

# Вводная лекция

# Вопросы:

- ◆ Какую научную работу написать?
- ◆ А что будет на кружке?

Наука



- Секция устных докладов
- Секция мультимедийных постеров

Реферативный доклад

Drug discovery pipeline

Клинические

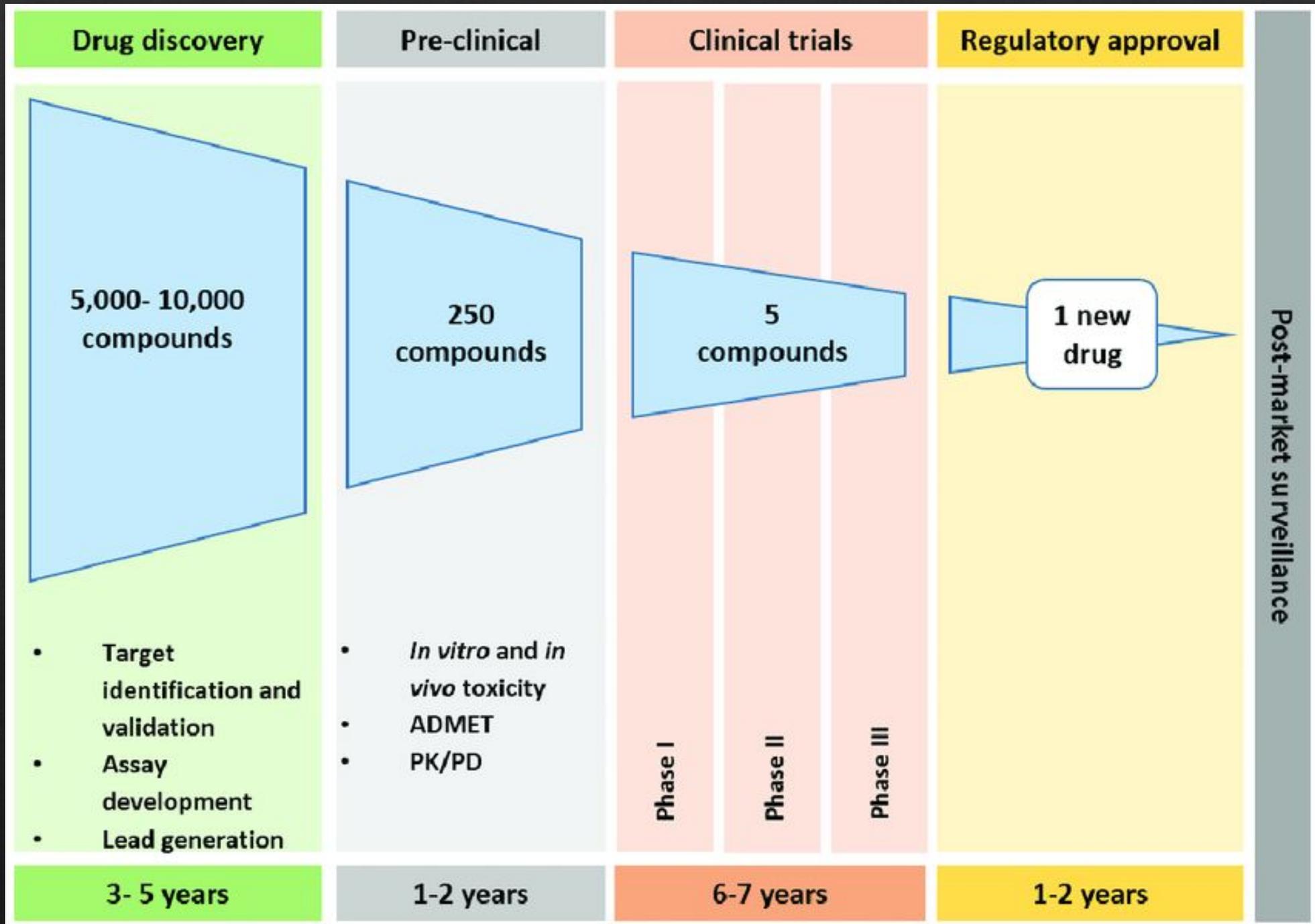
Теоретические











5,000- 10,000 compounds

- Target identification and validation
- Assay development
- Lead generation

3- 5 years

250 compounds

- *In vitro* and *in vivo* toxicity
- ADMET
- PK/PD

1-2 years

5 compounds

Phase I

Phase II

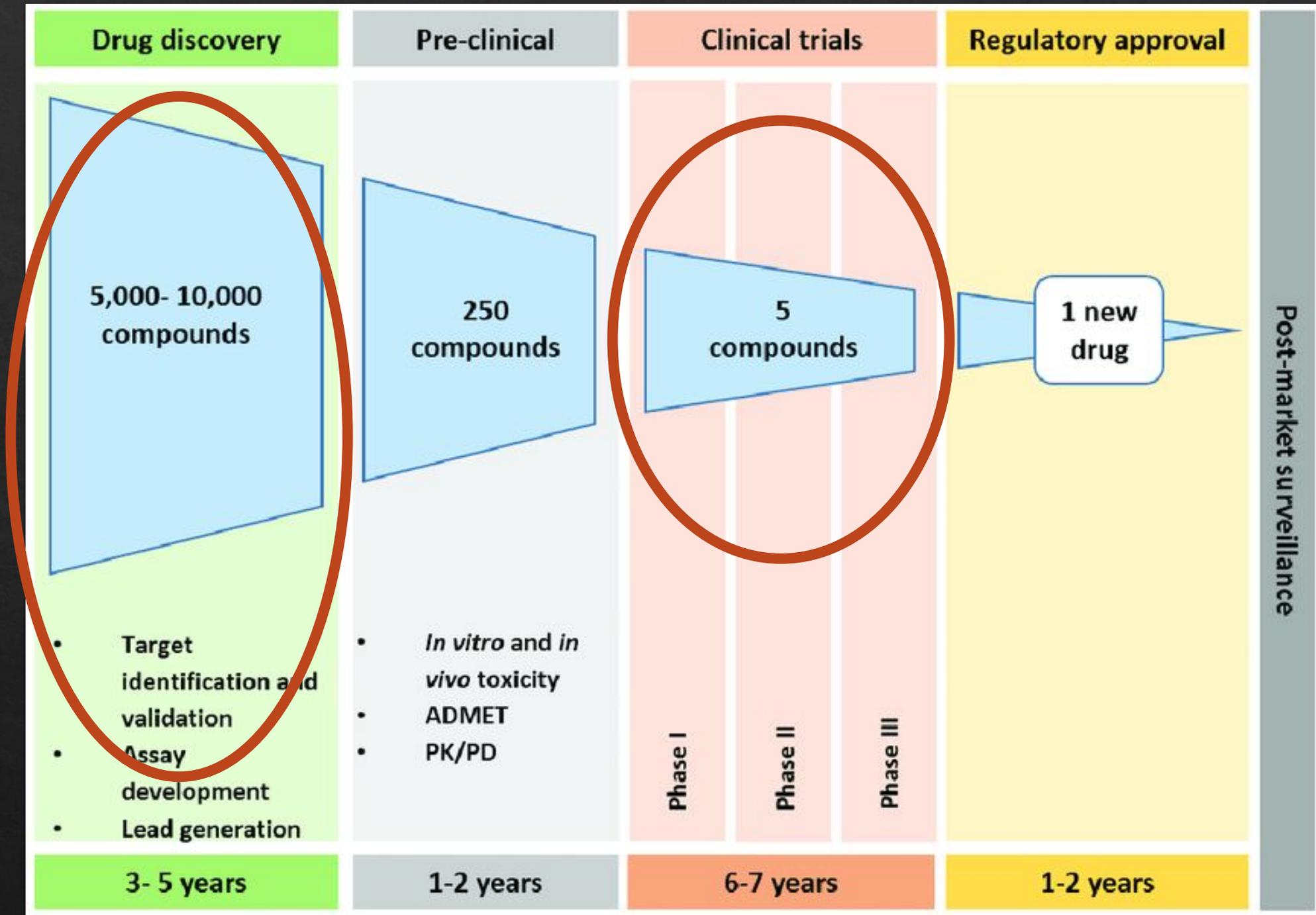
Phase III

6-7 years

1 new drug

1-2 years

Post-market surveillance



5,000- 10,000 compounds

- Target identification and validation
- Assay development
- Lead generation

3- 5 years

250 compounds

- *In vitro* and *in vivo* toxicity
- ADMET
- PK/PD

1-2 years

5 compounds

Phase I

Phase II

Phase III

6-7 years

1 new drug

1-2 years

Post-market surveillance

# Клинические работы

# Классическая работа

Идея: Сравнение методов лечения.

1. Препарат vs. Другой препарат,
2. Препарат vs. Комбинация препаратов,
3. Препарат vs. Немедикаментозный метод лечения.



## Этапы:

1. Сбор данных (Данные историй болезни, опрос пациентов)
2. Обработка данных: разбиваем пациентов на две группы, считаем описательные статистики, проверяем на нормальность.
3. Анализ данных: считаем средние значения сравниваемых показателей, подтверждаем наличие/отсутствие различий (t-критерий Стьюдента, U-критерий Манна — Уитни)
4. Делаем выводы.

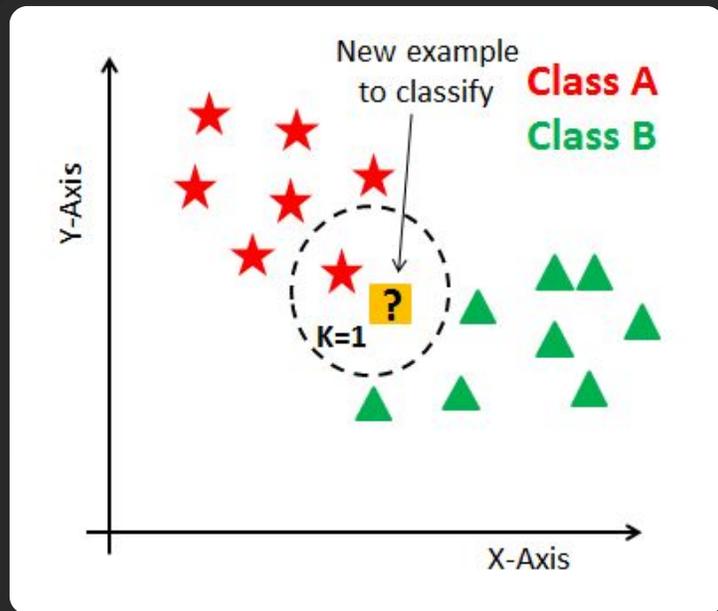


VS

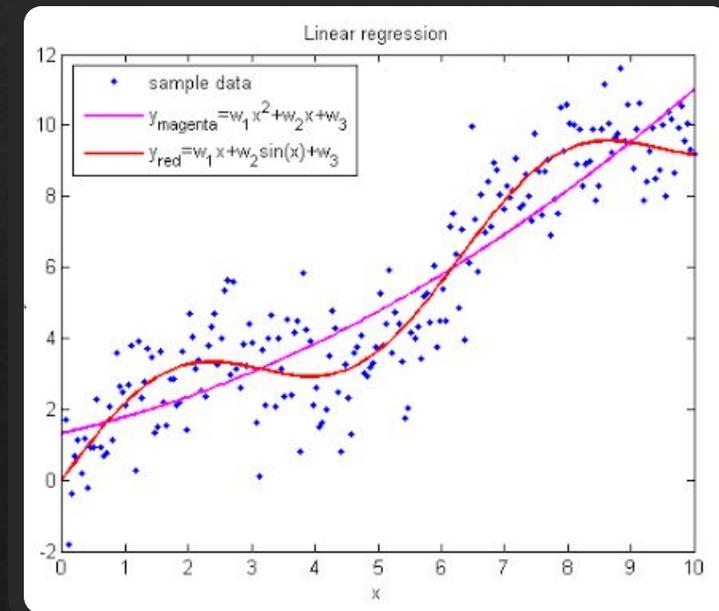


# Модели машинного обучения

## Классификация



## Регрессия



# Независимые переменные

V - 1	V - 2	V - 3	V - 4	V - 5	V - 6	V - 7	V - 8	V - 9	V - 10
value									
value									
value									
value									
value									
value									
value									

value									
-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

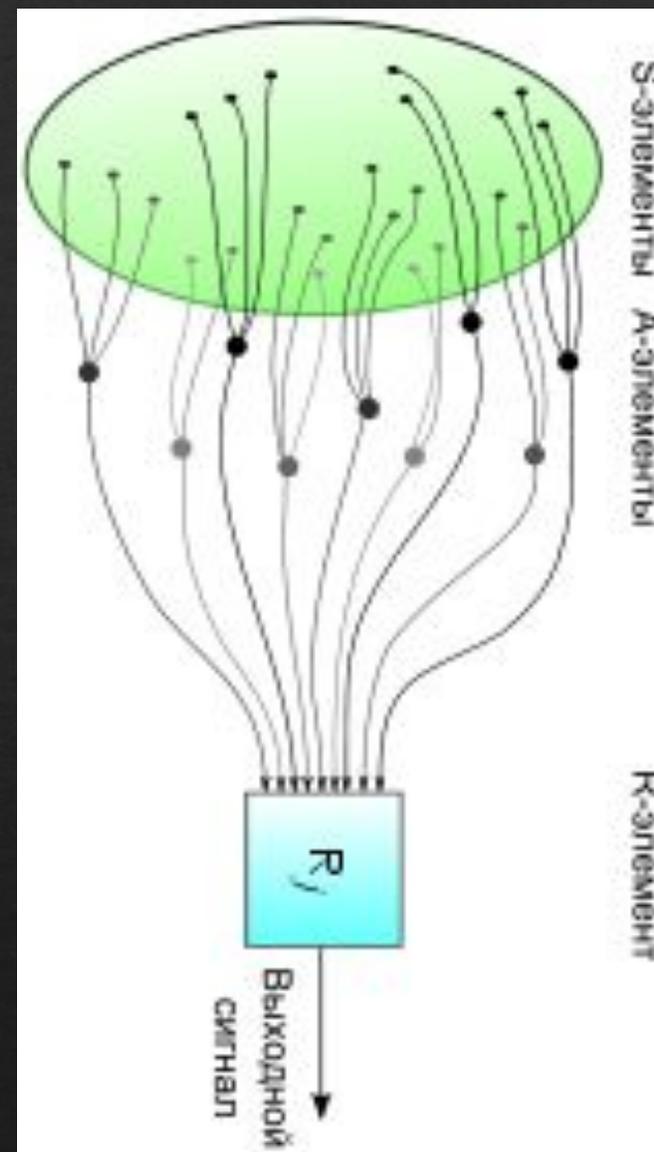
# Зависимая

Num	Class
1.003	cat
2.008	dog
7.256	dog
8.240	cat
3.001	cat
5.443	cat
2.754	dog

?	?
---	---

## Этапы:

1. Сбор данных (Данные историй болезни, опрос пациентов)
2. Обработка данных: разбиваем данные на две части — тренировочную и тестовую.
3. На тренировочных данных обучаем модель.
4. На тестовых данных предсказываем значение зависимой переменной.
5. Оцениваем эффективность модели, делаем выводы.

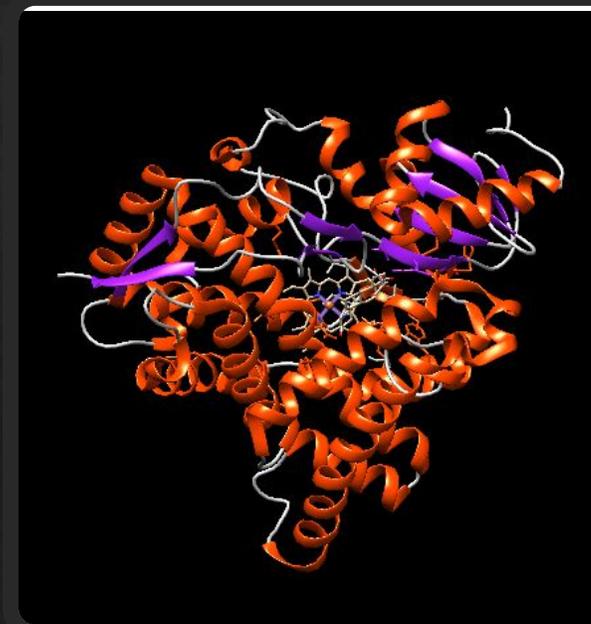
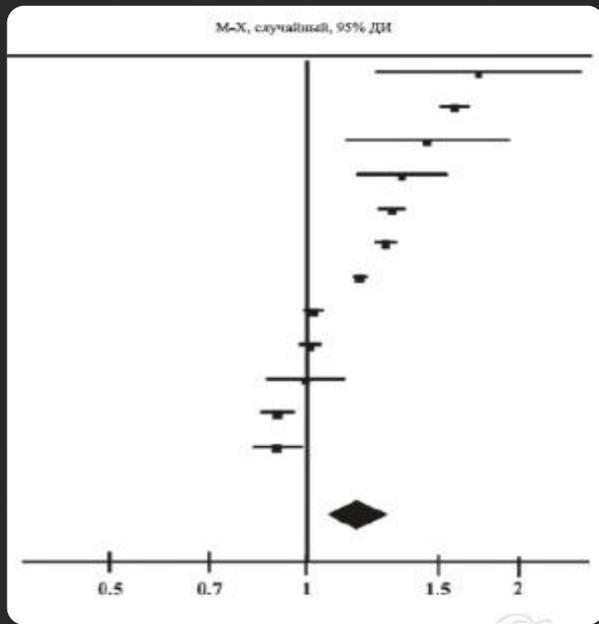


# Теоретические работы

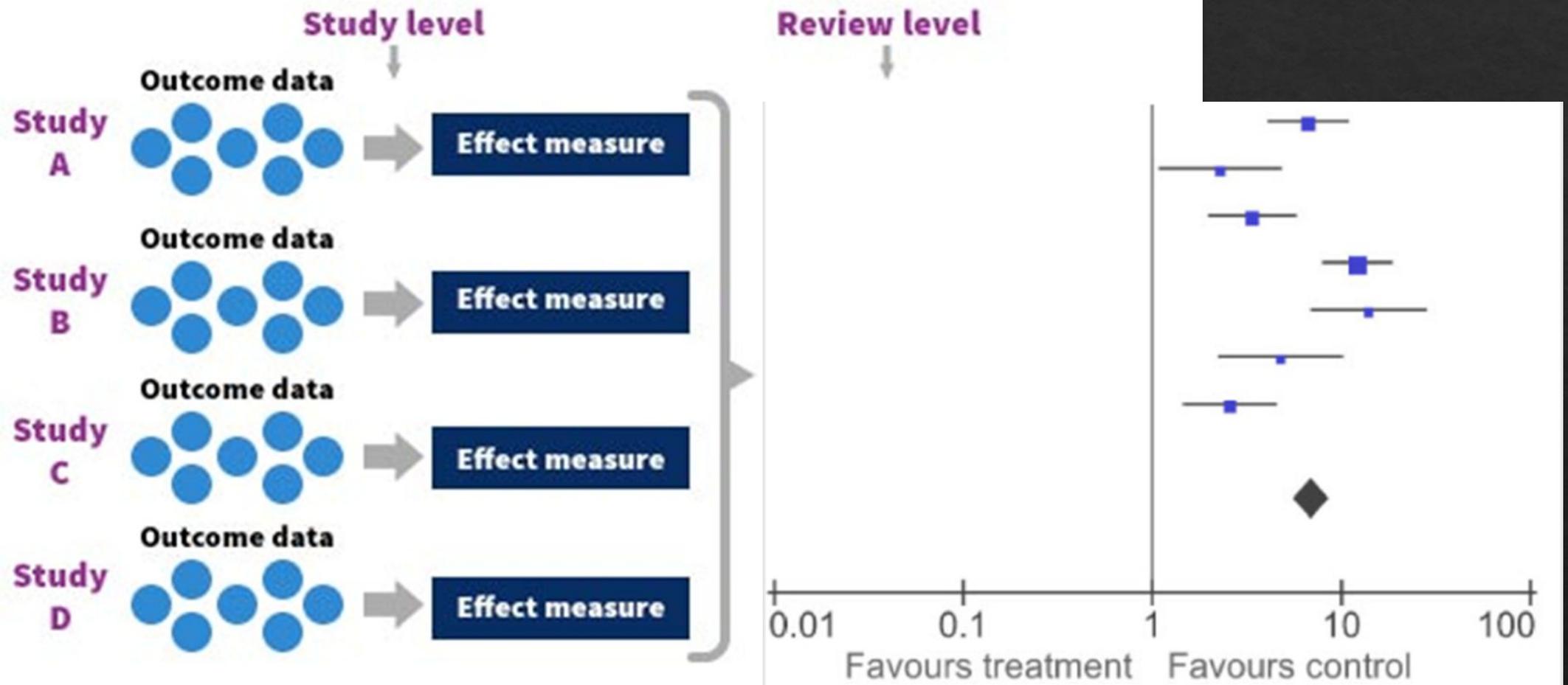
# Теоретические работы

Систематический обзор с  
выполнением мета-  
анализа

In silico drug design



# Мета-анализ



# Систематический обзор с выполнением мета-анализа

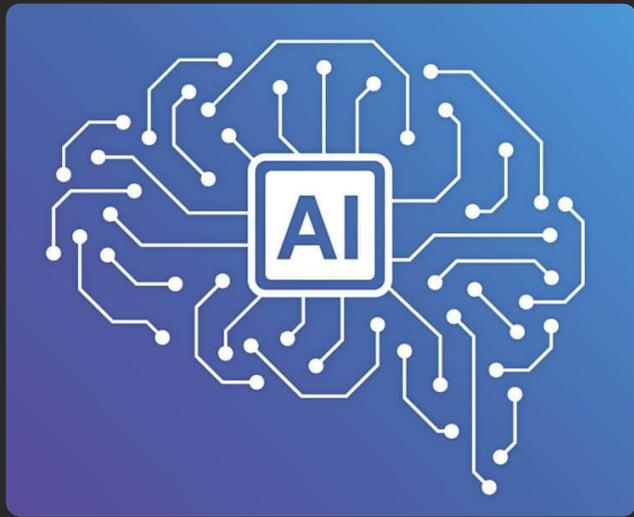
## Этапы:

1. Поиск и отбор статей по проблеме обзора
2. Отбор наиболее подходящих статей для включения в обзор
3. Выполнение мета-анализа
4. Оценка риска предвзятости (ROB 2.0)
5. Оценка качества доказательств (GRADE)
6. Выводы, оформление обзора

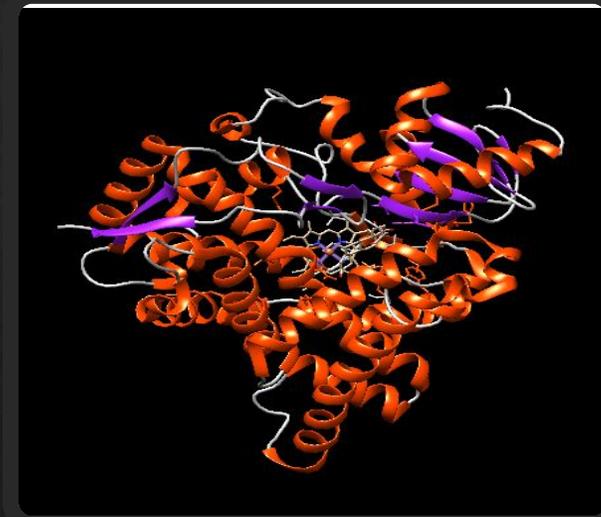


# In silico drug design

Модели машинного  
обучения



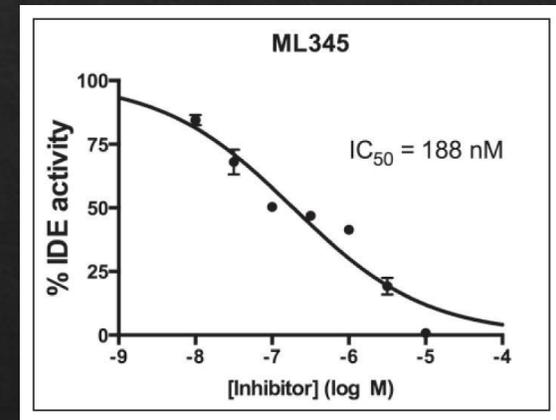
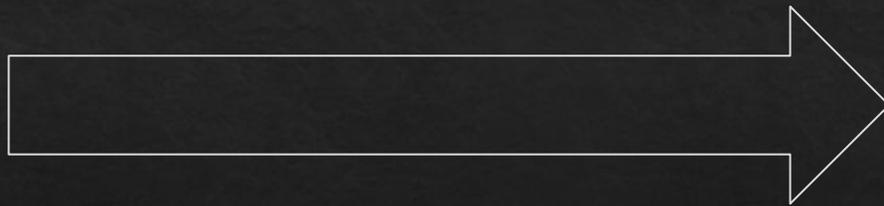
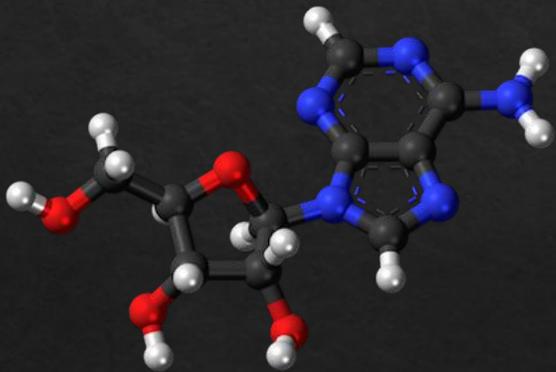
Проект кружка



# Модели машинного обучения

$$\text{QSAR} = \text{Q} + \text{SAR}$$

SAR = structure–activity relationship



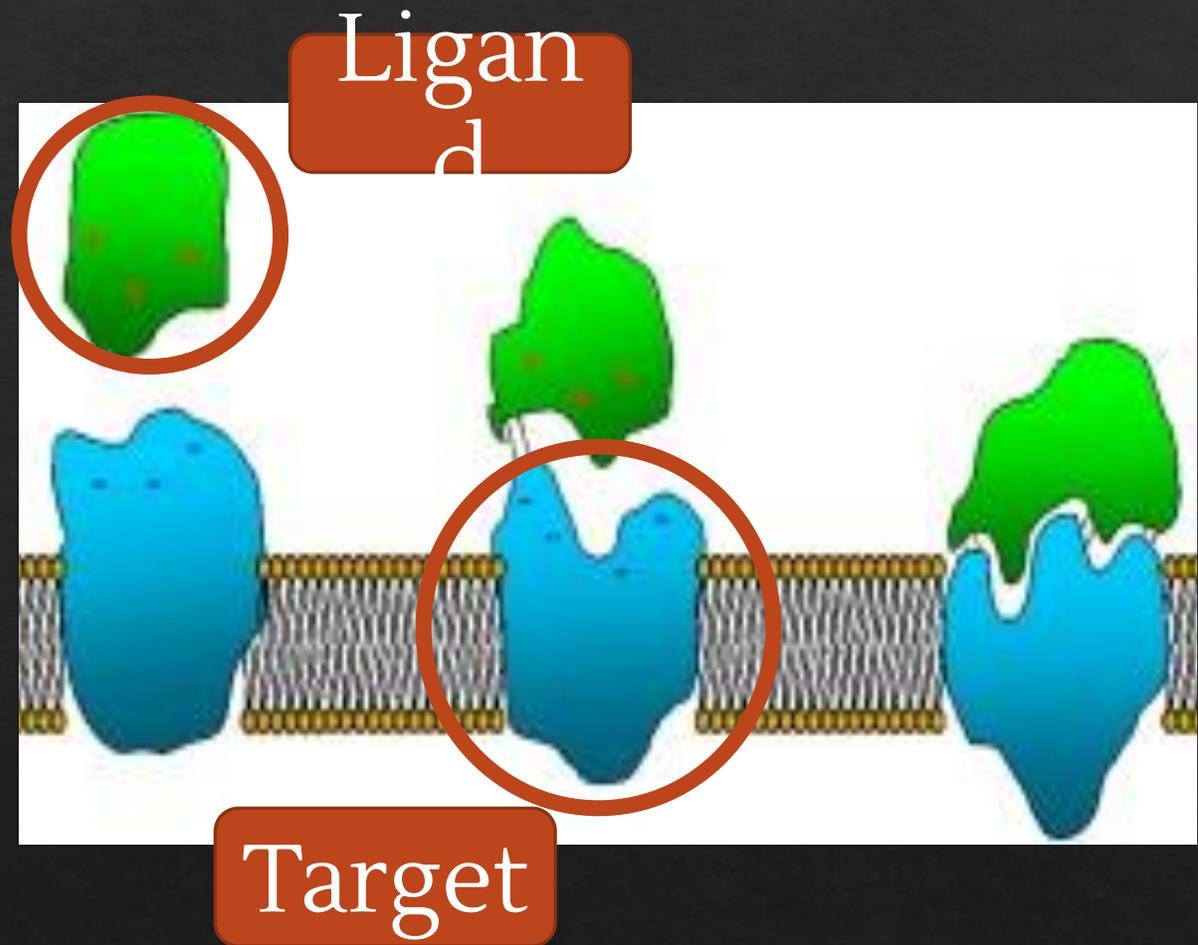
Q =

Quantitative

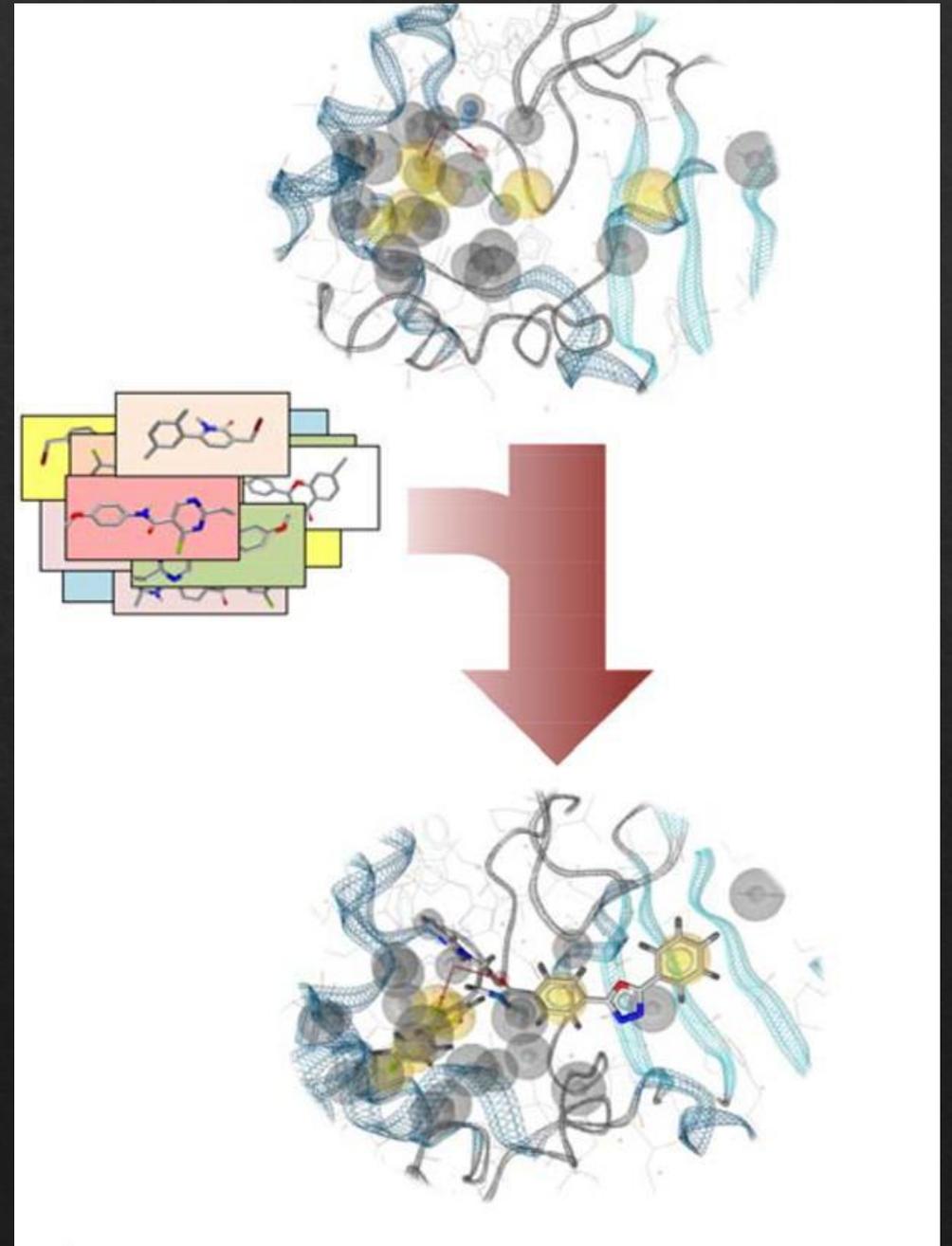
Кружок

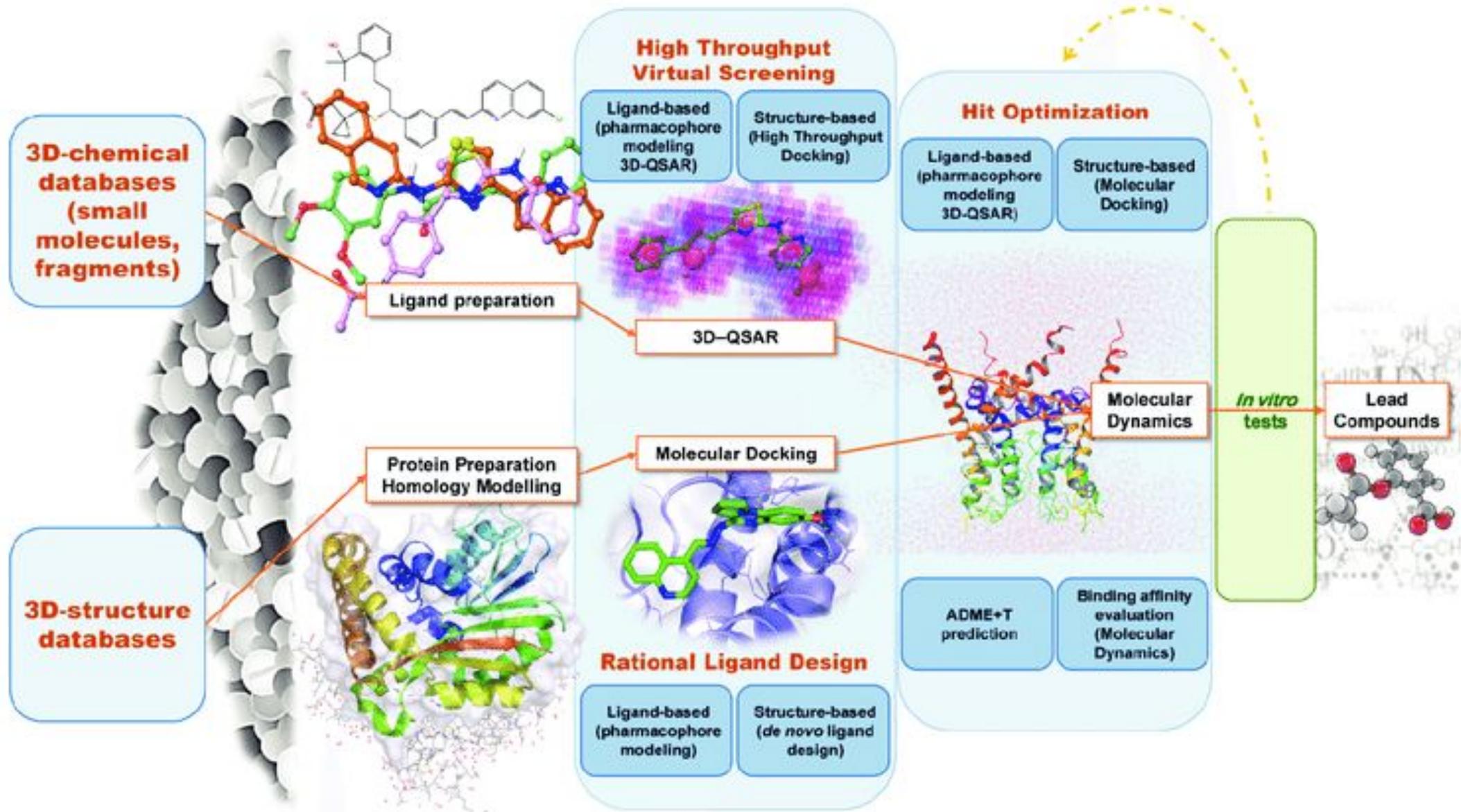


- ◆ Ligand based
- ◆ Target based
- ◆ Ligand + Target based



1. Выбираем “target” и “ligand”
2. Строим модель фармакофора
3. Отбираем молекулы, которые в большей степени подходят под структуру фармакофора
4. Фильтруем молекулы по ADME свойствам и токсичности
5. Молекулярный докинг





# Находим свою мишень “Target”

## Therapeutic Target Database



Therapeutic Target Database



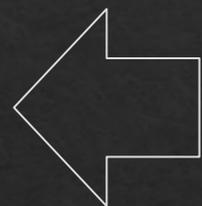
BIDD  
Bioinformatics and  
Drug Design group

PDB

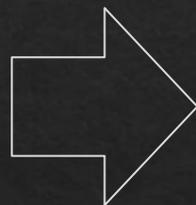
Advanced Search ▾ Target Group ▾ Drug Group ▾ Patient Data ▾ Model & Study ▾ Download ▾

	A	B	C
	<b>Entry</b>	<b><u>Protein names</u></b>	<b>Cross-reference (HGNC)</b>
1	<a href="#">P04229</a>	HLA class II histocompatibility antigen, DRB1...	<a href="#">HGNC:4948. HLA-DRB1.</a>
2	<a href="#">P26439</a>	3 beta-hydroxysteroid dehydrogenase/Delta 5--...	<a href="#">HGNC:5218. HSD3B2.</a>
3	<a href="#">P08888</a>	5-hydroxytryptamine receptor 1A	<a href="#">HGNC:5286. HTR1A.</a>
4	<a href="#">P28222</a>	5-hydroxytryptamine receptor 1B	<a href="#">HGNC:5287. HTR1B.</a>
5	<a href="#">P28223</a>	5-hydroxytryptamine receptor 2A	<a href="#">HGNC:5293. HTR2A.</a>
6	<a href="#">P41595</a>	5-hydroxytryptamine receptor 2B	<a href="#">HGNC:5294. HTR2B.</a>
7	<a href="#">P28335</a>	5-hydroxytryptamine receptor 2C	<a href="#">HGNC:5295. HTR2C.</a>

ChEMBL DB



Мой скриншот



TTD\_mod

	A	B	C
1	Entry	Protein names	ligand count
2	P43220	Glucagon-like peptide 1 receptor	109090
3	P08684	Cytochrome P450 3A4	49298
4	Q03431	Parathyroid hormone/parathyroid hormone-relat...	47085
5	P10635	Cytochrome P450 2D6	31461
6	P16473	Thyrotropin receptor	29930
7	P11473	Vitamin D3 receptor	26176
8	P00533	Epidermal growth factor receptor	24402
9	Q12809	Potassium voltage-gated channel subfamily H m...	22156

MEGA/TTD\_mod

- ◆ PDB (<https://www.rcsb.org/>)
- ◆ UniProt (<https://www.uniprot.org/>)
- ◆ ChEMBL  
(<https://www.ebi.ac.uk/chembl/>)

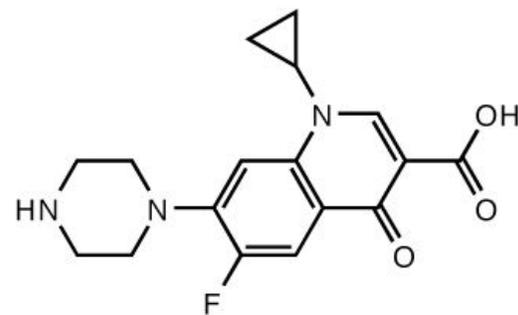
# SMILES 😊

Simplified  
molecular-input  
line-entry system

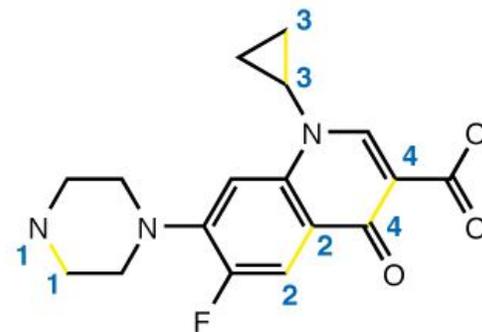
(Упрощенная система  
молекулярного ввода-

SMILES	Name
CC	ethane
O=C=O	carbon dioxide
C#N	hydrogen cyanide
CCN(CC)CC	triethylamine
CC(=O)O	acetic acid
C1CCCCC1	cyclohexane
c1ccccc1	benzene

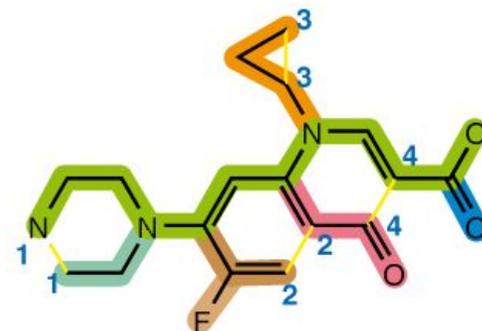
A



B



C



D

N1CCN(CC1)C(C(F)=C2)=CC(=C2C4=O)N(C3CC3)C=C4C(=O)O



colab +  GitHub

Спасибо за внимание!