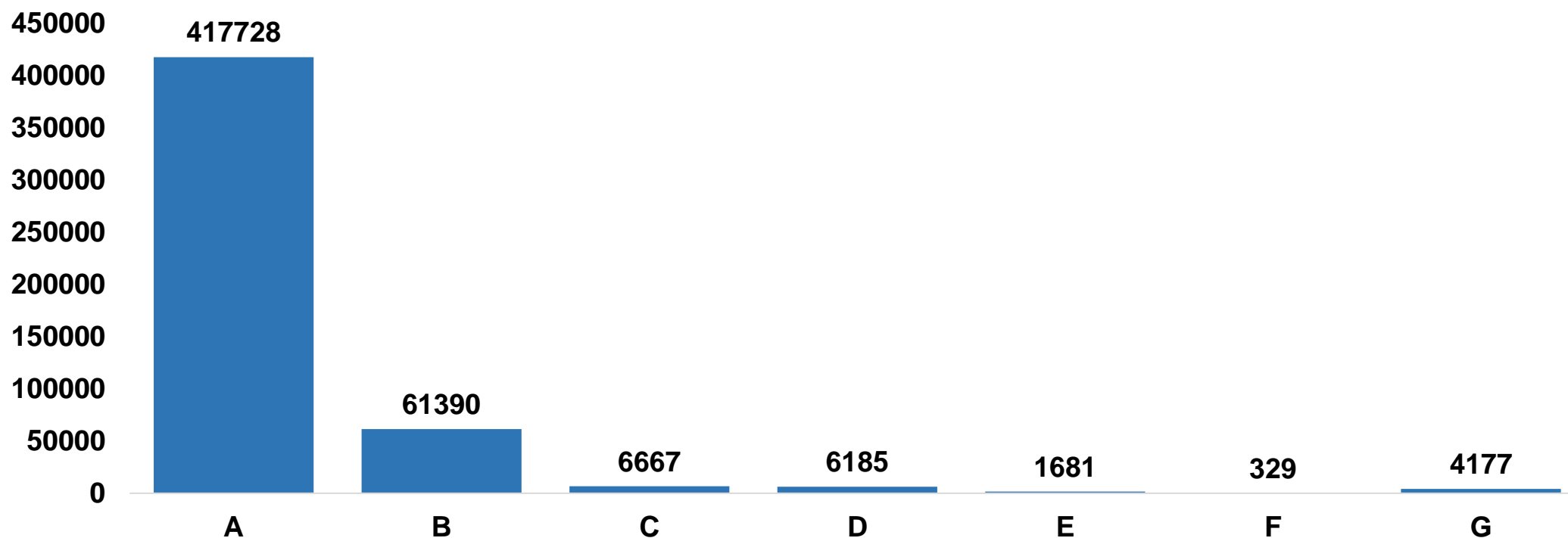


Yap и Workman, *Annu.Rev.Pharmacol.Toxicol.*, 2012

Wagner J. et al. *Nature Reviews Drug Discovery*, 2017. doi: 10.1038/nrd.2017.217

Poroikov V., Druzhilovskiy D. In: *In Silico Drug Design*, Ed. Runal Roy, Academic Press, 2019, p.3-17.

Поройков В.В. В кн.: *Фармацевтическая разработка*. Под ред. В.В. Береговых, Изд-во РАН, 2019.



Количество ХС, находящихся на различных стадиях тестирования с целью создания лекарственных препаратов:

A - биологические испытания;

B - доклинические исследования;

C, D и E - клинические испытания (фазы I, II и III, соответственно);

F - регистрация;

G – получено разрешение для медицинского применения

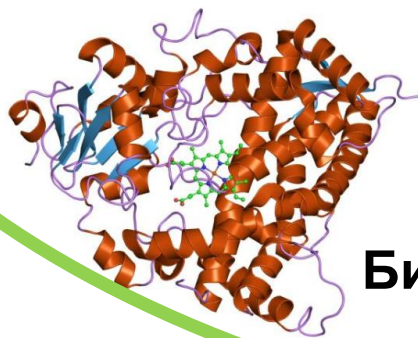
Большие данные химико-биологических взаимодействий



Болезнь

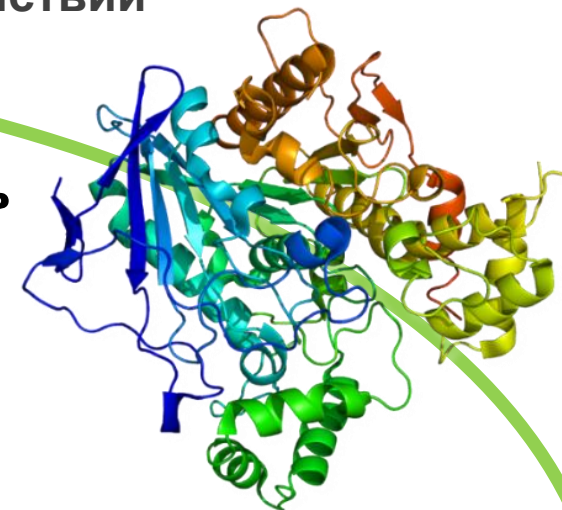
Лекарственно-подобных органических соединений

- 100 млн синтезировано
- 1.75 млрд SAVI-2020
- 166 млрд GDB-17
- 10^{30} – 10^{60} возможных

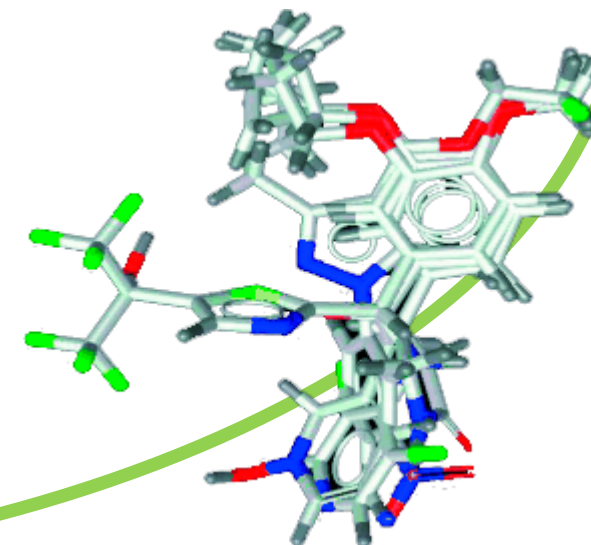


Биотрансформация

Мишень



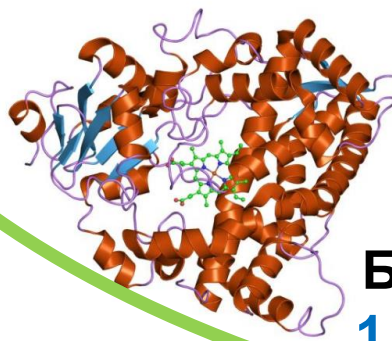
Лиганды



Большие данные химико-биологических взаимодействий



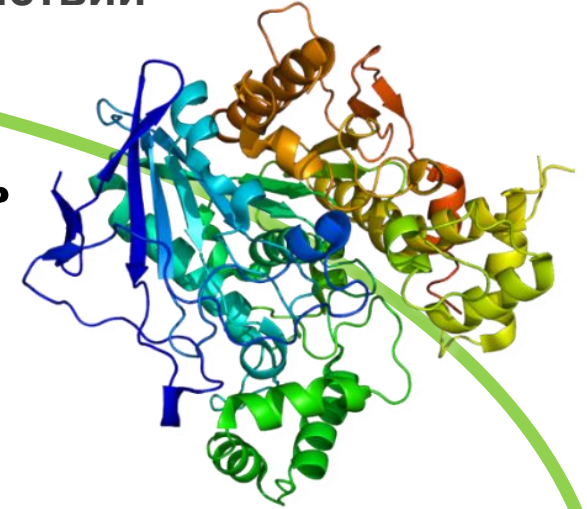
Болезнь
~500



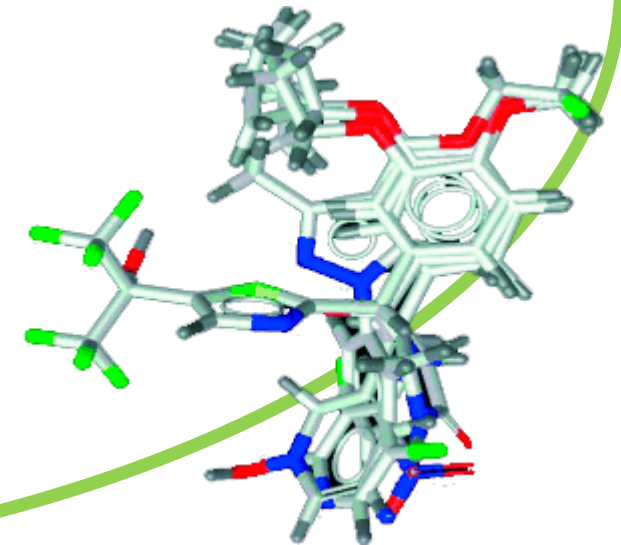
Биотрансформация
1 ~ >10 метаболитов



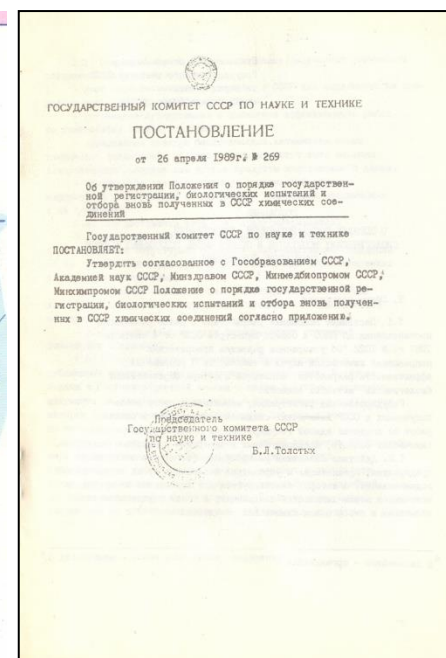
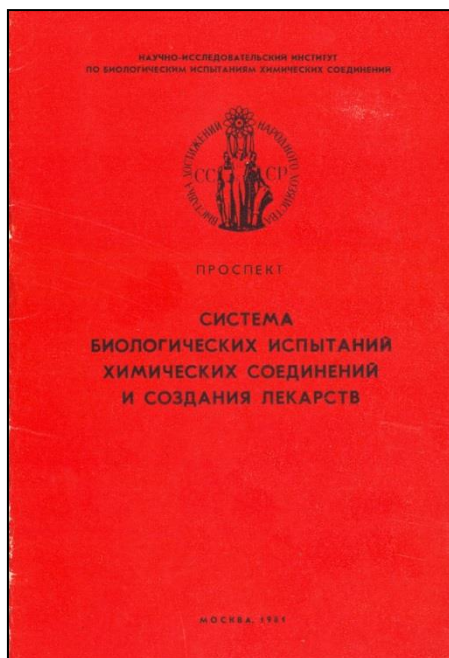
Мишень
~2 000



Лиганды
~100 000 000



Система государственной регистрации и биологических испытаний новых химических соединений, полученных в СССР (1972 – 1991)



1937 - 2013

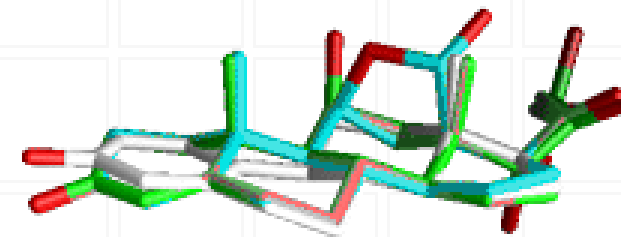
“The system of biological tests of physiologically active substances developed under the supervision of Piruzyan allows the use of relatively simple methods of investigation requiring minimal expenditures of time and substances to eliminate substances that are of little promise for pharmacology from a great number of new compounds and select those that may become medicinal drugs”.

Lev Aramovich Piruzyan (On His 70th birthday) <https://doi.org/10.1134/S0362119707050192>

Как отбирать наиболее перспективные вещества среди зарегистрированных соединений?

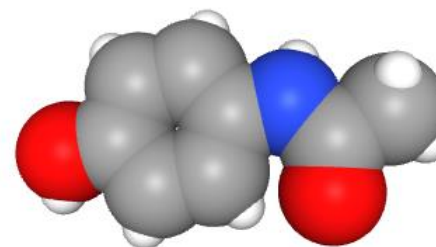
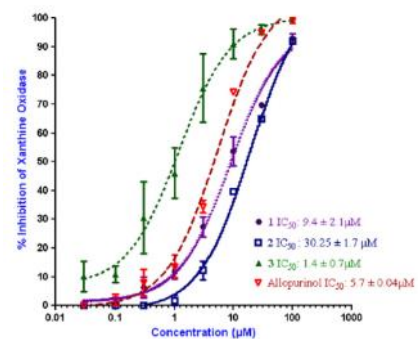
– Необходимо создать компьютерную систему, прогнозирующую профили биологической активности по структурной формуле химического соединения.

Основная догма SAR/QSAR/QSPR:

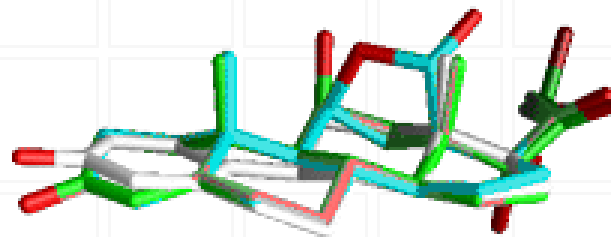


$$\text{Activity} = f(\text{Structure})$$

Активность = Функция (Структура)



Основная догма SAR/QSAR/QSPR:

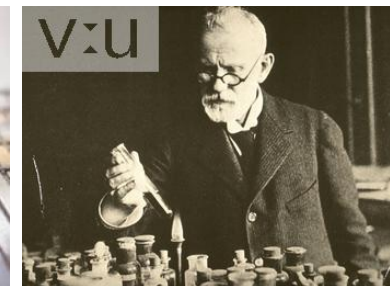
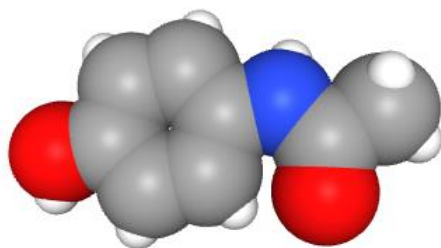


$$\text{Activity} = f(\text{Structure})$$

Активность = Функция (Структура, Мишень, Объект, Метод, Условия, ...)

26 of 550 Possible Pharmacological Effects :

0.879 0.008 Immunostimulant (HIV)
0.860 0.003 Antidote, cyanide
0.852 0.004 Alopecia treatment
0.788 0.014 Antiseborrheic
0.702 0.006 Anterior pituitary suppressant
0.683 0.022 Fibrinolytic
0.687 0.027 Sialagogue
0.662 0.004 Leukopoiesis inhibitor
0.693 0.049 Ionotropic
0.656 0.019 Mucositis treatment
0.652 0.025 Antiviral (Arbovirus)
0.625 0.007 Antipyretic
0.609 0.016 Gastroprotector
0.606 0.014 Immunomodulator (HIV)
0.592 0.008 Erythropoiesis stimulant
0.593 0.012 Leukopoiesis stimulant
0.579 0.029 Gastric secretion stimulant
0.565 0.018 Vasodilator, renal
0.546 0.004 Antiuremic
0.539 0.005 Allergic conjunctivitis treatment
0.545 0.023 Respiratory analeptic
0.529 0.026 Skin diseases treatment
0.517 0.016 Natural killer cell stimulant
0.516 0.018 Cardiovascular analeptic
0.512 0.026 Antiinflammatory, intestinal
0.515 0.030 Potassium sparing diuretic

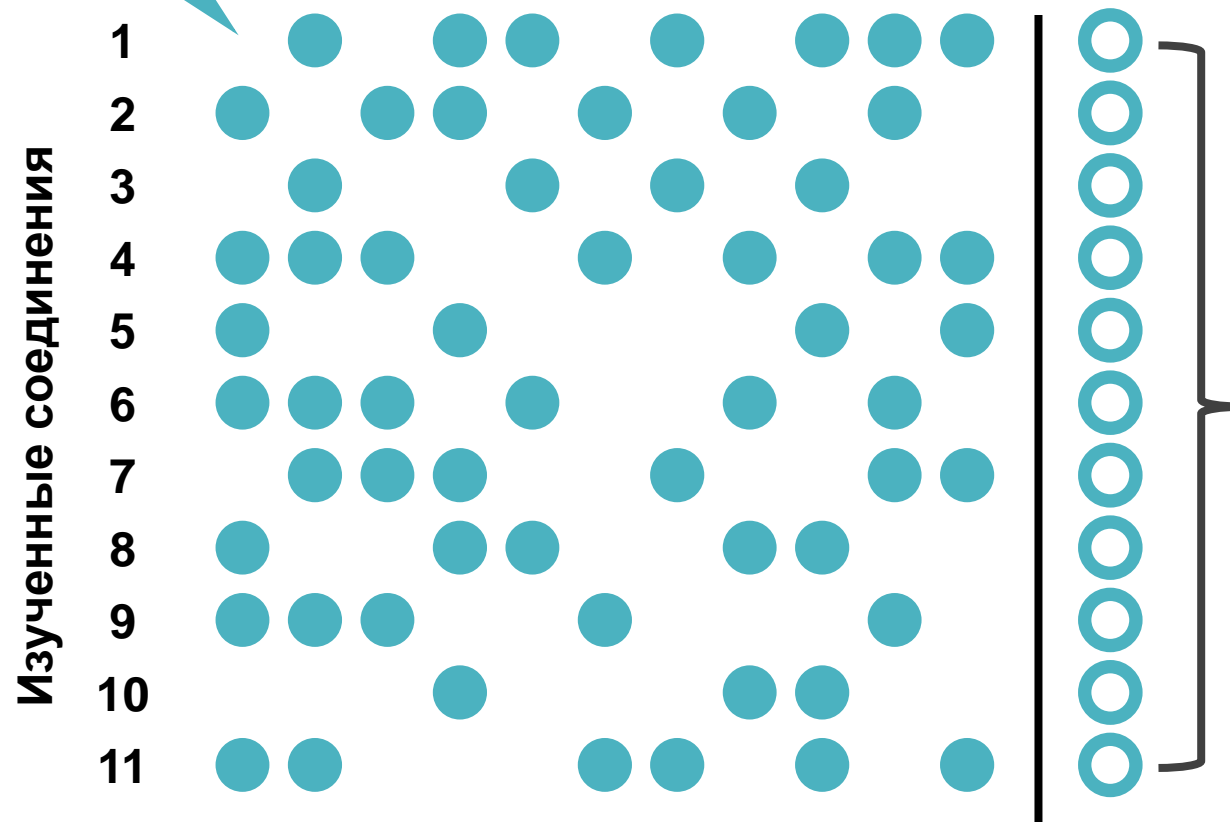


Tarasova O.A. et al., *J. Chem. Inform. Model.*, 2015, 55(7), 1388-99. doi: 10.1021/acs.jcim.5b00019.

Известные данные

Известные мишени A B C D E F G H I J K

Новая мишень



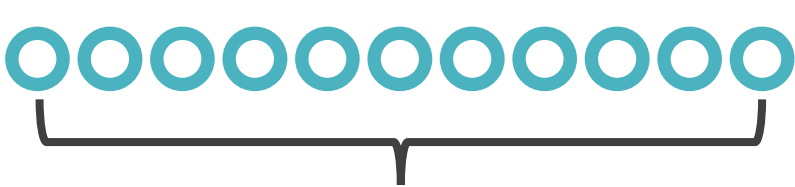
Виртуальный скрининг

Задача виртуального скрининга
 – отобрать среди множества возможных органических соединений наиболее отвечающие сформулированным требованиям.

Задача прогноза спектра биологической активности
 – отобрать среди множества возможных биологических активностей наиболее отвечающие сформулированным требованиям.

В обоих случаях требуется отобрать **наилучшее подмножество** из множества многих возможностей.

Новое соединение



Прогноз спектра активности

○ Оценка активности

= Функция (Структура, Мишень)

Poroikov, Filimonov. «Multitarget profiling» in: Muratov E.N. et al. QSAR without borders. Chemical Society Reviews, 2020, 49, 3525-3564. DOI: 10.1039/d0cs00098a.

Way2Drug PREDICTIVE SERVICES
 Understanding Chemical-Biological Interactions

Prediction of Activity Spectra for Substances



Refined Version 2020

Copyright © 1992-2020
 V.Poroikov, D.Filimonov & Associates
www.way2drug.com

1992
 Первая версия PASS:
 9 314 БАВ, 114 АК, 0.85 ТП

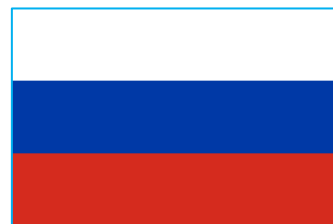
1998
 Новая версия PASS:
 31 000 БАВ, 435 АК, 0.94 ТП

1999
 Первая в Интернете программа PASS Online

2009
 Новая версия PASS:
 206 000 БАВ, 3 750 АК, 0.95 ТП

2021
 Новая версия PASS:
 1 370 000 БАВ, 7 500 АК, 0.95 ТП

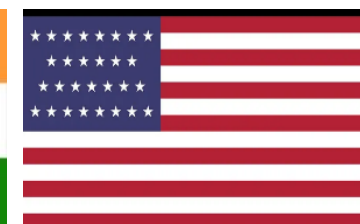
2021: 20 веб-сервисов
Статистика PASS Online:
 1 137 046 прогнозов;
 > 1 200 публикаций;
 30 758 пользователей из 104 стран мира



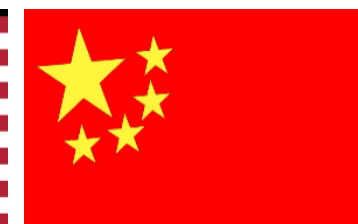
Россия (27,6%)



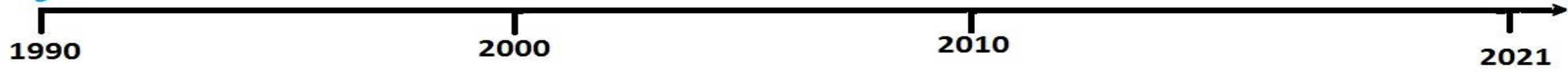
Индия (12,4%)



США (8%)



Китай (5,3%)



БАВ – к-во соединений в обучающей выборке; **АК** – к-во прогнозируемых видов активности; **ТП** – средняя инвариантная точность прогноза

C:\4444\Chemistry\World Wide Approved Drugs\WWAD2020.sdf

5x5 | 4x4 | 3x3 | 2x2 | Molecular Structure | MNA

Known | Effects | Mechanisms | Toxicity | Antitargets | Metabolism | Gene Expression

369

370

371

372

373

374

> <Name>
paracetamol

27 of 561 Possible Pharmacological Effects at Pa > 0.500

0.882	0.008	Immunostimulant (HIV)
0.853	0.004	Alopecia treatment
0.789	0.015	Antiseborrheic
0.714	0.012	Kidney function stimulant
0.688	0.013	Antineoplastic (basal cell carcinoma)
0.674	0.004	Leukopoiesis inhibitor
0.682	0.023	Fibrinolytic
0.694	0.050	Iontropic
0.654	0.026	Antiviral (Arbovirus)
0.627	0.015	Immunomodulator (HIV)
0.619	0.018	Mucositis treatment
0.603	0.008	Antipyretic
0.611	0.016	Gastroprotector
0.592	0.008	Erythropoiesis stimulant
0.594	0.012	Leukopoiesis stimulant
0.557	0.005	Allergic conjunctivitis treatment
0.565	0.017	Vasodilator, renal
0.547	0.004	Antidiuretic
0.528	0.017	Hepatitis alcoholic treatment
0.535	0.025	Respiratory analeptic
0.531	0.026	Skin diseases treatment
0.519	0.017	Natural killer cell stimulant
0.518	0.018	Cardiovascular analeptic
0.542	0.052	Hepatitis treatment
0.515	0.026	Antiinflammatory, intestinal
0.515	0.031	Potassium sparing diuretic
0.560	0.107	Antineoplastic (Kaposi's sarcoma)

24 Substructure Descriptors; 0 new. Mol.Mass: 151.16.
There are 775 known activities.
2348 of 8438 Possible Activities
27 of 561 Possible Pharmacological Effects
245 of 4617 Possible Mechanisms of Action
344 of 692 Possible Toxic and Adverse Effects
2 of 161 Possible Antitargets
24 of 225 Possible Metabolism-Related Actions
1719 of 2414 Possible Gene Expression Regulation
0 of 86 Possible Transporters-Related Actions

371/4223 | 0.882 0.008 Immunostimulant (HIV)

World Wide Approved Drugs (П. И. Савосина, Д. С. Дружиловский)

4 217 структур

167 292 активностей, найденных в PASS

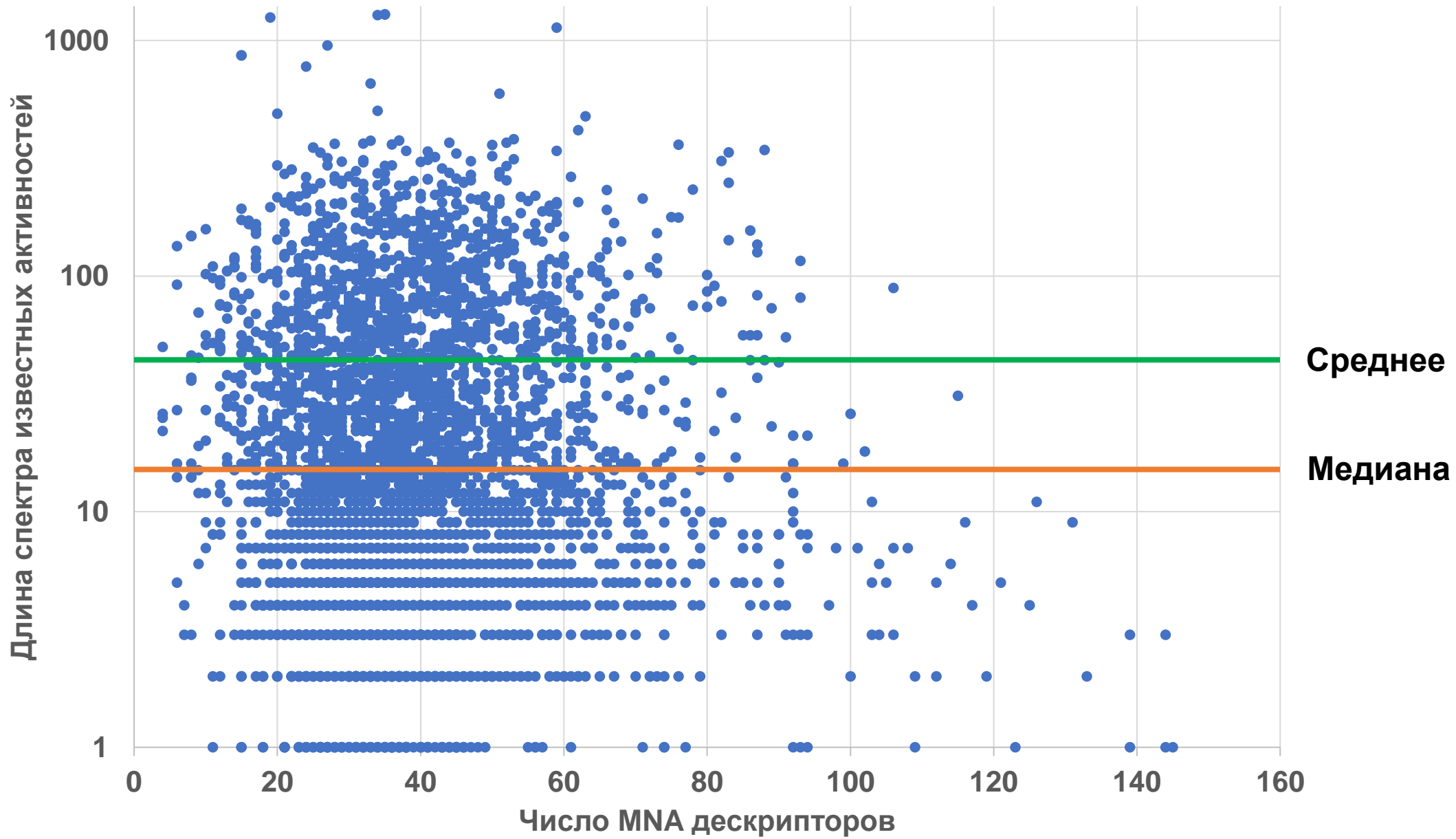
Известных активностей у одного лекарства:

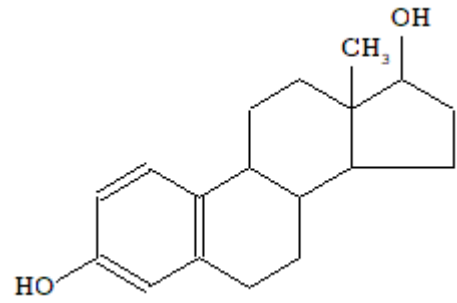
медиана 15

среднее 42

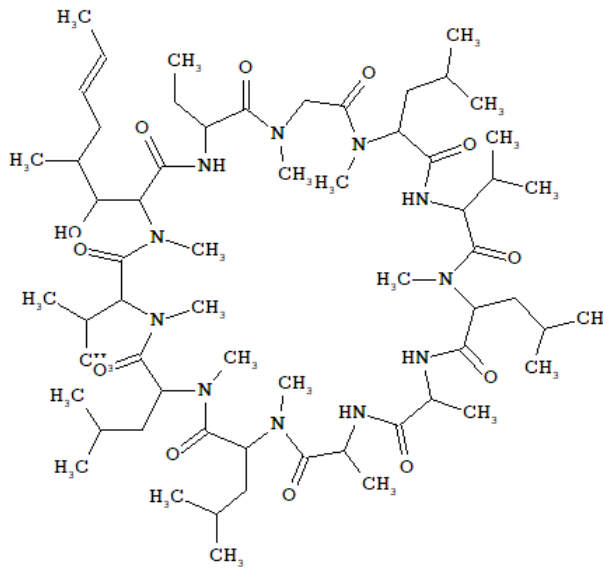
Категория	Записей	Активностей	Акт./Соед.	Соед./Акт.
Pharmacological Effects	21967	499	5.88	44.02
Mechanisms of Action	32810	3129	8.79	10.49
Toxic and Adverse Effects	73297	643	19.63	113.99
Antitargets	3325	144	0.89	23.09
Metabolism-Related Actions	7297	201	1.95	36.30
Gene Expression Regulation	16997	2355	4.55	7.22
Transporters-Related Actions	1731	43	0.46	40.26

World Wide Approved Drugs

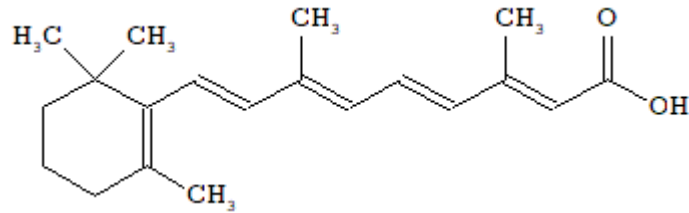




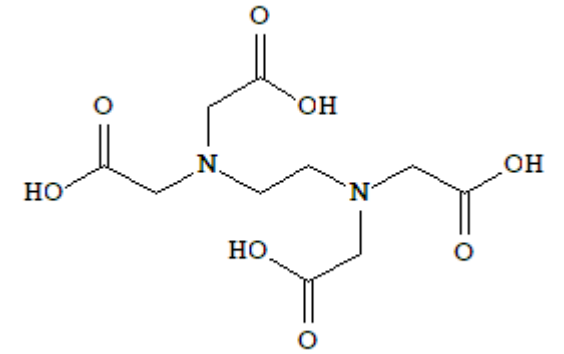
Estradiol 35 MNA 1289 AK



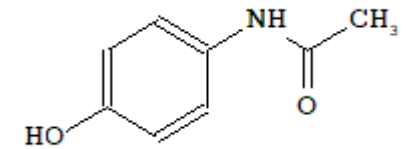
Ciclosporin 59 MNA 1134 AK



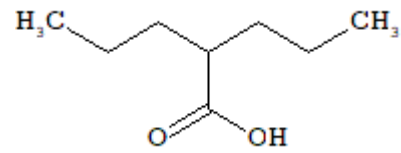
Tretinoin 34 MNA 1282 AK



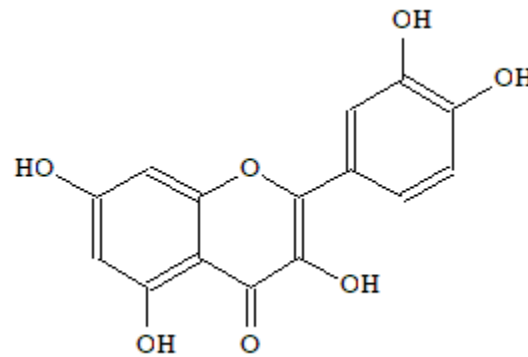
Edetic Acid 15 MNA 864 AK



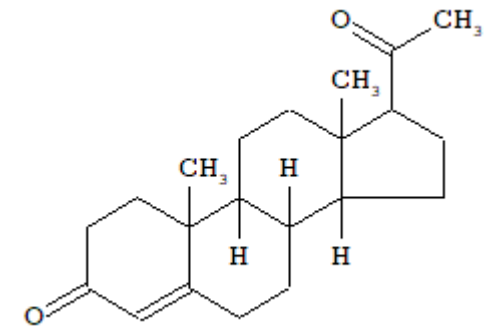
Paracetamol 24 MNA 775 AK



Valproic Acid 19 MNA 1251 AK

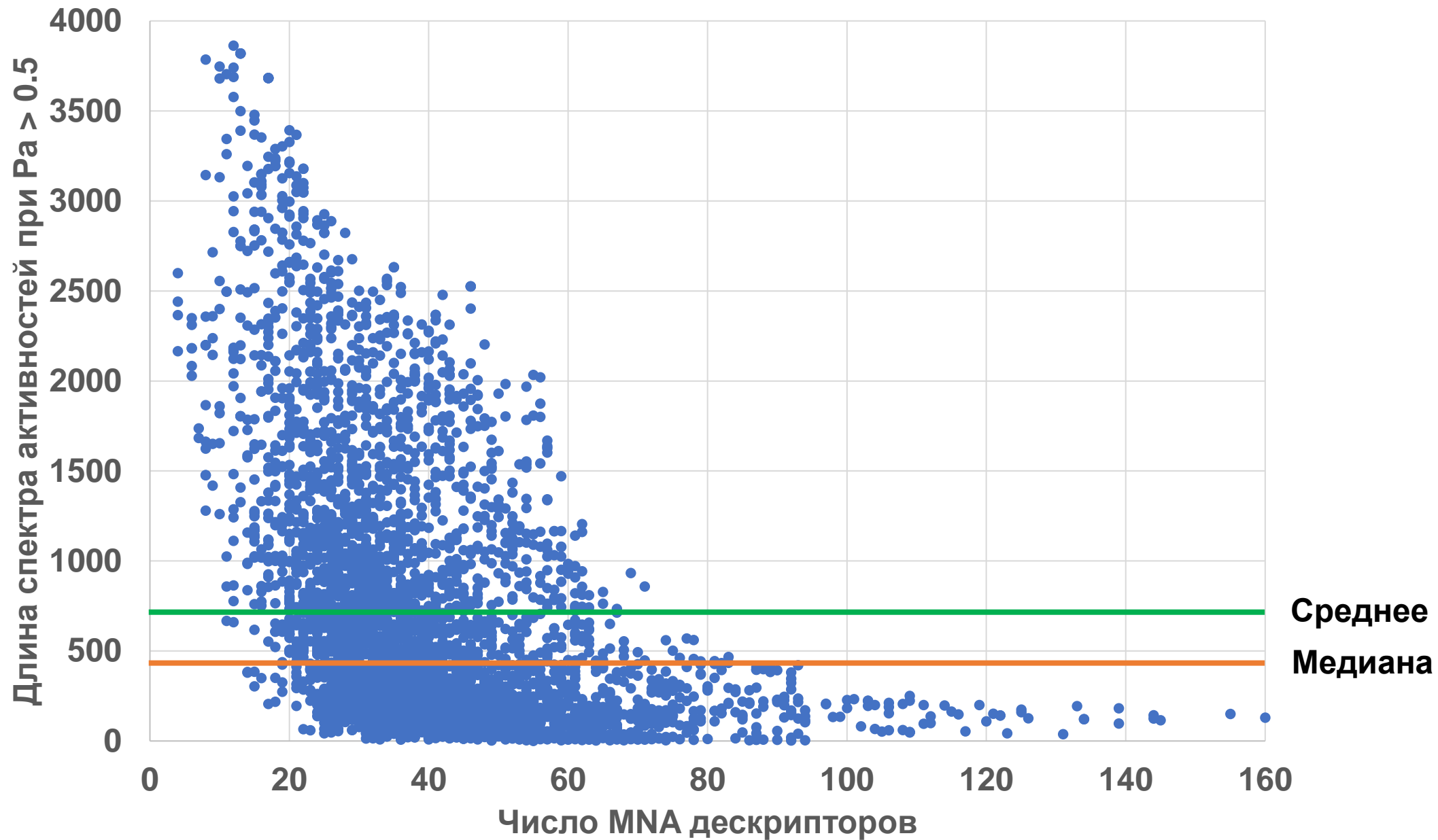


Quercetin 27 MNA 952 AK

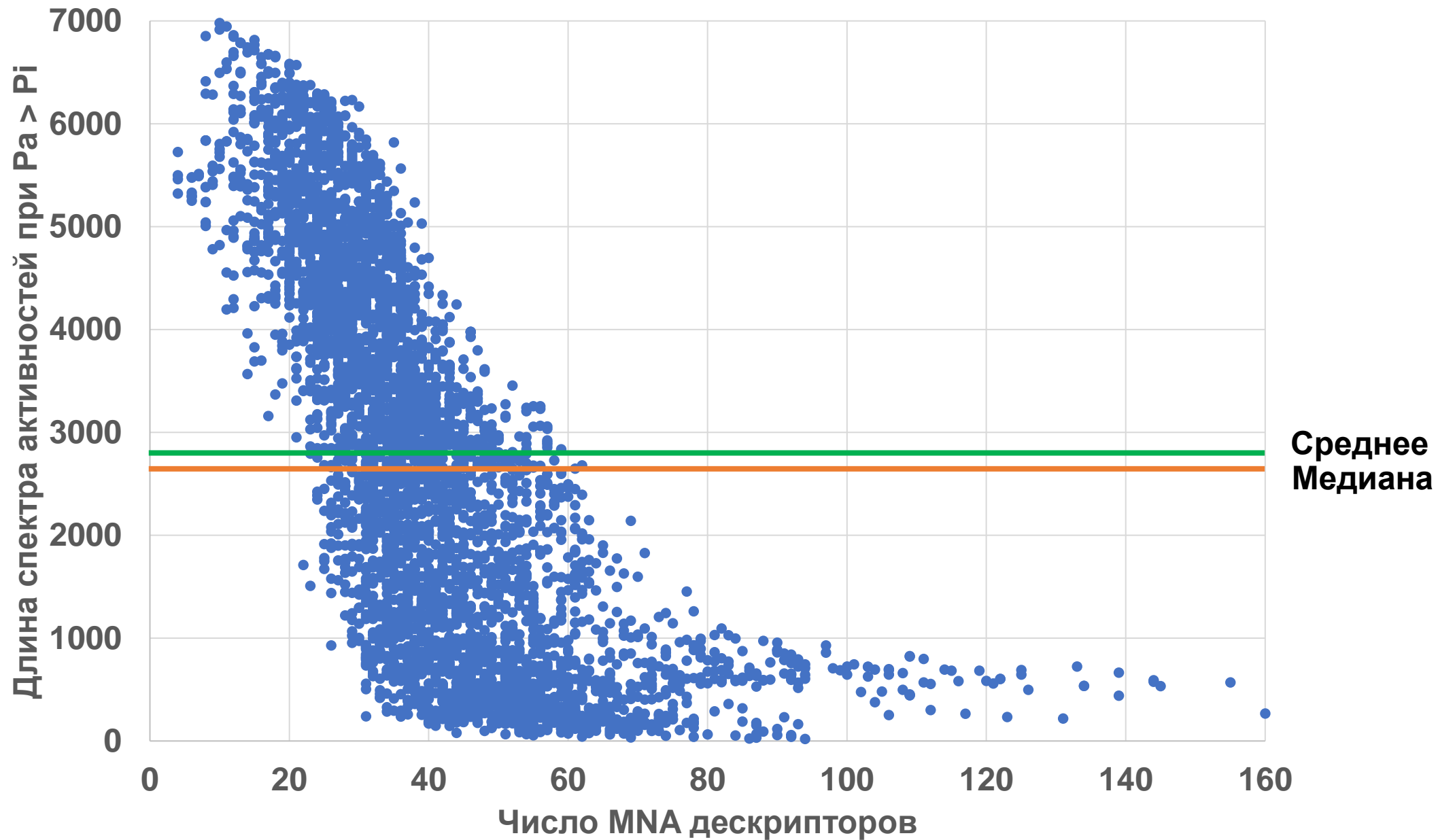


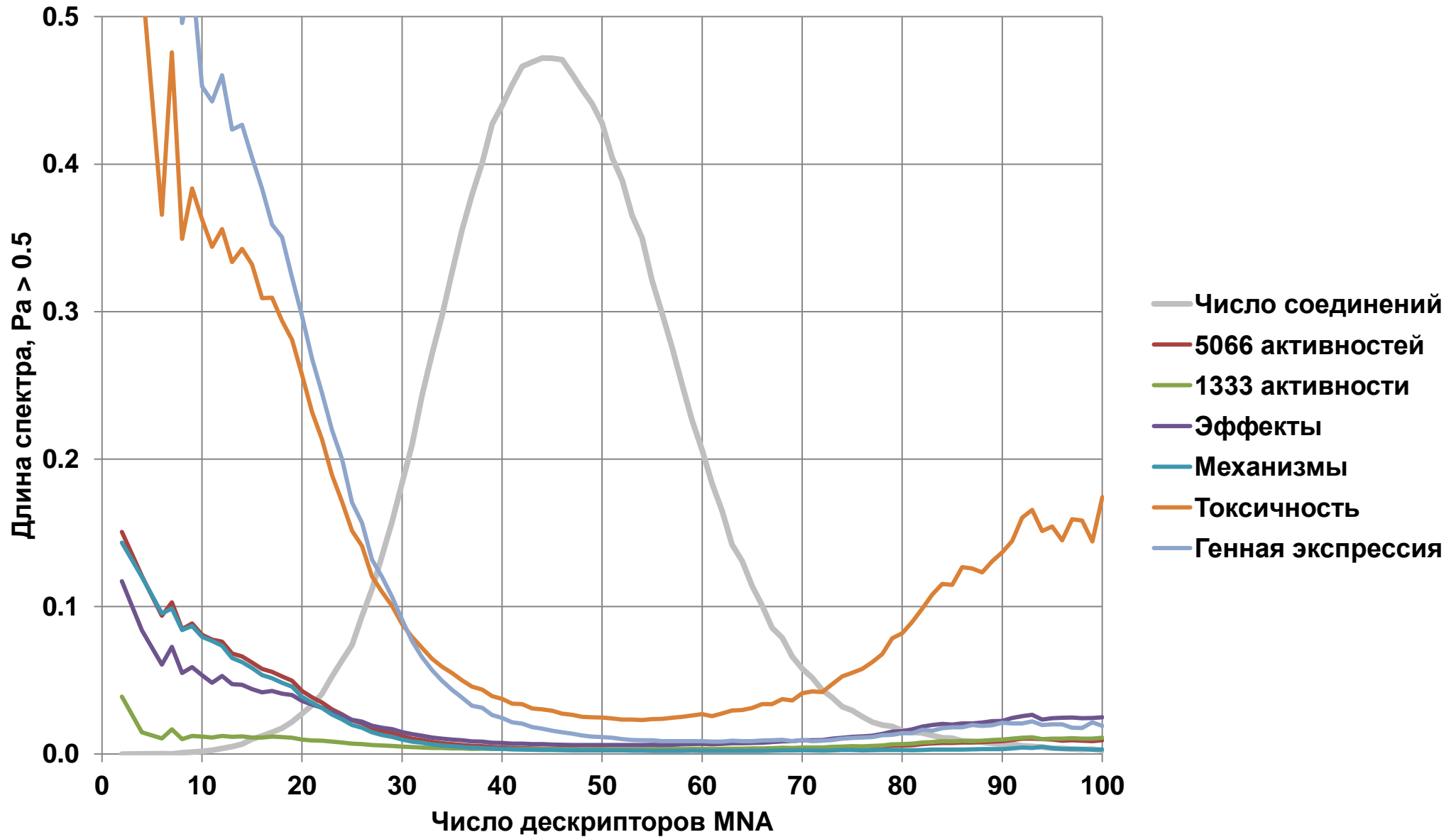
Progesterone 33 MNA 656 AK

World Wide Approved Drugs



World Wide Approved Drugs



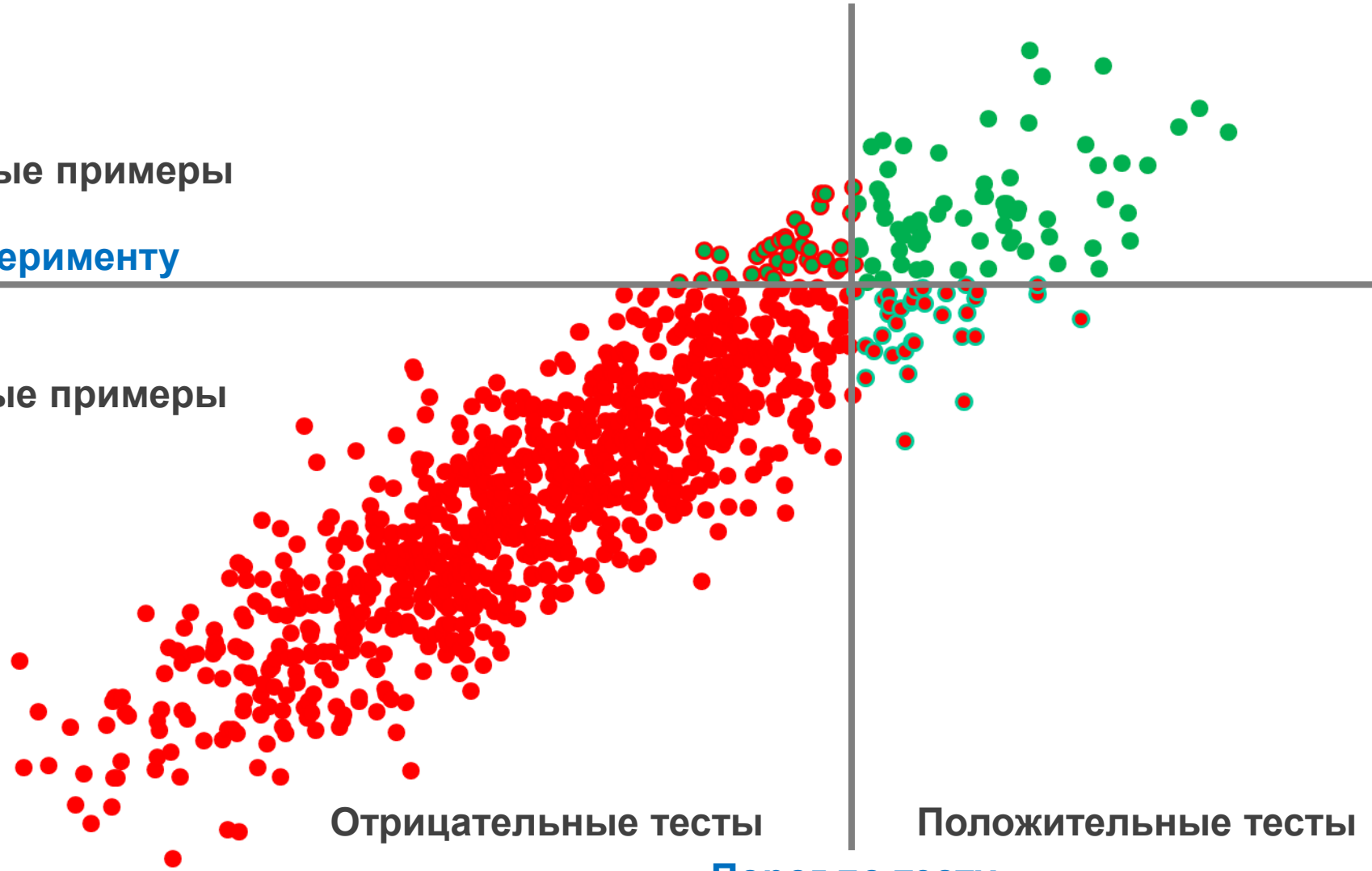


Задача классификации

Положительные примеры

Порог по эксперименту

Отрицательные примеры

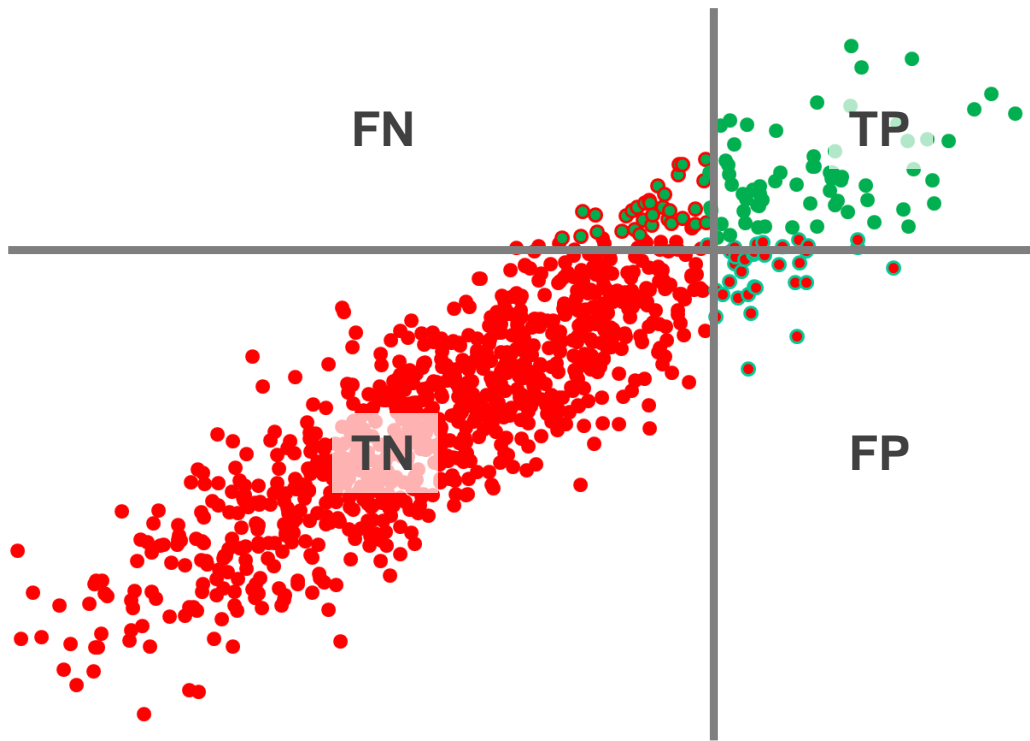


Отрицательные тесты

Положительные тесты

Порог по тесту

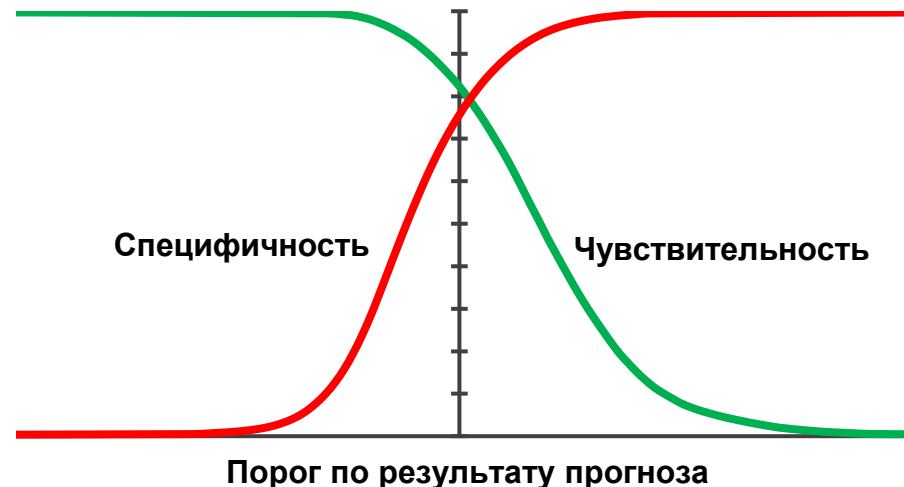
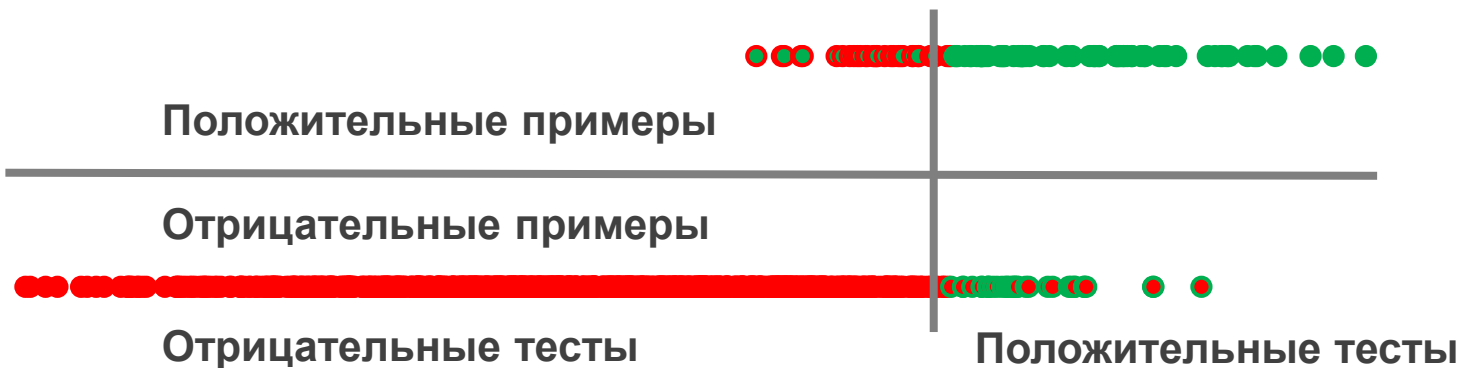
Критерии качества классификации



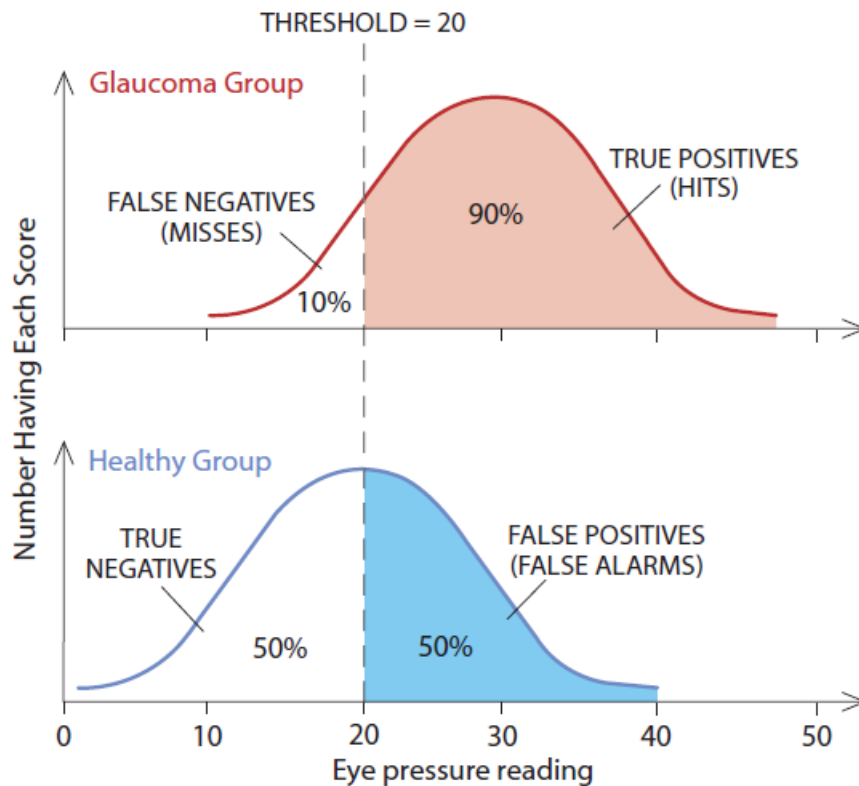
TP – True Positives
TN – True Negatives
FP – False Positives
FN – False Negatives

Sensitivity (Recall)	$= TP / (TP + FN) = 1 - Error_1$
Specificity (Selectivity)	$= TN / (TN + FP) = 1 - Error_2$
Balanced Accuracy	$= (Sensitivity + Specificity) / 2$
Accuracy (Concordance)	$= (TP + TN) / N$
Precision	$= TP / (TP + FP)$
False Negative Rate	$= FN / (TP + FN) = Error_1$
False Positive Rate	$= FP / (TN + FP) = Error_2$
Positive Likelihood	$= SENS / (1 - SPEC)$
Negative Likelihood	$= (1 - SENS) / SPEC$
...	...

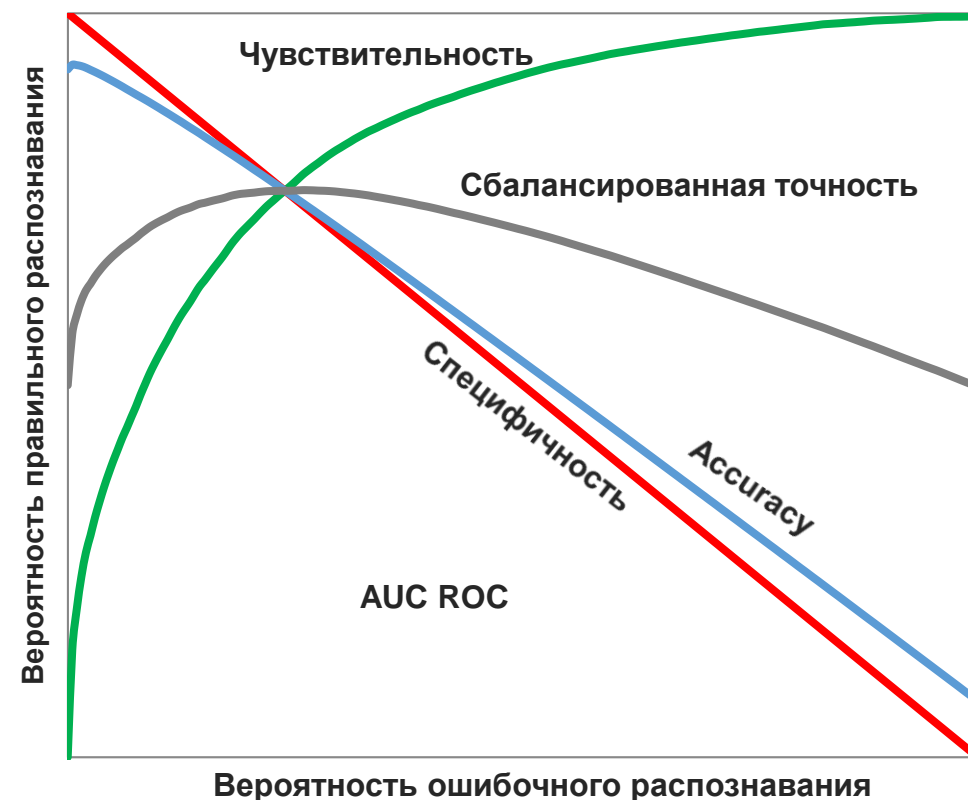
Критерии качества классификации



Положительные примеры

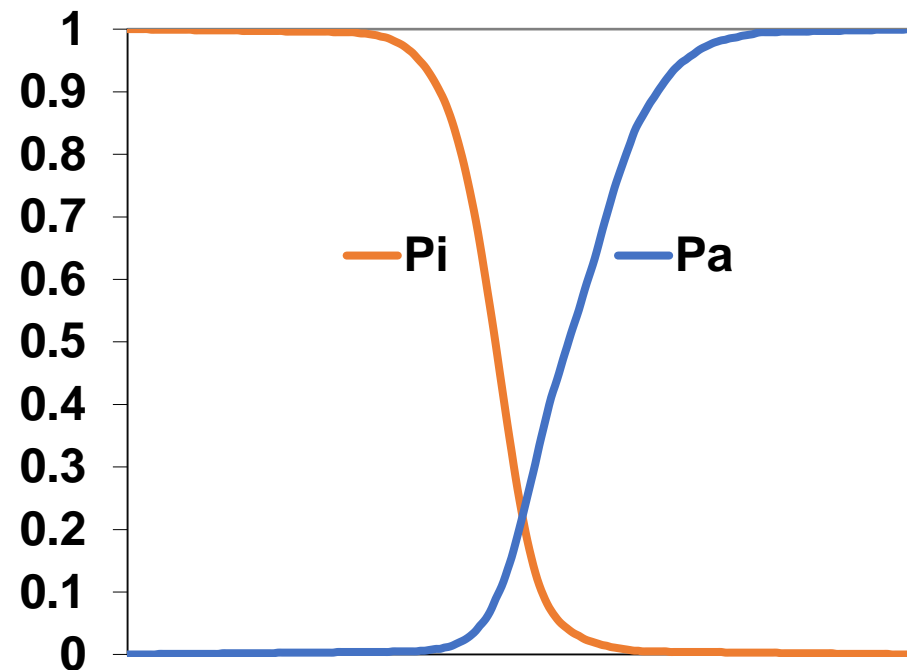
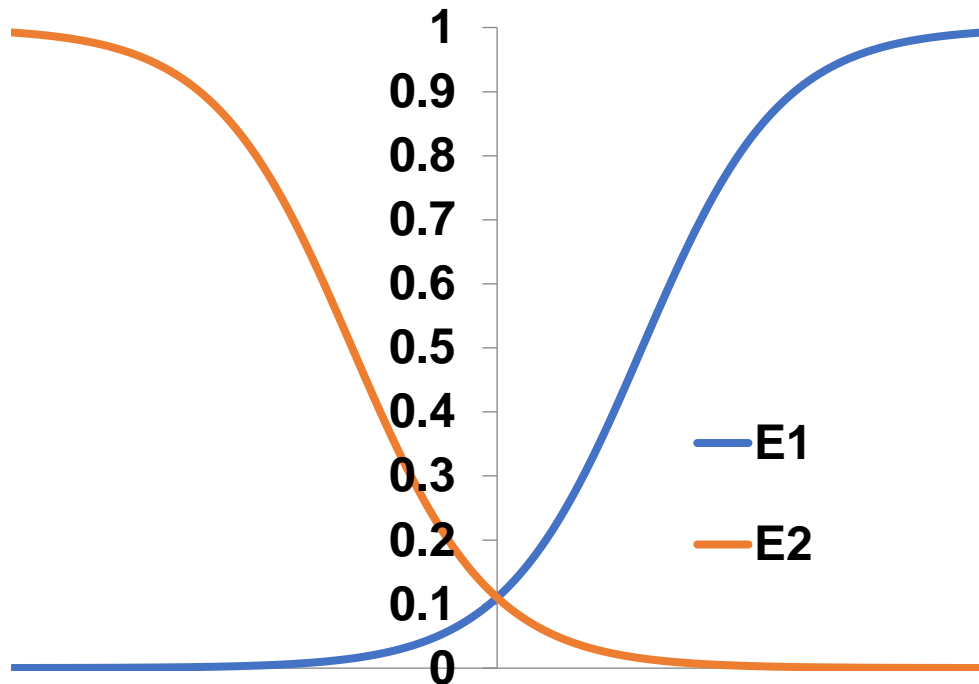


Отрицательные примеры



Аппроксимация зависимости $E_1(\theta)$ и $E_2(\theta)$ от порога отсечения θ :

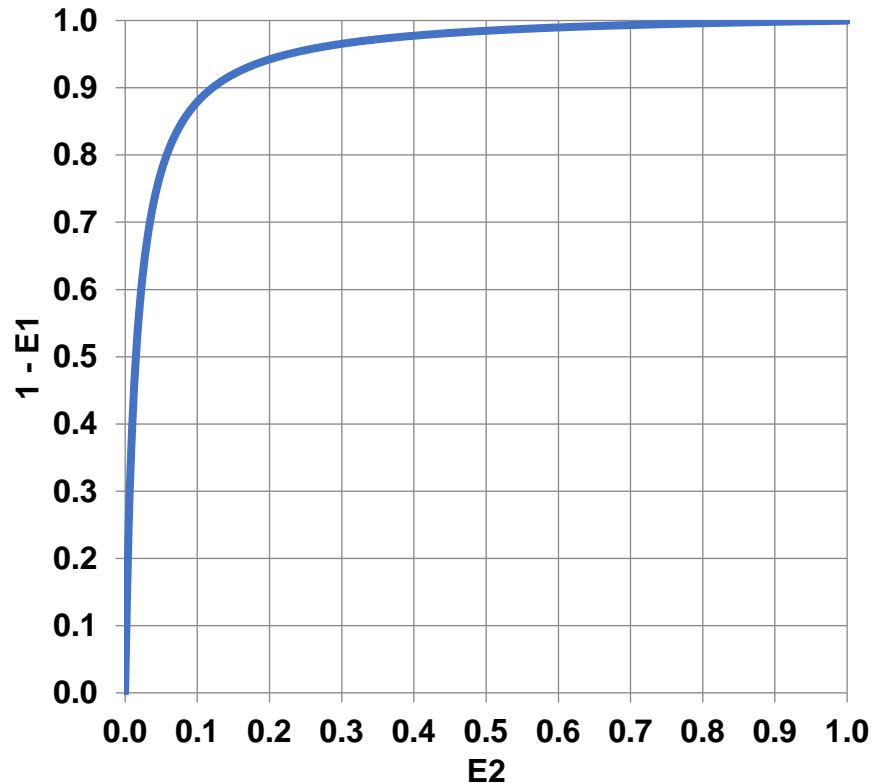
$$E_1(\theta) = \frac{1}{1+e^{a-b\theta}} \quad \text{и} \quad E_2(\theta) = \frac{1}{1+e^{a+b\theta}}; \quad E_1 = \frac{1-E_2}{1+CE_2}, \quad E_2 = \frac{1-E_1}{1+CE_1}, \quad C = e^{2a} - 1$$



Слева – зависимости $E_1(\theta)$ и $E_2(\theta)$ от порога отсечения θ для $a = 2.09$ и $b = 7$.

Справа – оценки вероятностей P_a и P_i как функции В-статистики. Кривые построены для активности «Antihypertensive» по данным SAR Base PASS.

Как выбрать лучший порог?



Аппроксимация зависимости вероятности правильного распознавания $1 - E_1$ от вероятности ошибки второго рода E_2 при $a = 2.09$, $C = 64.37$, $IAP = 0.95$.

27 of 561 Possible Pharmacological Effects at $P_a > 0.500$

0.882	0.008	Immunostimulant (HIV)
0.853	0.004	Alopecia treatment
0.789	0.015	Antiseborrheic
0.714	0.012	Kidney function stimulant
0.688	0.013	Antineoplastic (basal cell carcinoma)
0.674	0.004	Leukopoiesis inhibitor
0.682	0.023	Fibrinolytic
0.694	0.050	Inotropic
0.654	0.026	Antiviral (Arbovirus)
0.627	0.015	Immunomodulator (HIV)
0.619	0.018	Mucositis treatment
0.603	0.008	Antipyretic
0.611	0.016	Gastroprotector
0.592	0.008	Erythropoiesis stimulant
0.594	0.012	Leukopoiesis stimulant
0.557	0.005	Allergic conjunctivitis treatment
0.565	0.017	Vasodilator, renal
0.547	0.004	Antiuremic
0.528	0.017	Hepatitis alcoholic treatment
0.535	0.025	Respiratory analeptic
0.531	0.026	Skin diseases treatment
0.519	0.017	Natural killer cell stimulant
0.518	0.018	Cardiovascular analeptic
0.542	0.052	Hepatitis treatment
0.515	0.026	Antiinflammatory, intestinal
0.515	0.031	Potassium sparing diuretic
0.560	0.107	Antineoplastic (Kaposi's sarcoma)

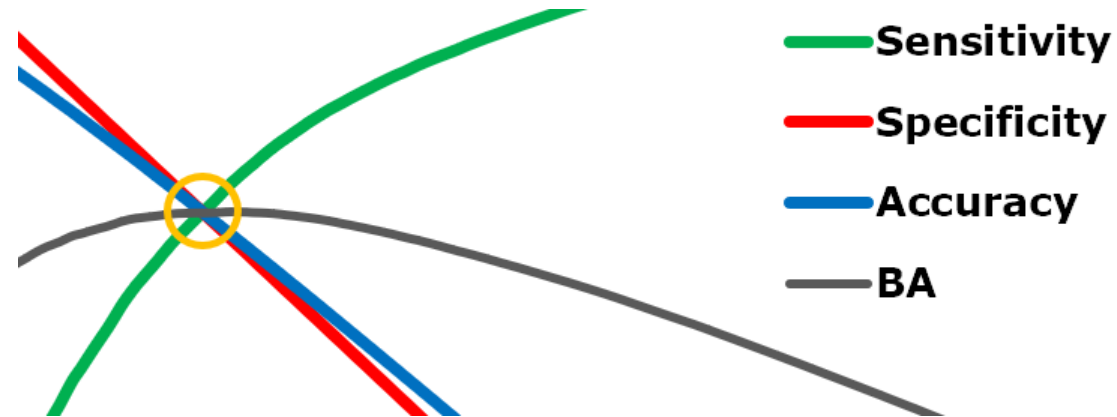
Как выбрать лучший порог?

При отсутствии знаний априорной вероятности найти соединение с заданной активностью и цены правильных и ошибочных решений можно использовать принцип минимакса:

минимальные потери при максимально плохих условиях

которому соответствуют равенства

“Sensitivity = Specificity = Accuracy” или $E_1 = E_2$ или критерий $P_a > P_i$



Есть две основные практические задачи:

Задача анализа спектра прогнозируемой биологической активности

– это задача выбора наиболее подходящих видов активности из множества возможных.

Проблема виртуального скрининга

– это проблема выбора наиболее подходящих соединений из множества возможных.

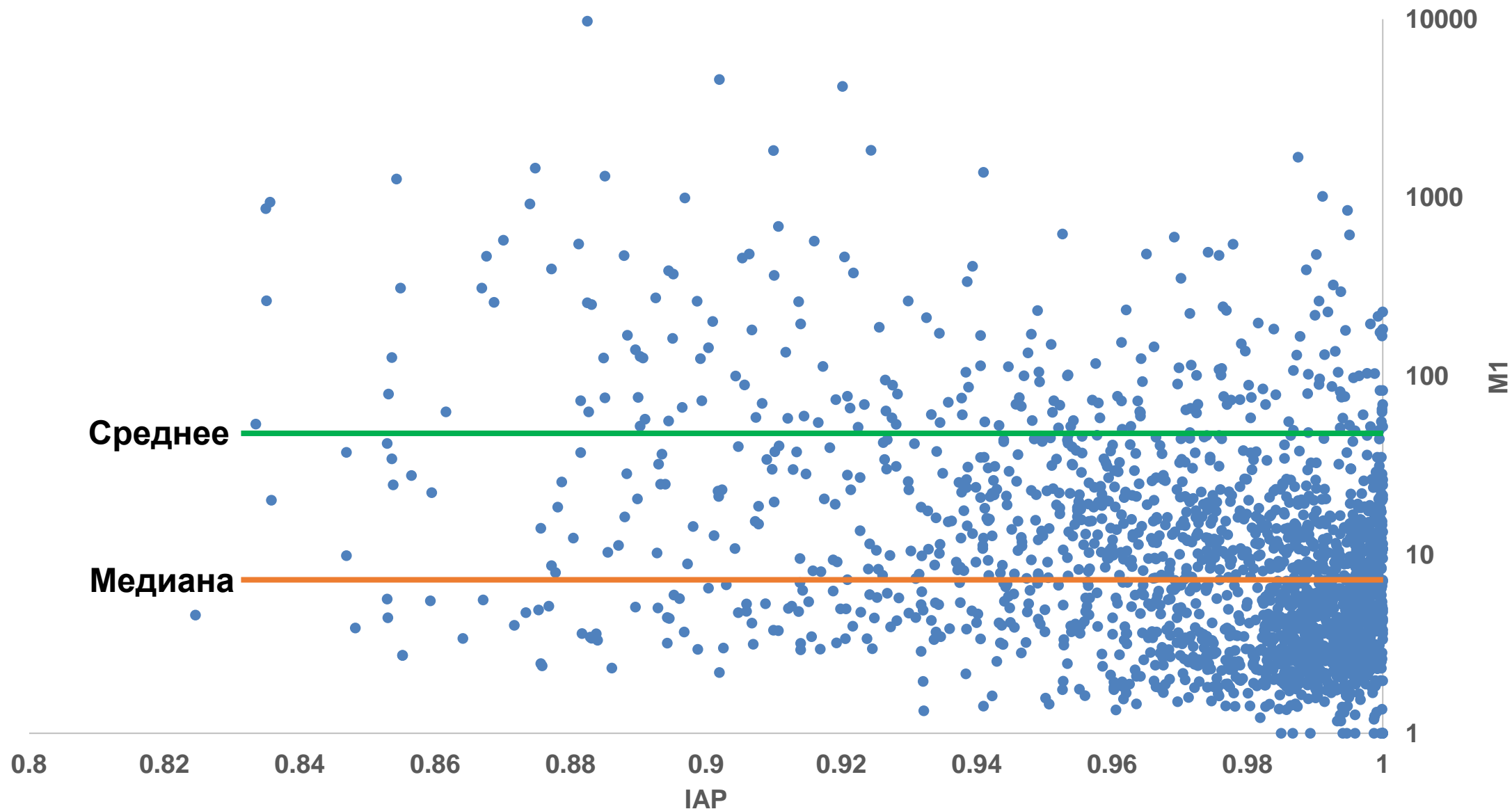
Оба случая представляют собой задачу выбора наилучшего подмножества из M , $M < N$ множества N оценок активности – например оценок P_a .

Число M – это еще одно представление порогового значения:

$$M = TP + FP = N_1 \cdot (1 - E_1) + (N - N_1) \cdot E_2 = TP \cdot \left(1 + \frac{(N - N_1) \cdot E_2}{N_1 \cdot (1 - E_1)} \right) \equiv \frac{TP}{Precision}$$

Достижению первого успеха соответствует $TP = 1$:

$$M_1 = 1 + FP|_{TP=1} = 1 + (N - N_1) \cdot E_2|_{N_1 \cdot (1 - E_1) = 1}$$



Оценки значений M_1 и IAP в SAR Base PASS 2020 для 1945 отобранных видов активности

Для 1945 отобранных видов активности, которые включают основные биологические активности, наиболее важные для медицинских химиков, фармакологов и врачей,
медианное значение M_1 равно 7, среднее значение M_1 равно 43

Это означает, что из прогнозируемого спектра биологической активности необходимо проверить около 7-10 из наиболее вероятных активностей, чтобы иметь хорошие шансы найти одну из них в биологическом тестировании.

Это также означает, что при виртуальном скрининге соединений необходимо рассмотреть около 7-10 или более “лучших” соединений как возможно активных, чтобы иметь хорошие шансы найти одно активное соединение при тестировании.

При расширении списка активностей более редкими и прогнозируемыми с низкой точностью количество ложноположительных прогнозов возрастает более чем квадратично.

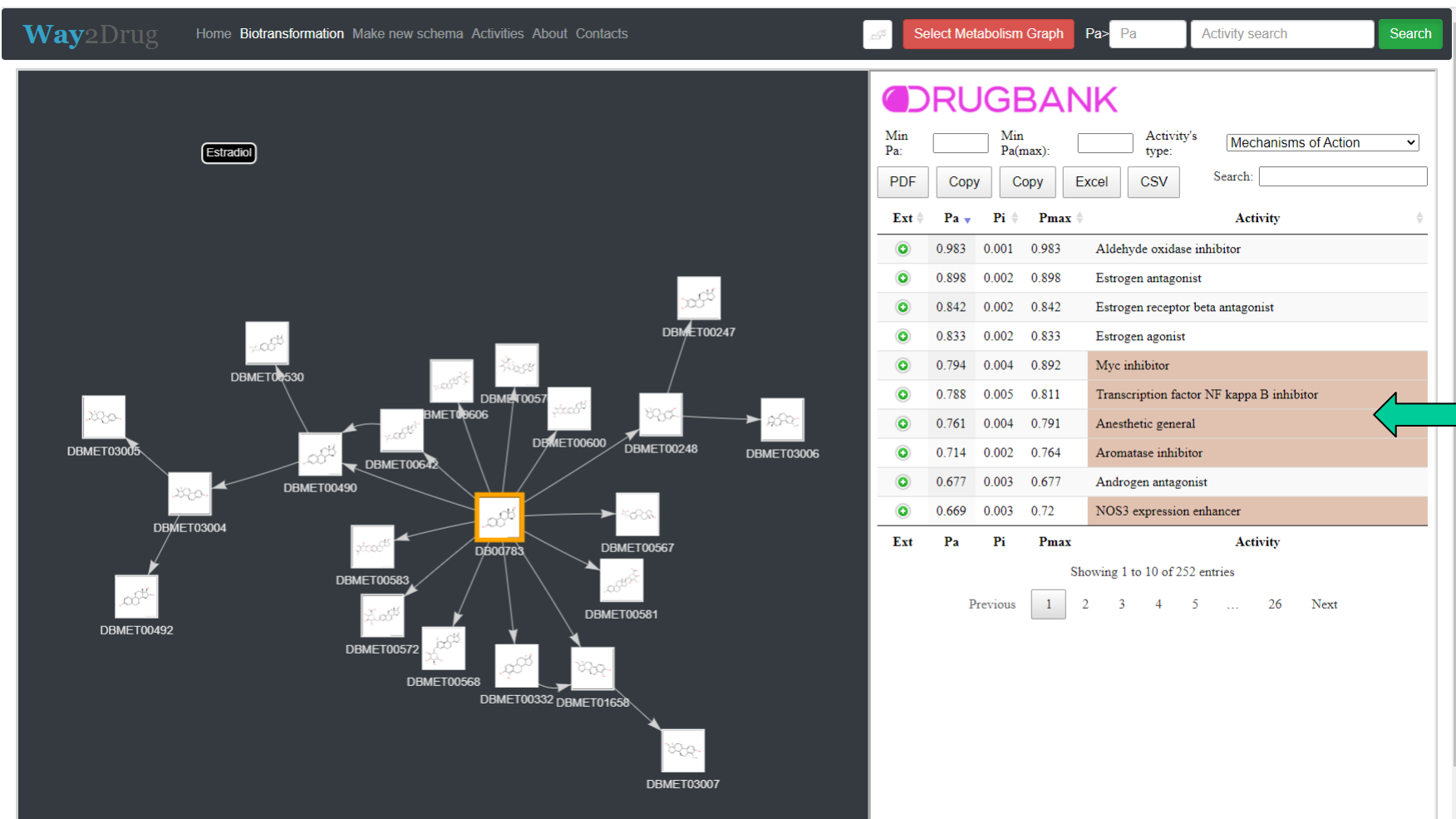
**Лучшая стратегия – проверять экспериментально активности/соединения
в порядке убывания предсказанной активности, шаг за шагом**

Последовательная стратегия: анализ прогноза спектра активности
Paracetamol, фармакологические эффекты, порог $P_a > P_i$

Номер по порядку среди активностей			P_a	P_i	Активность
18 известных	364 отобранных	561 прогнозируемых			
1	1	12	0.603	0.008	Antipyretic
2	7	42	0.426	0.018	Anesthetic general
3	17	75	0.333	0.066	Immunostimulant
4	35	121	0.305	0.120	Antibacterial
5	70	181	0.169	0.055	Anesthetic
6	86	201	0.182	0.087	Non-steroidal antiinflammatory agent
7	104	233	0.187	0.125	Antiosteoporotic
8	120	253	0.190	0.146	Bone diseases treatment
9	149	298	0.068	0.058	Antibiotic
10			0.098	0.103	Bone formation stimulant
11			0.165	0.244	Antiviral
12			0.147	0.247	Immunosuppressant
13			0.140	0.345	Antiarthritic
14			0.155	0.393	Antiinflammatory
15			0.088	0.449	Analgesic
16			0.091	0.607	Antineoplastic
17			0.036	0.654	Analgesic, non-opioid
18			0.019	0.750	Signal transduction pathways inhibitor

Последовательная стратегия: анализ прогноза спектра активности с учетом биотрансформации

Estradiol, 1289 известных активностей, 20 известных метаболитов

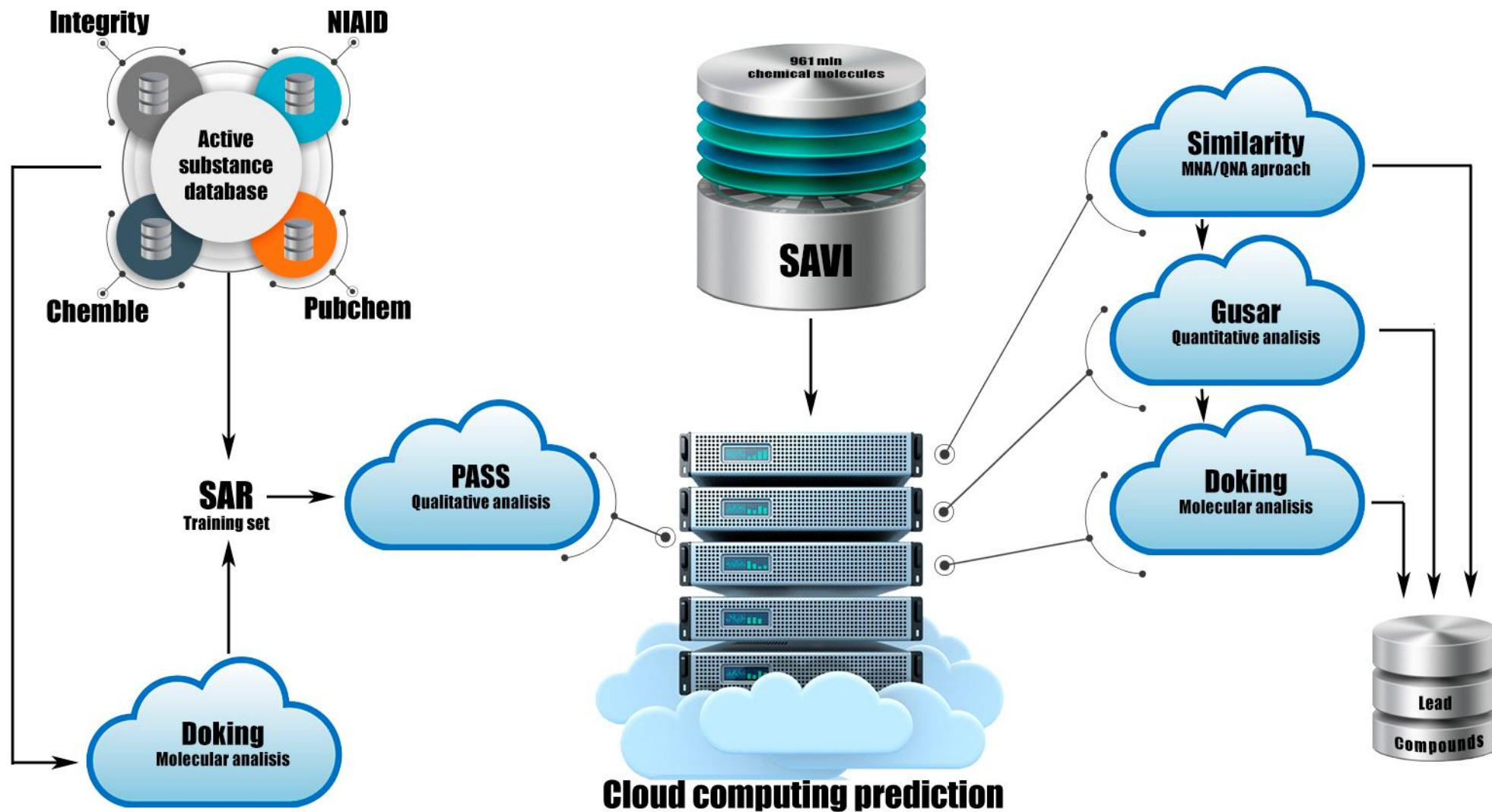


Pa метаболита больше
Pa исходного соединения

<http://way2drug.com/metapass/>

Rudik A.V. et al. Molecular Informatics, 2020; <https://doi.org/10.1002/minf.202000231>

Последовательная стратегия: скрининг химических Big Data



Прогноз PASS во многих случаях позволил отобрать для синтеза наиболее перспективные соединения и определить приоритетные направления тестирования их биологической активности

Bioorganic Chemistry 96 (2020) 103610
Contents lists available at ScienceDirect
journal homepage: www.elsevier.com/locate/bioorg

Future VIROLOGY
Research Article
For reprint orders, please contact: reprints@futuremedicine.com
Do fever-relieving medicines have anti-COVID activity: an *in silico* insight
Fahad Hassan Shah¹, Kyeong Ho Lim² & Song Ja Kim^{*,1}
¹Department of Biological Sciences, College of Natural Sciences, Kongju National University, Gongju 32588, Republic of Korea
²Department of Construction and Environmental Engineering, College of Engineering, Kongju National University, Cheonan 31060, Republic of Korea

Heliyon xxx (xxxx) xxx
Contents lists available at ScienceDirect
journal homepage: www.cell.com/heliyon

CellPress
Research article
Synthesis, characterization, *in silico*, and *in vitro* biological screening of coordination compounds with 1,2,4-triazine based biocompatible ligands and selected 3d metal ions

Advances in Traditional Medicine
<https://doi.org/10.1007/s13596-020-00530-x>
RESEARCH ARTICLE
Network pharmacology of *Withania somnifera* against stress associated neurodegenerative diseases
Taaza Duyu¹ · Pukar Khanal¹ · Ya...
Received: 16 March 2020 / Accepted: 17 Nov...
© Institute of Korean Medicine, Kyung Hee U...

Springer

ACS OMEGA
<http://pubs.acs.org/journal/acsodf>
Article
New Reaction Products of Acetylacetone with Semicarbazide Derivatives
Vera S. Glukhacheva,^{*} Sergey G. Il'yasov, Igor V. Kazantsev, Elena O. Shestakova, Dmitri S. Il'yasov, Iliia V. Eltsov, Andrey A. Nefedov, and Yuri V. Gatilov
Cite This: <https://doi.org/10.1021/acsomega.1c00518>
Read Online

pharmaceuticals
Article
Antidepressant, anxiolytic, antipyretic, and thrombolytic profiling of methanol extract of the aerial part of *Piper nigrum*: In vivo, in vitro, and in silico approaches
© 2020. Published by The Company of Biologists Ltd | Disease Models & Mechanisms (2020) 13, dmm044040. doi:10.1242/dmm.044040

antibiotics
Article
Antimicrobial Activity of Nitrogen-Containing

MDPI

>1000 публикаций

Спасибо за внимание!



Российский
научный
фонд

Работа выполнена при поддержке Российского научного фонда грант No. 19-15-00396