Applied Mathematics and Informatics In Drug Discovery (2024)

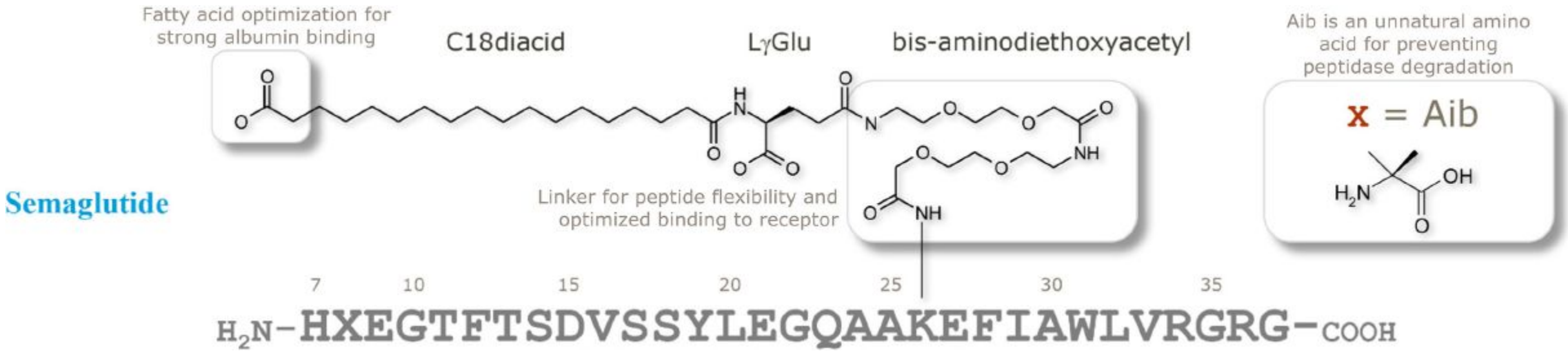


Figure from Knudsen *et al.*, 2019

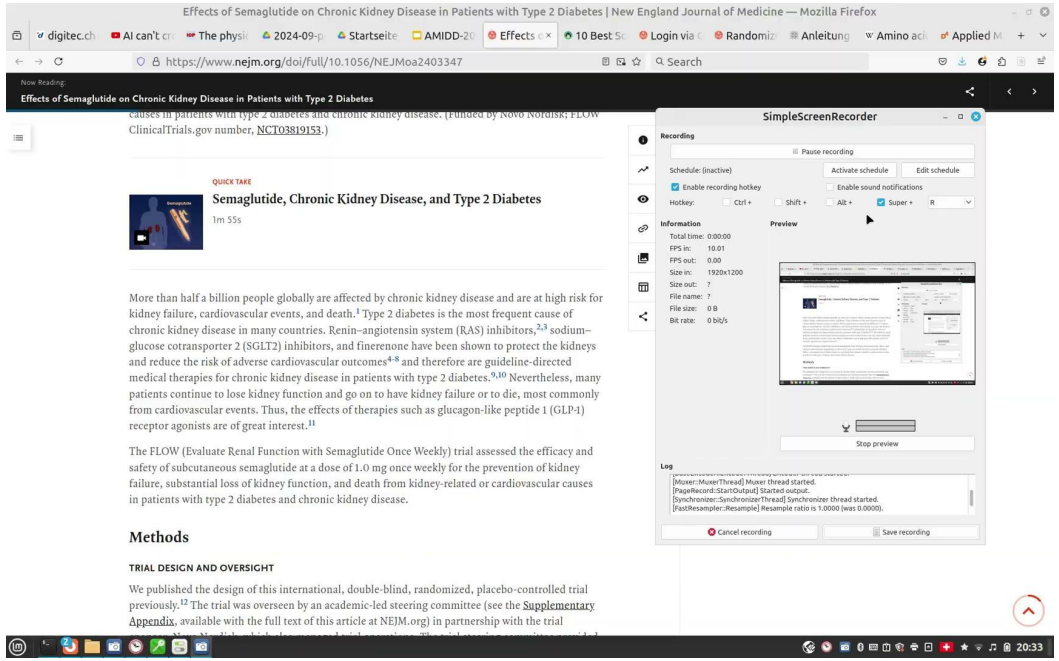
Dr. Jitao David Zhang, Computational Biologist

[1] Pharmaceutical Sciences, Pharma Research and Early Development, Roche Innovation Center Basel, F. Hoffmann-La Roche;
[2] Department of Mathematics and Computer Sciences, University of Basel

Outline

1. Learning drug discovery backwards
2. Economy and productivity of drug discovery
3. Why mathematics and computer science in drug discovery?
4. About the course

FLOW: a Phase 3 clinical trial for Semaglutide 1.0 mg once a week for people with type 2 diabetes and chronic kidney disease



Q1: What is semaglutide?
Q2: What are the proven benefits of semaglutide for patients with Type II diabetes?
Q3: Why was the clinical trial conducted?
Q4: How many patients participated in the trial?
Q5: What was the primary outcome of the trial?
Q6: How many patients experienced serious adverse effects?
Q7: What was the conclusion of the authors?

Perkovic, V. et al. Effects of Semaglutide on Chronic Kidney Disease in Patients with Type 2 Diabetes. New England Journal of Medicine 391, 109–121 (2024). ClinicalTrials.gov: NCT03819153

Glucagon-like peptide 1 (GLP-1) and its receptor GLP1R

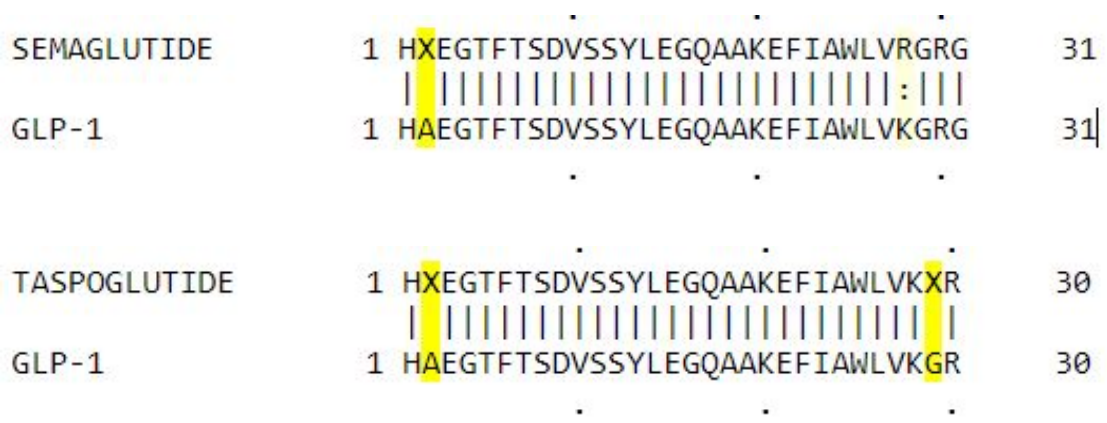
GLP-1 is a 30- or 31-amino acid long peptide. It is derived from the proglucagon peptide. Proglucagon peptide is a product of the cleavage of preproglucagon, which is a protein encoded by the gene GCG in human.

Endogenous GLP-1 has a half-life of about 2 minutes. Synthesized agonists have similar sequences but much longer half-life due to chemical modifications. Examples include:

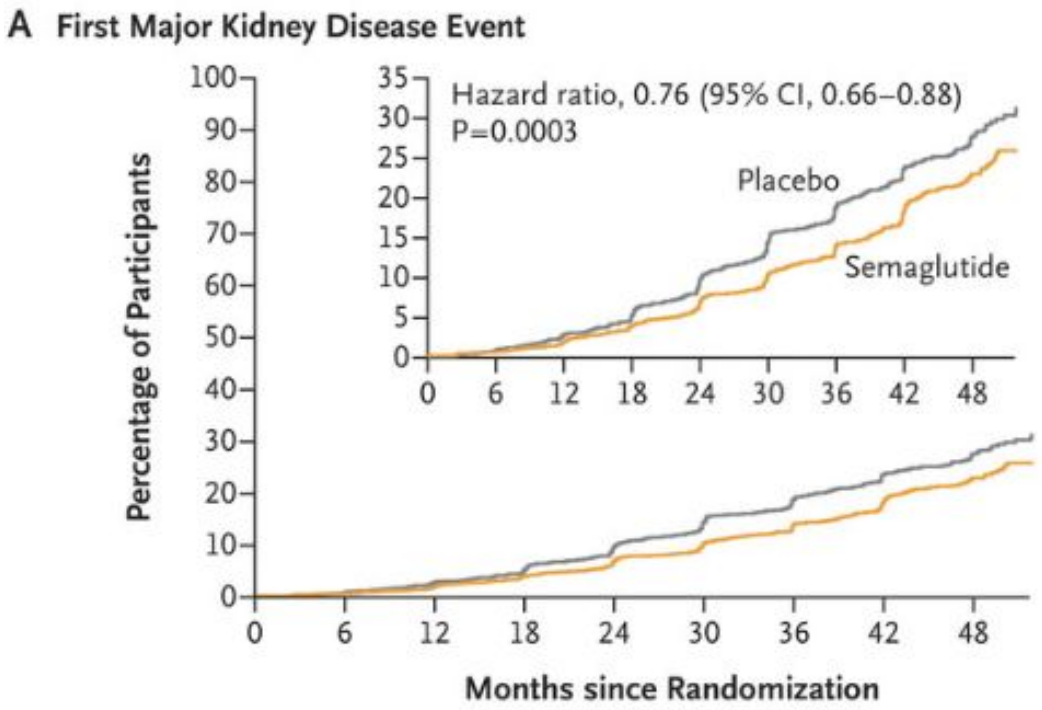
- Semaglutide (right top panel, approved)
- Taspoglutide (right middle panel, tested, not approved).

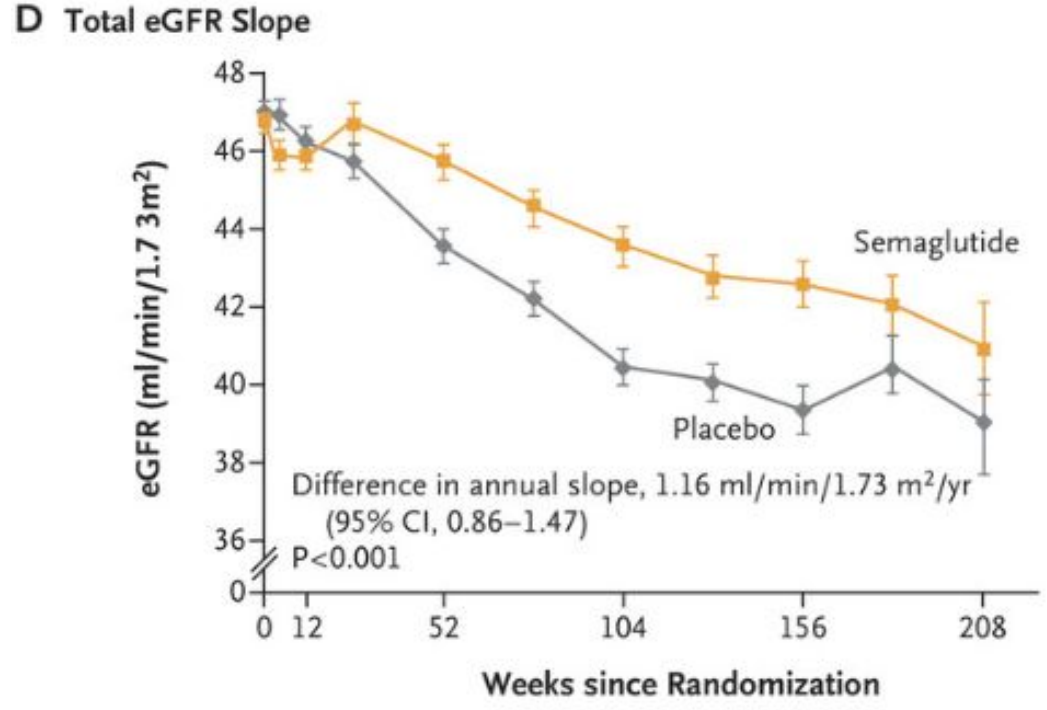
GLP1R is a GPCR receptor found on beta cells of the pancreas and on neurons of the brain. It is involved in the regulation of blood sugar level by enhancing insulin secretion (*incretin signaling).*

Right bottom: crystal structure of GLP-1 (orange) bound to GLP-1R (green) (PDB: 3IOL)



The primary outcome (major kidney disease events) and one confirmatory secondary outcome (reduction of the estimated glomerular filtration rate, or eGFR) of the FLOW trial

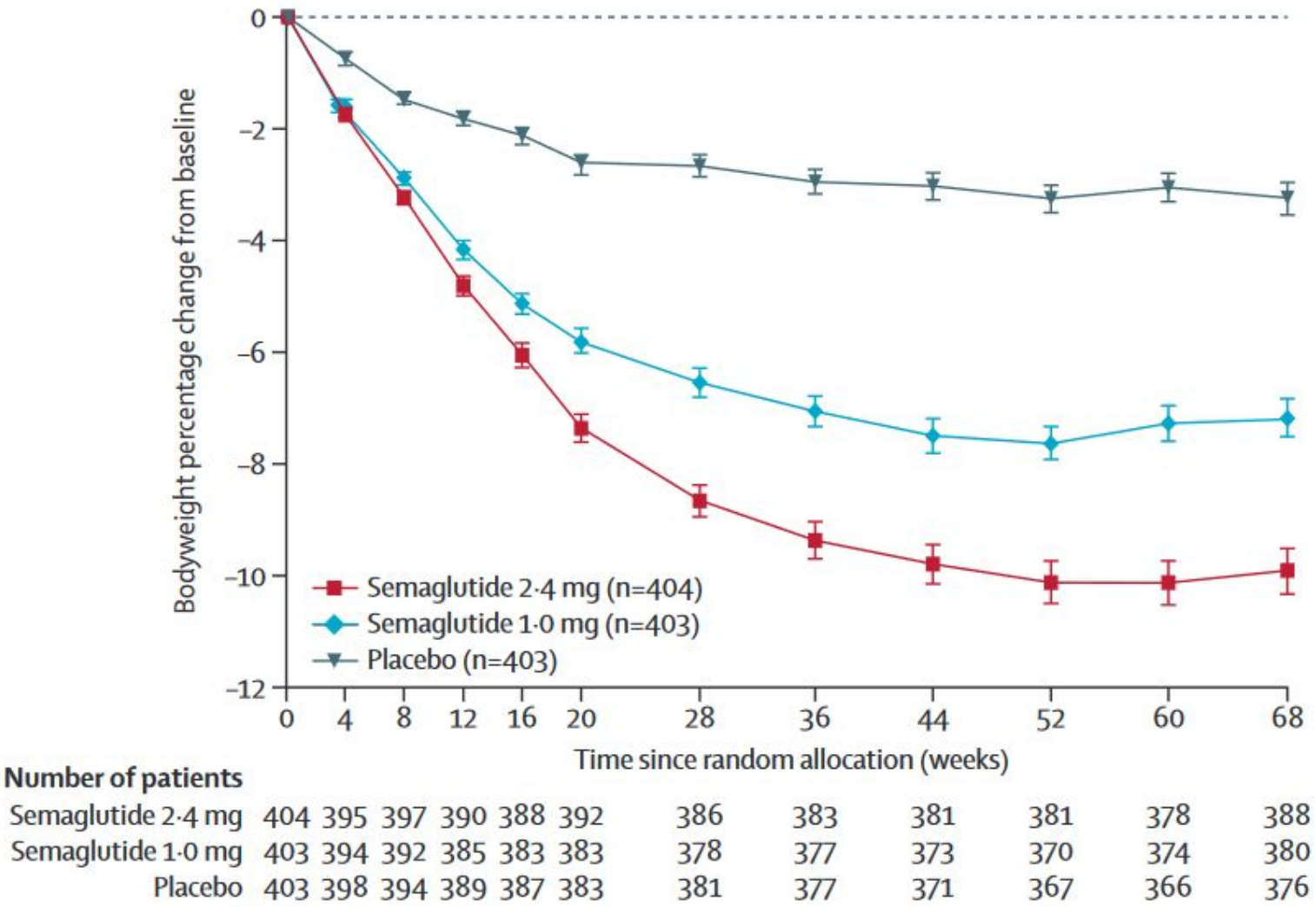




Safety outcome

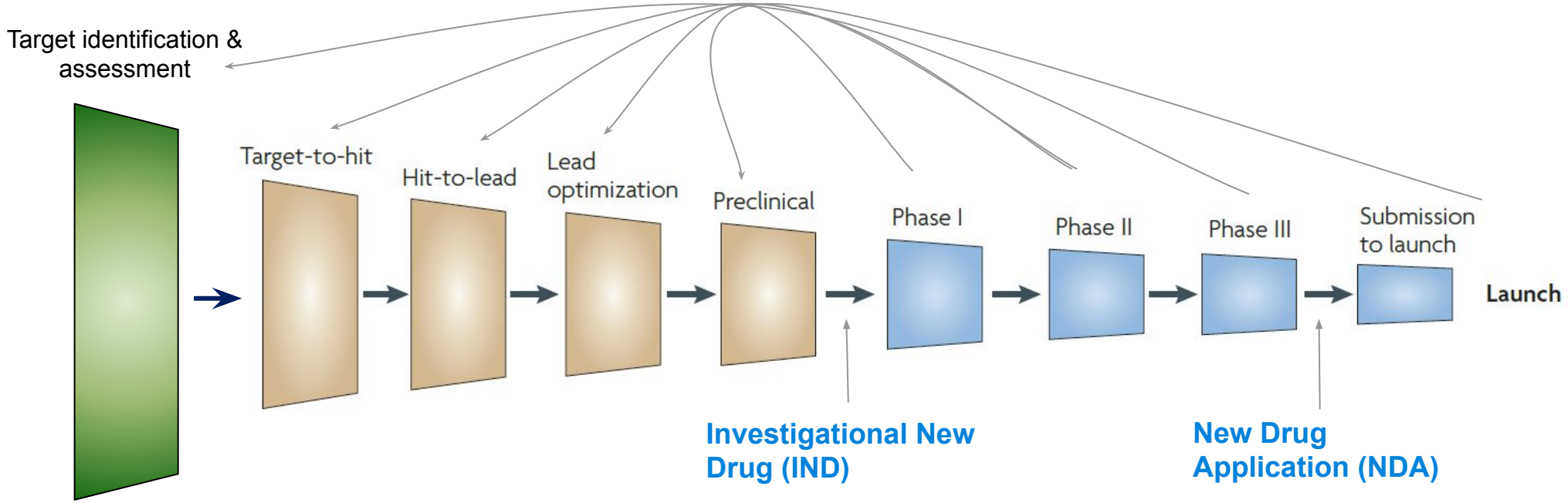
Adverse Event	Semaglutide (N=1767)	Placebo (N=1766)
	no. of participants (%)	
Serious adverse event	877 (49.6)	950 (53.8)
Adverse event leading to permanent discontinuation of semaglutide or placebo	233 (13.2)	211 (11.9)
Prespecified adverse events of special interest		
Diabetic retinopathy*	402 (22.8)	398 (22.5)
Covid-19–related disorder	358 (20.3)	404 (22.9)
Serious adverse event: cardiovascular disorder	273 (15.4)	319 (18.1)
Heart failure*	133 (7.5)	175 (9.9)
Acute kidney failure*	172 (9.7)	182 (10.3)
Malignant tumor*	120 (6.8)	104 (5.9)
Serious adverse event: gastrointestinal disorder	95 (5.4)	94 (5.3)
Serious adverse event: rare event	48 (2.7)	57 (3.2)
Acute gallbladder disease*	32 (1.8)	39 (2.2)
Severe hypoglycemia*	37 (2.1)	37 (2.1)
Medication error*	15 (0.8)	13 (0.7)
Serious adverse event: hepatic disorder	18 (1.0)	20 (1.1)
Acute pancreatitis*	10 (0.6)	7 (0.4)
Serious adverse event: allergic reaction	6 (0.3)	9 (0.5)
Serious adverse event: abuse and misuse	1 (0.1)	4 (0.2)
Serious adverse event: suspected transmission of infectious agent through semaglutide or placebo	0	1 (0.1)

STEP2: a clinical trial for Semaglutide 2.4mg once a week for overweight, obesity, and type-2 diabetes



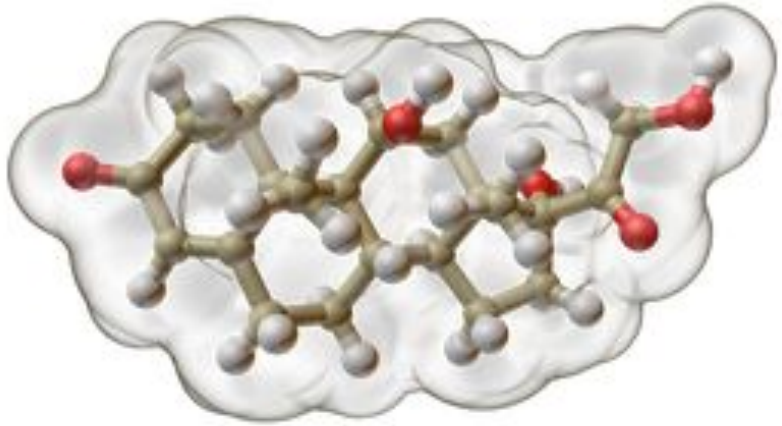
Semaglutide 2·4 mg once a week in adults with overweight or obesity, and type 2 diabetes (STEP 2): a randomised, double-blind, double-dummy, placebo-controlled, phase 3 trial, Davies *et al.*, 2021

The linear view of drug discovery

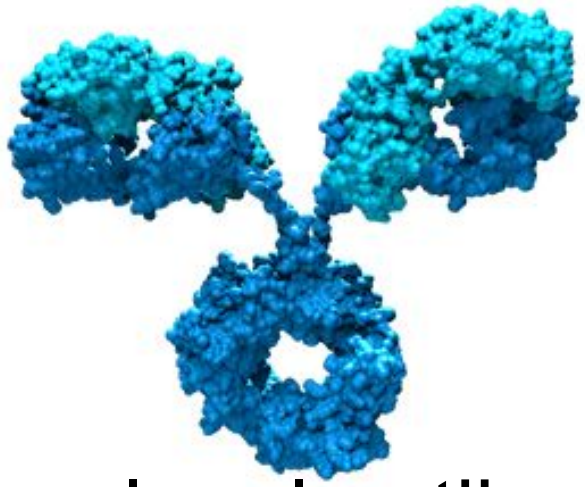


Adapted from Paul *et al.* "How to Improve R&D Productivity: The Pharmaceutical Industry's Grand Challenge." Nature Reviews Drug Discovery, 2010

A zoo of modalities

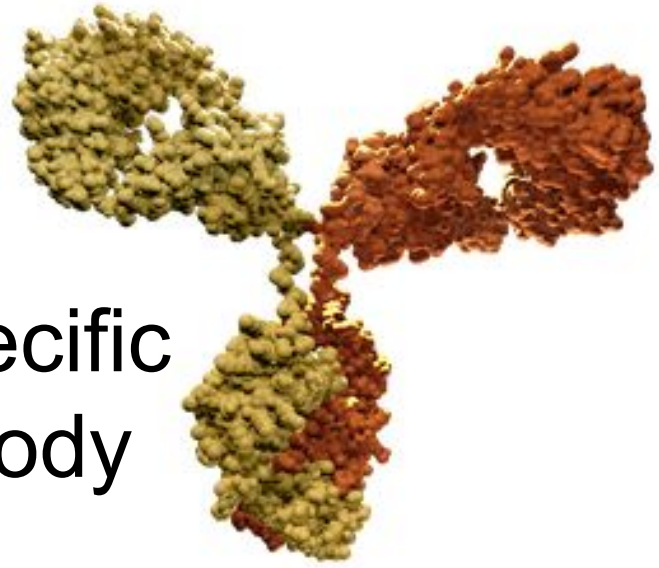


Small molecule

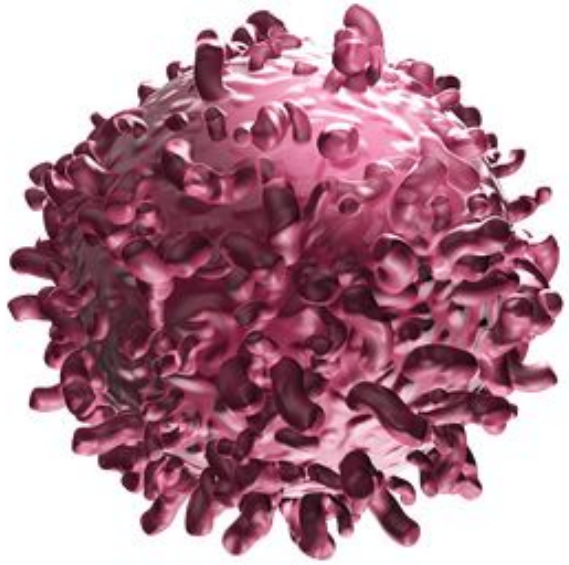


Monoclonal antibody

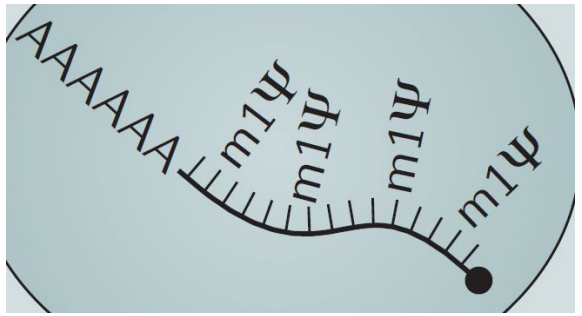
Oligonucleotides



Bispecific antibody



Chimeric Antigen Receptor (CAR) T-cells

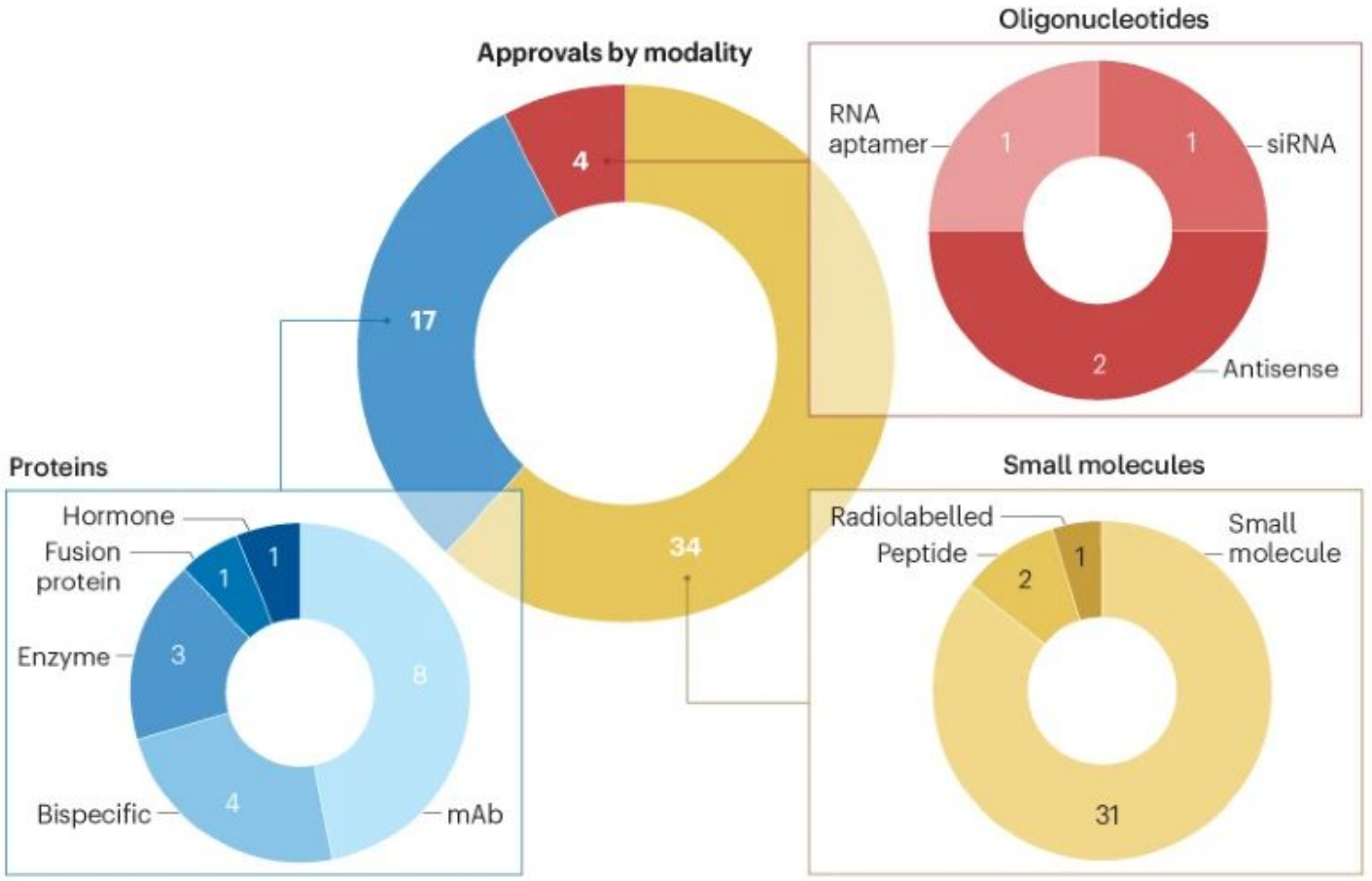


mRNA vaccines

Novel drugs approved by the FDA's Center for Drug Evaluation and Research (CDER) in 2023

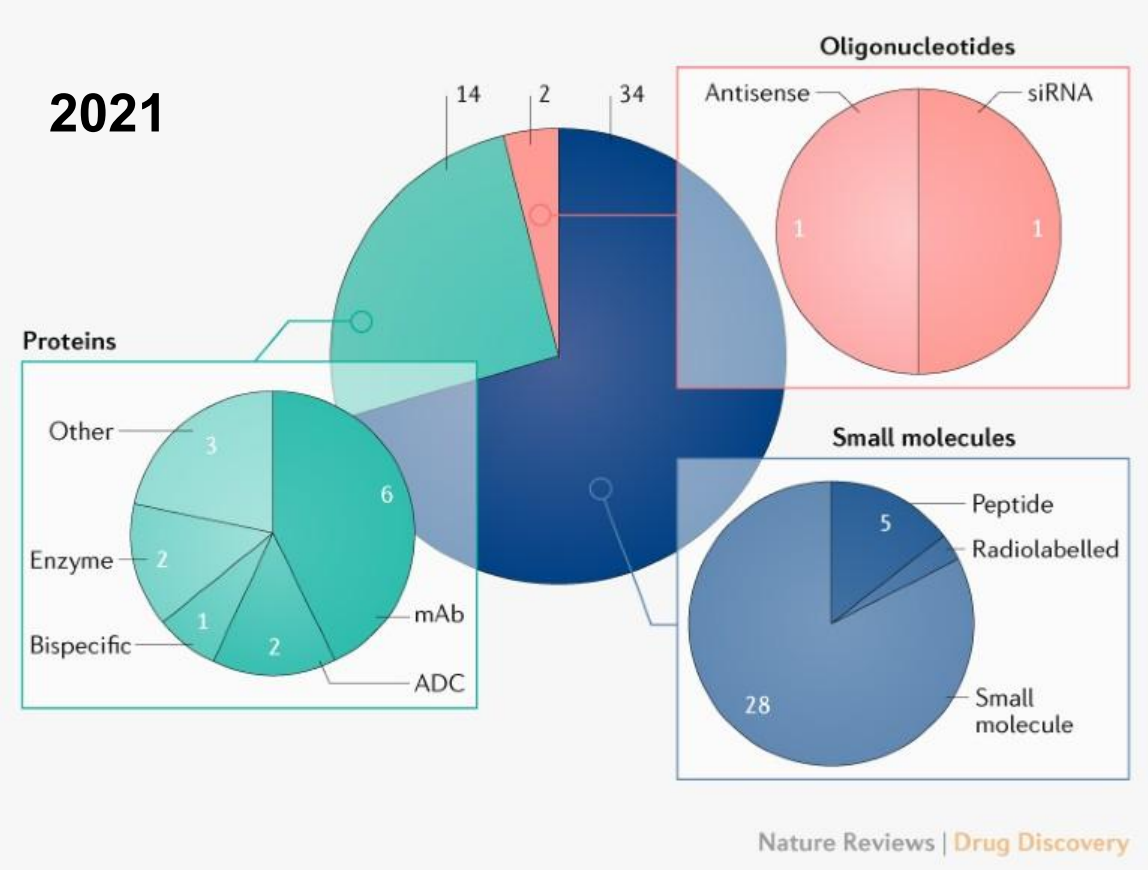
Three modalities

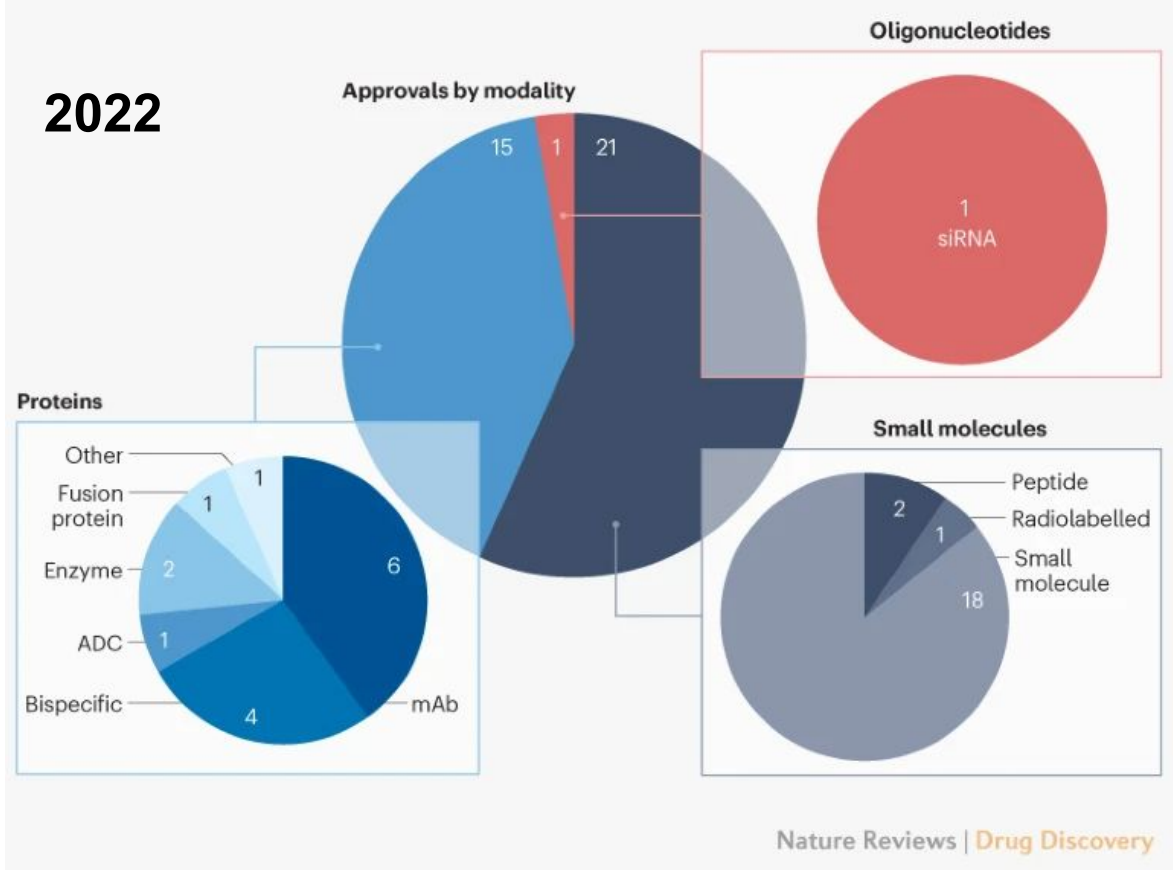
- Small molecules (molecular weight under 900 daltons)
- Proteins
  - mAb: monoclonal antibody
  - Bispecific: bispecific antibodies
- Oligonucleotides
  - siRNA: small interfering RNA



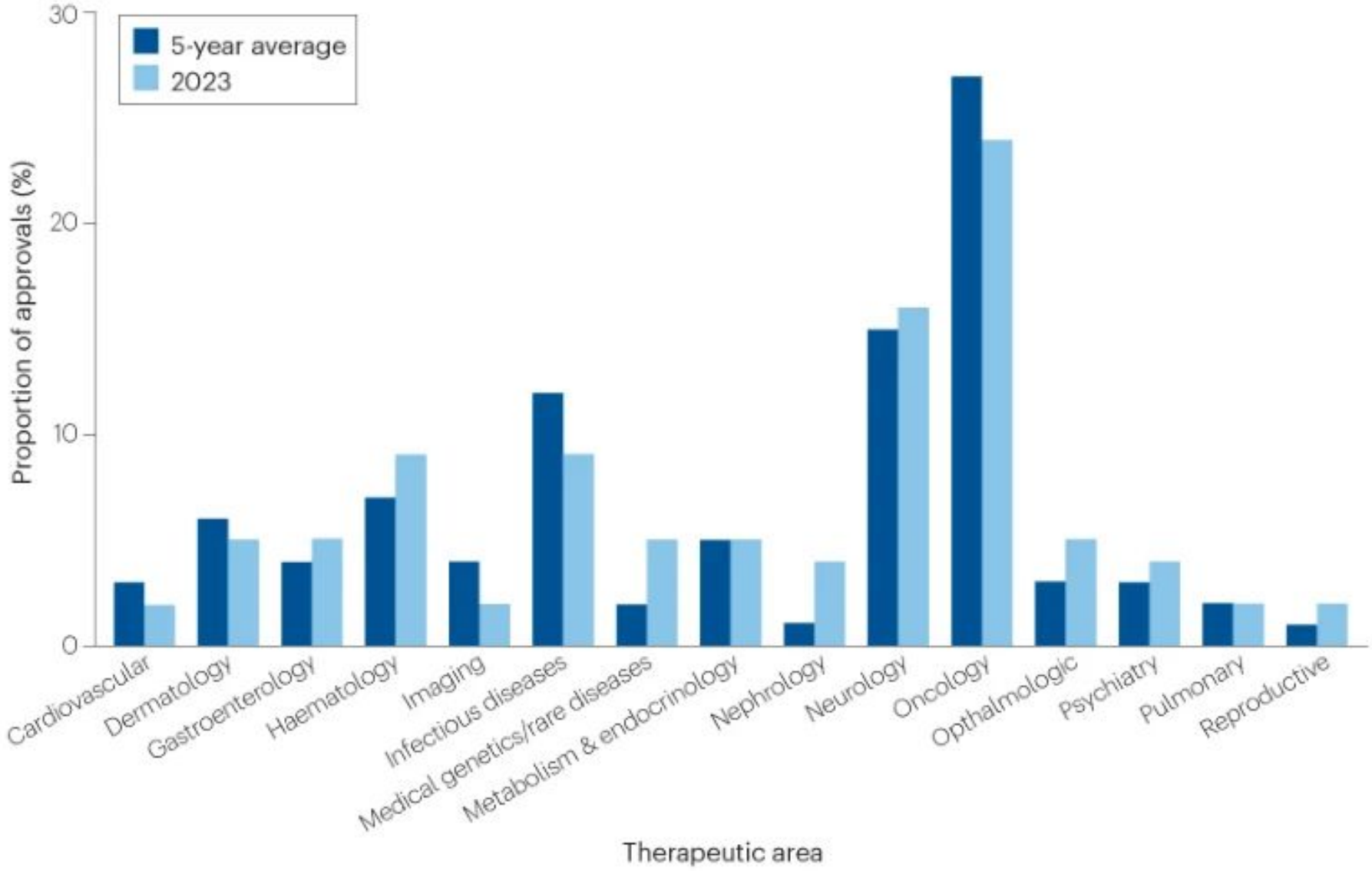
Source: Asher Mullard, Nature Reviews Drug Discovery, 2024. The list can be found on FDA's website

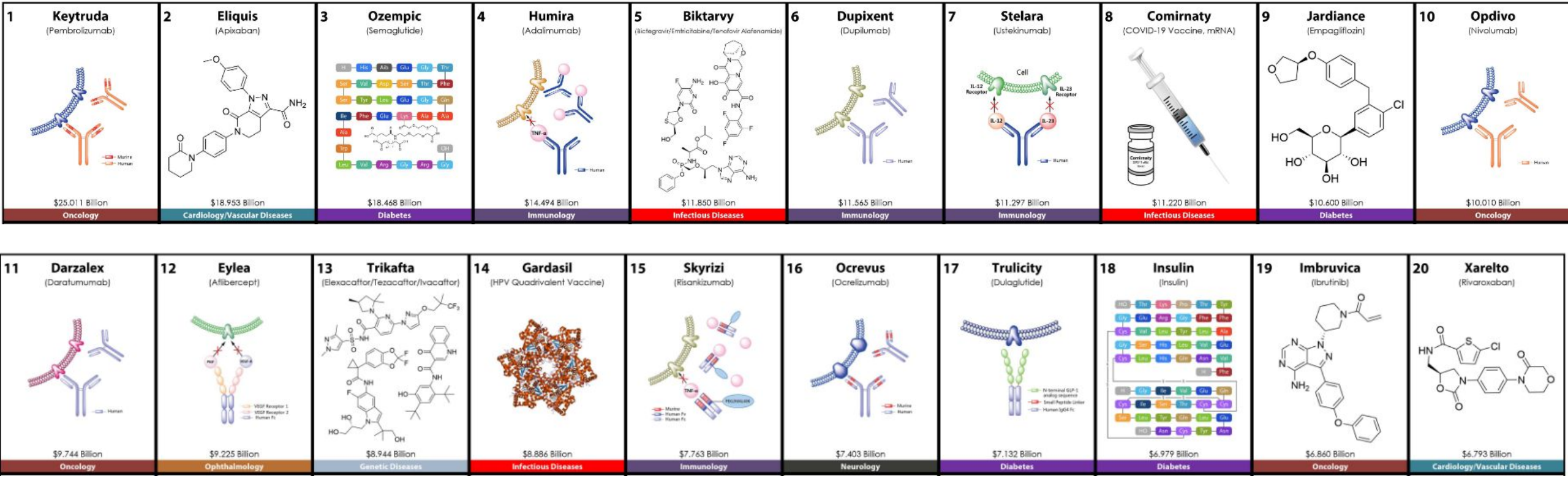
Relative contributions of modalities remain constant in the past three years





New drug approvals vary between disease areas





Top 20 pharmaceuticals by sales in 2023

Poster compiled by the Jon Njardarson group at University of Arizona (https://njardarson.lab.arizona.edu). Citation: J. Chem. Ed. 2010, 87, 1348.

Questions: (1) How many are small molecules, proteins, and oligonucleotides each? (2) Are there other modalities? (3) What patterns do you observe? (4) Do you have explanations for these patterns?

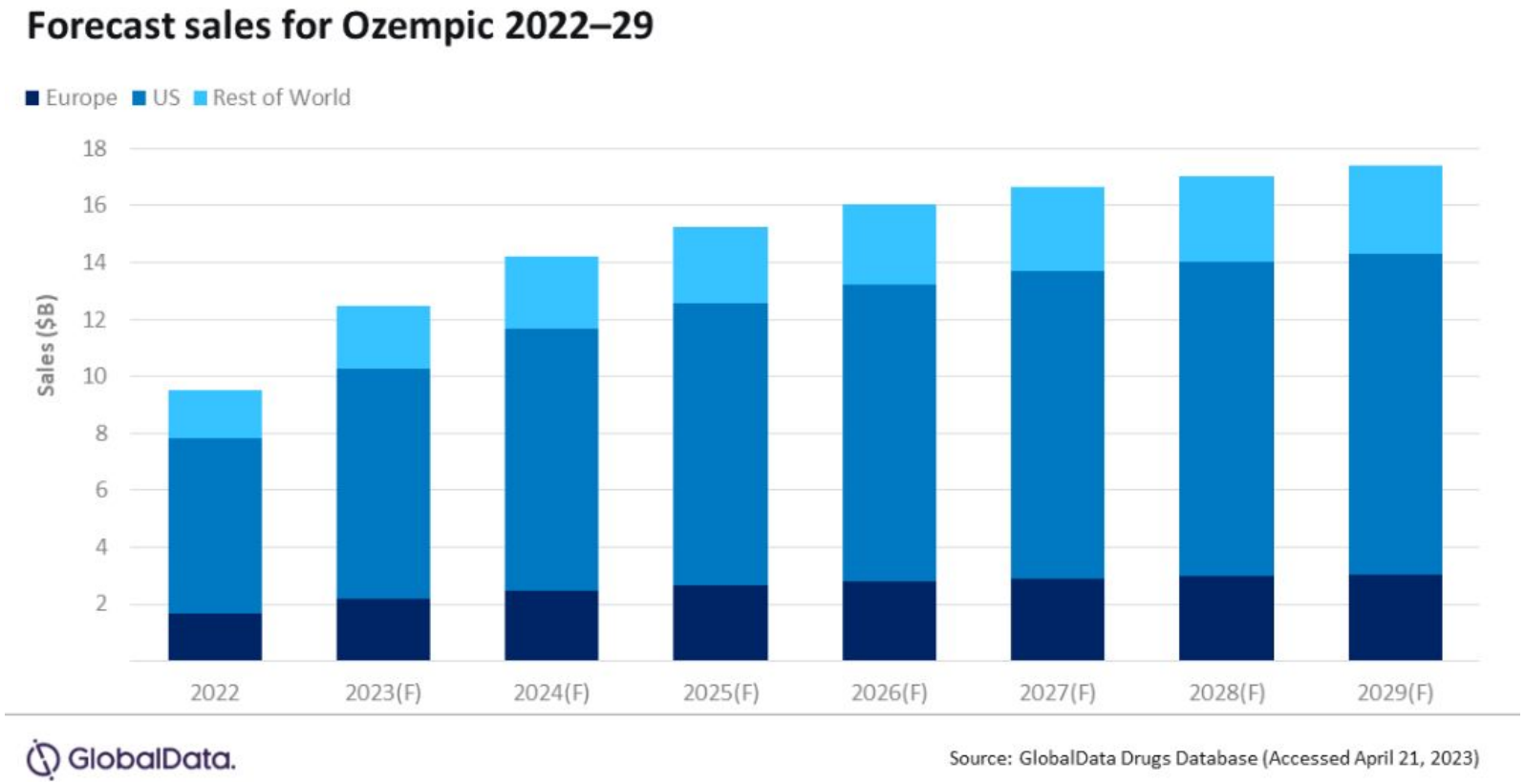
Forms and market size of Semaglutide

Ozempic, as well as Wegovy and Rybelsus, are **brand names** of semaglutide.

Ozempic was approved by the FDA for type 2 diabetes.

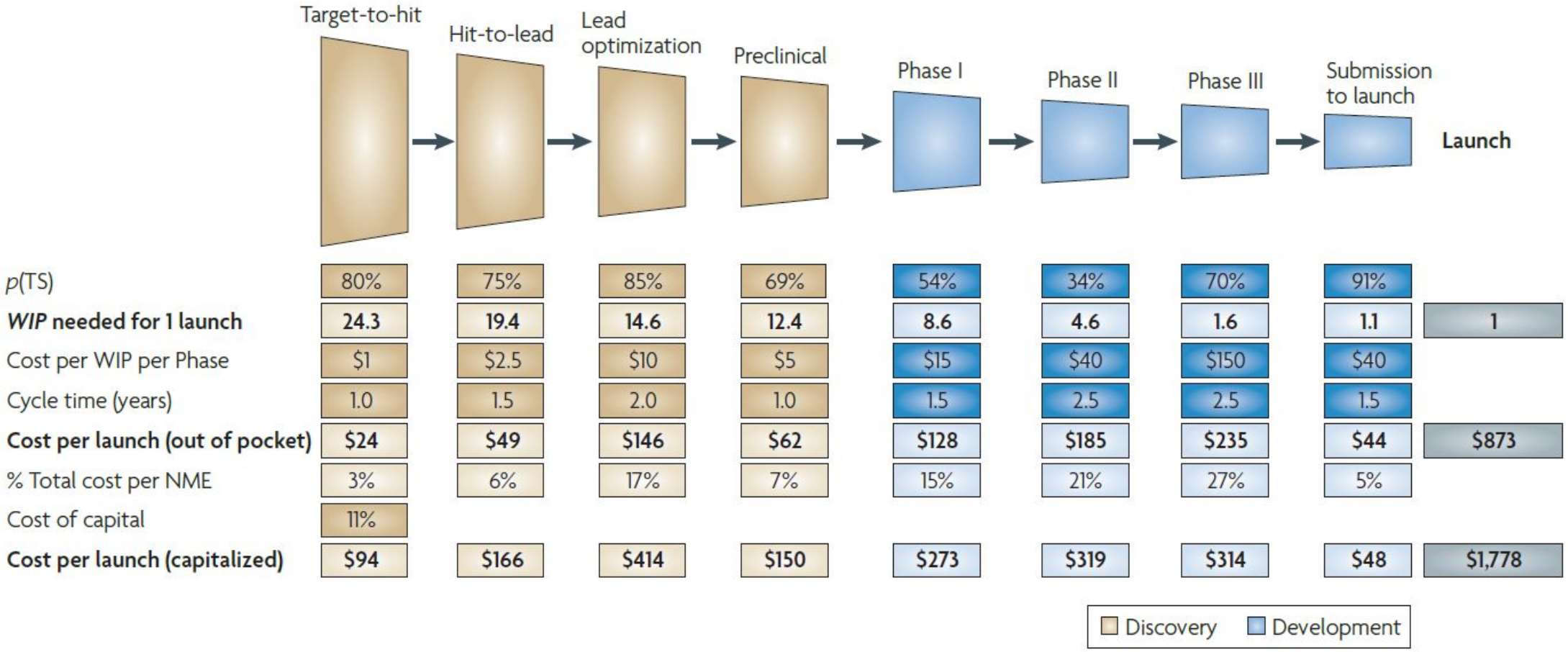
Wegovy was approved by the FDA for weight management at once-weekly 2.4 mg injectable doses in 2022.

Rybelsus tablets are approved by the FDA used for adults with type 2 diabetes to control blood sugar levels.



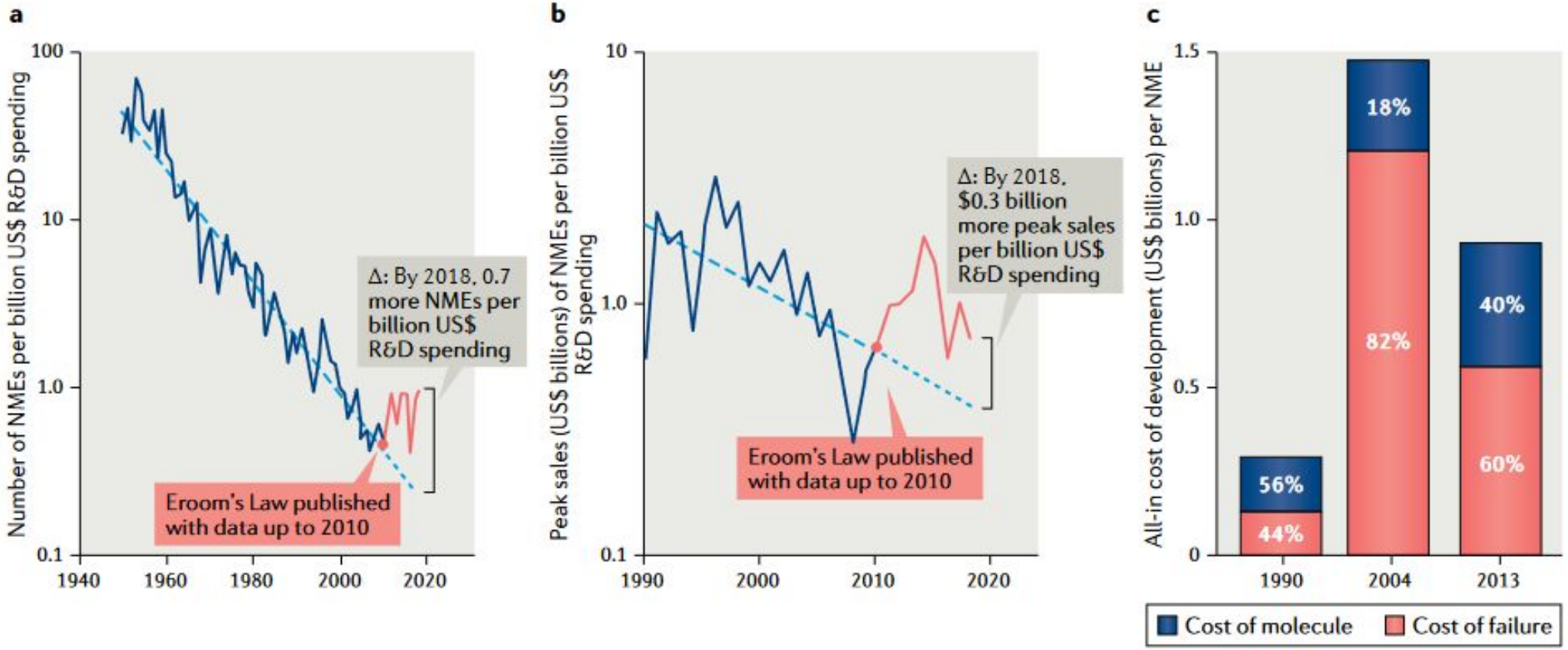
Data source: GlobalData

Risks and costs associated with each stage of the linear view of drug discovery



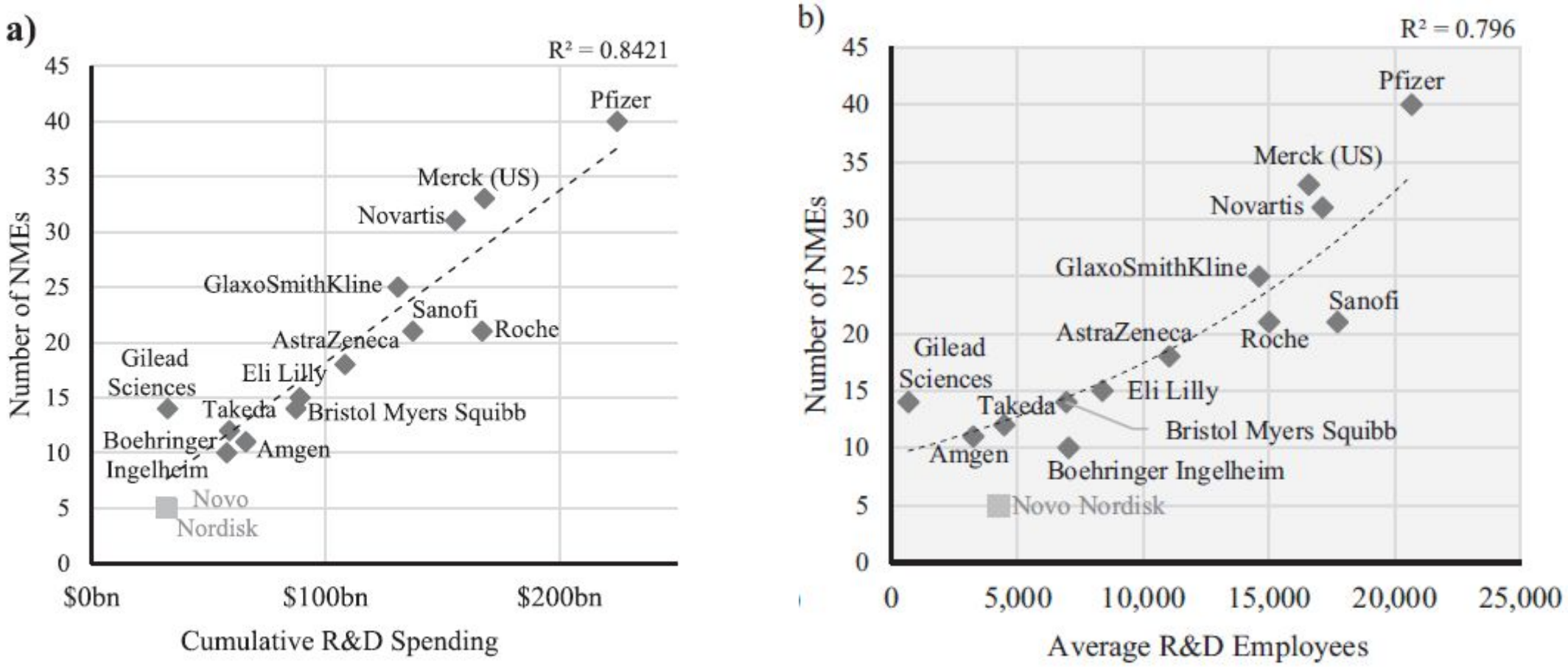
pTS: probability of technical success. **WIP**: work in progress; **Capitalized cost**: Out-of-pocket cost corrected for cost of capital, standard for long-term investments; **Out-of-pocket cost**: total cost required to expect one drug launch, taking into account attrition, but not the cost of capital; **Cost of capital**: annual rate of return expected by investors based on the level of risk of the investment. Paul *et al.*, Nature Reviews Drug Discovery, 2010.

(Breaking?) The Eroom's Law



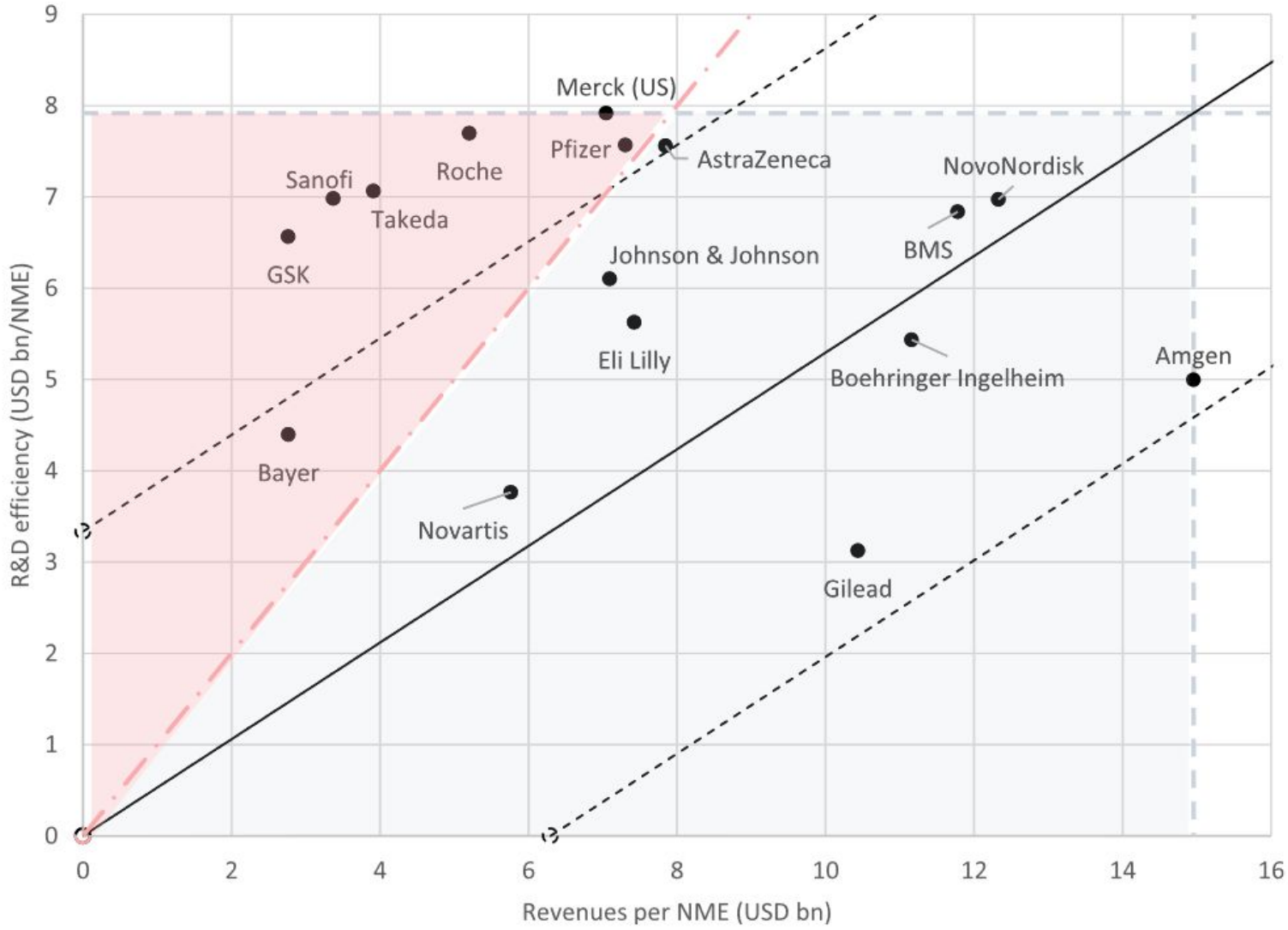
Ringel, Michael S., Jack W. Scannell, Mathias Baedeker, and Ulrik Schulze. "Breaking Eroom's Law." Nature Reviews Drug Discovery 19, no. 12 (April 16, 2020): 833–34.

Drug discovery and development require huge investment and large interdisciplinary teams



Schuhmacher, Alexander, Lucas Wilisch, Michael Kuss, Andreas Kandelbauer, Markus Hinder, and Oliver Gassmann. "R&D Efficiency of Leading Pharmaceutical Companies – A 20-Year Analysis." *Drug Discovery Today* 26, no. 8 (August 1, 2021): 1784–89. https://doi.org/10.1016/j.drudis.2021.05.005.

Profits generated by new molecule entities (NMEs) cannot cover the cost in some companies in the last 20 years

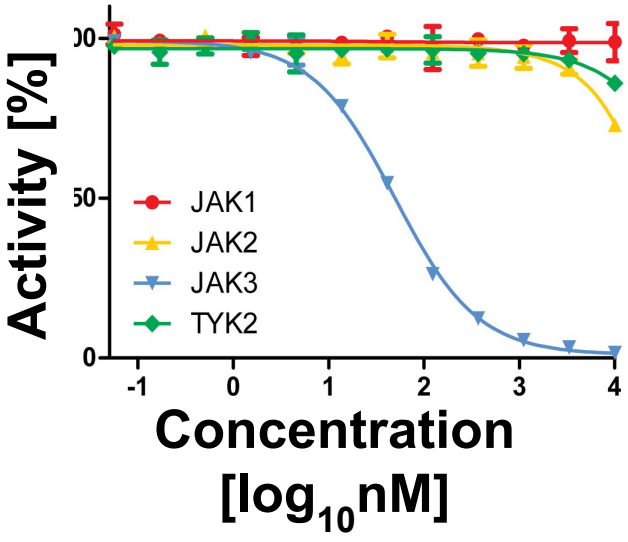


危机

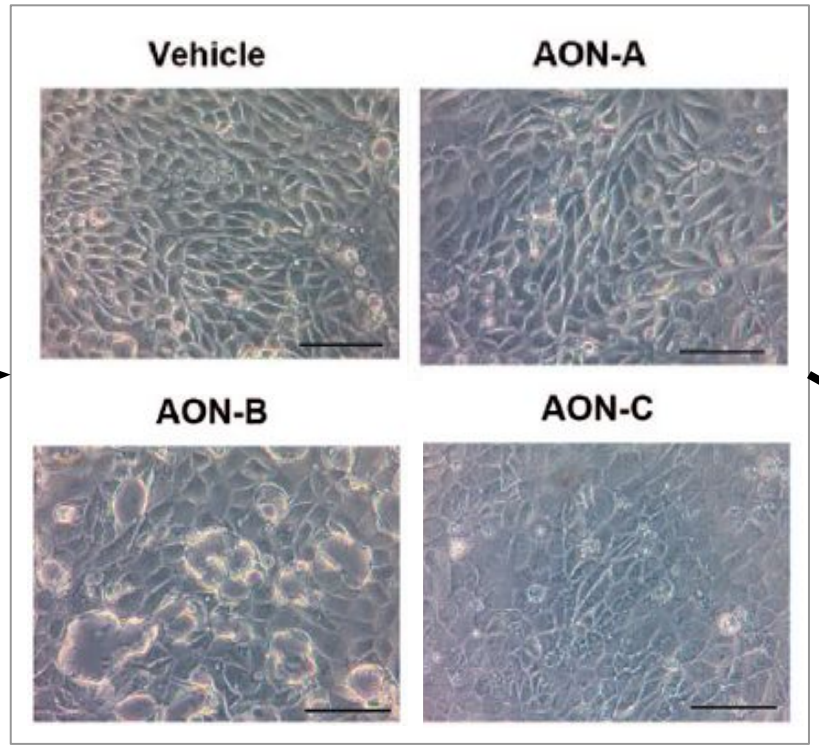
~~*n.* crisis~~

Danger + Opportunity

Classical workflow of efficacy and toxicity assessment

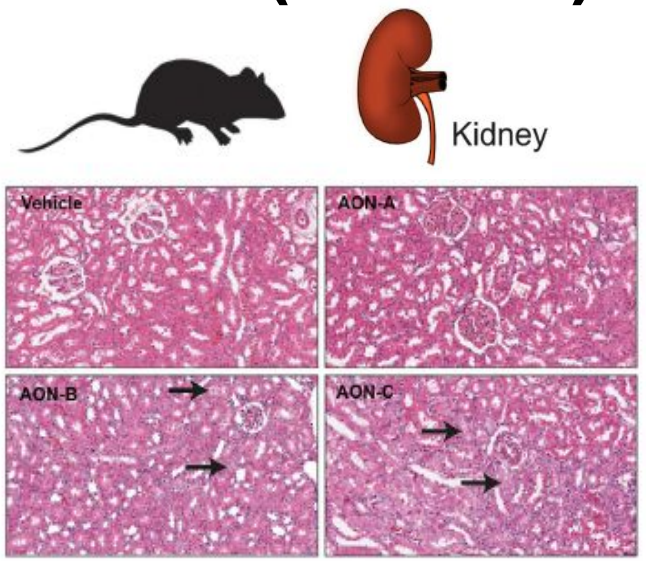


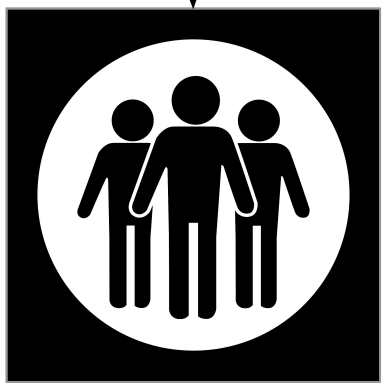
Biochemical & biophylical assays



Cellular assays (*in vitro*)

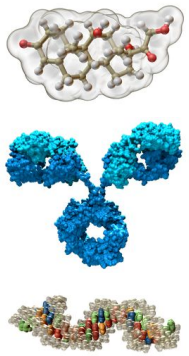
Animal experiments (*in vivo*)

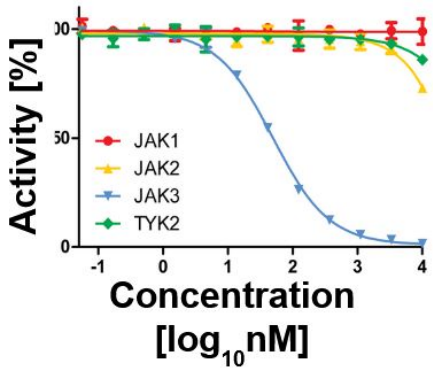


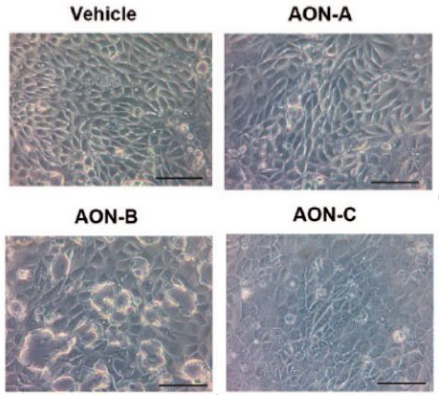


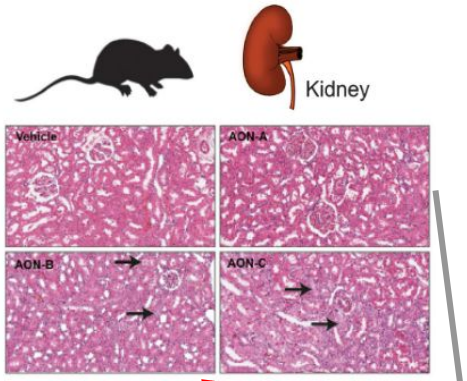
Clinical trials

Computational methods empower efficacy and toxicity assessment



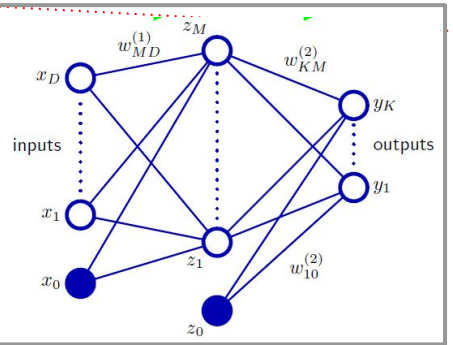






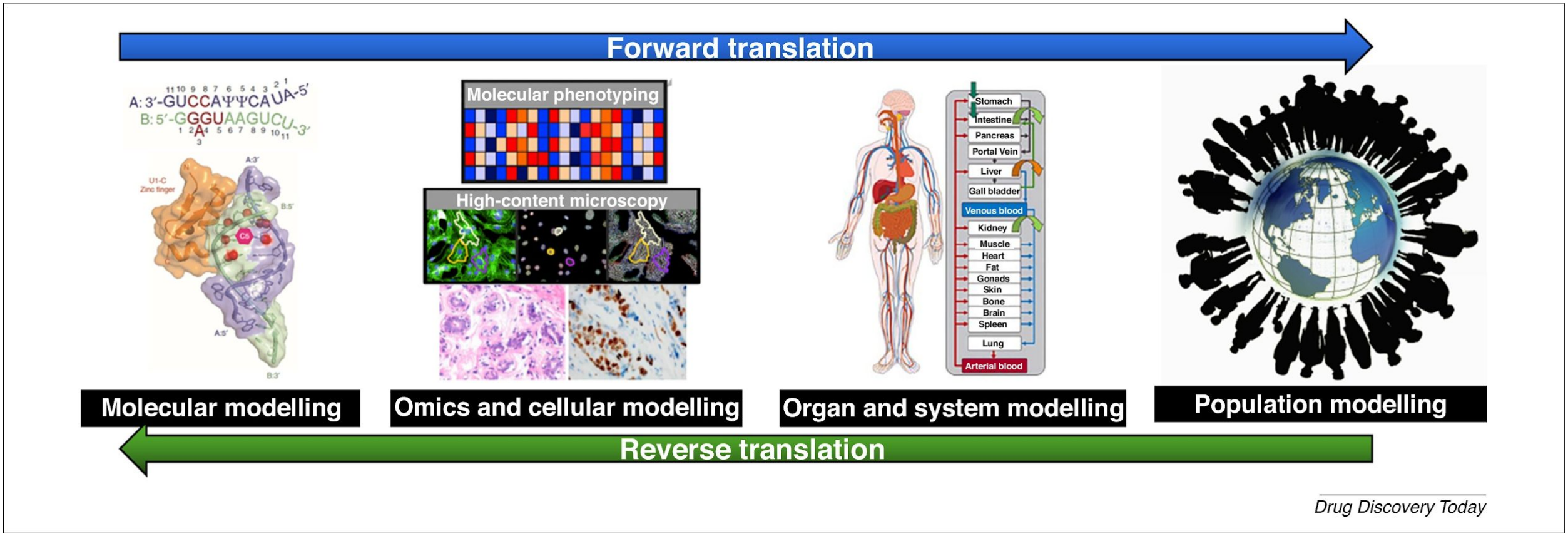
High-throughput technologies (omics, microscopy, etc.)

$$\frac{\partial u}{\partial t} = \underbrace{D\frac{\partial^2 u}{\partial x^2}}_{\text{Diffusion}} + \underbrace{ku}_{\text{Binding}}$$



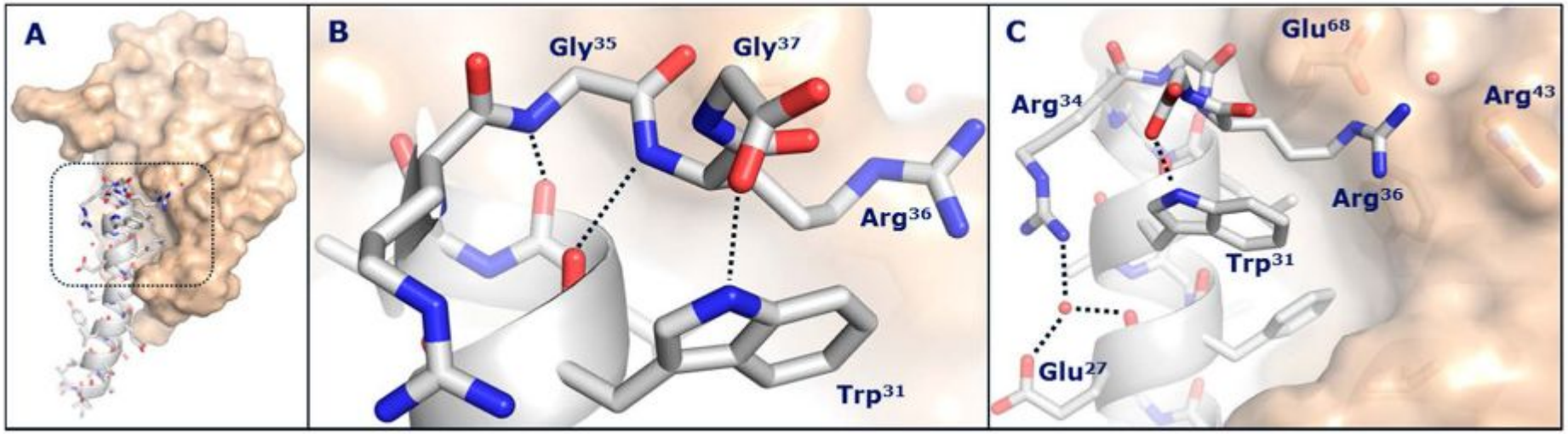
Mechanistic, causal, and statistical models

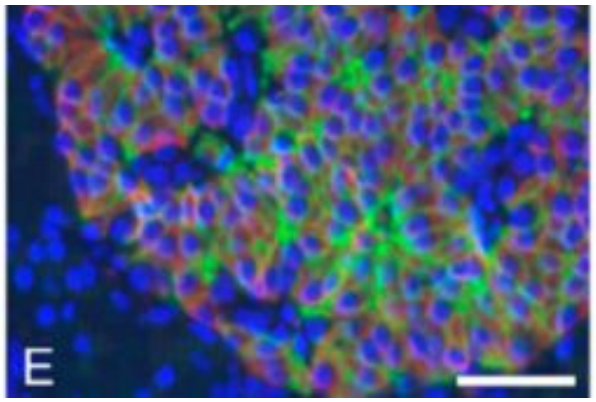
The multiscale modelling view of drug discovery

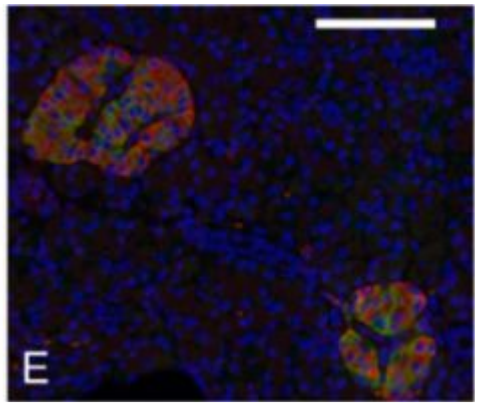


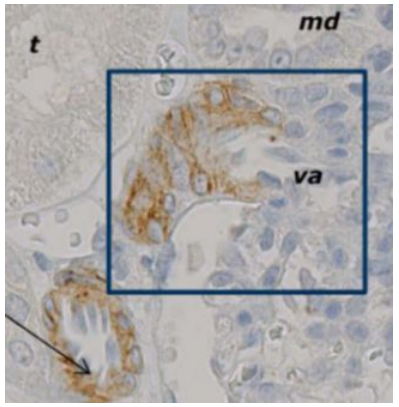
Zhang, Jitao David, Lisa Sach-Peltason, Christian Kramer, Ken Wang, and Martin Ebeling. 2020. “Multiscale Modelling of Drug Mechanism and Safety.” Drug Discovery Today 25 (3): 519–34. https://doi.org/10.1016/j.drudis.2019.12.009.

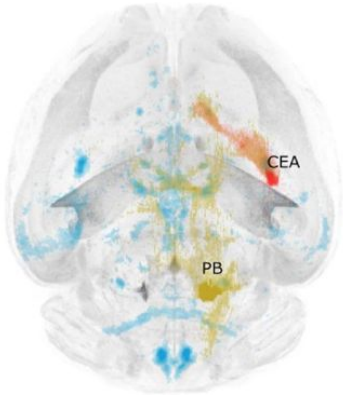
An example of multiscale understanding with semaglutide







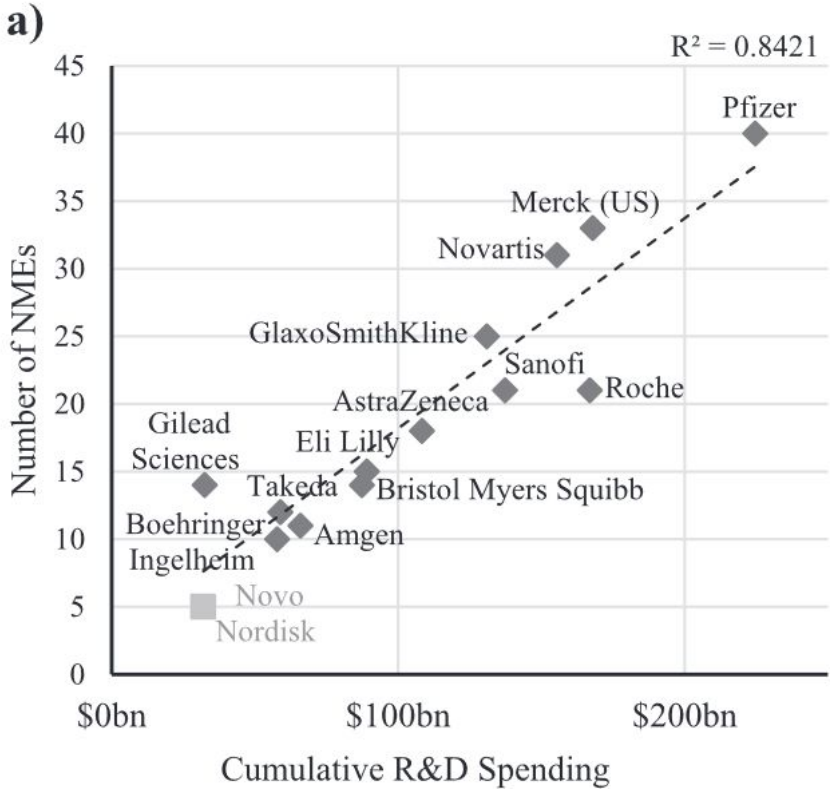




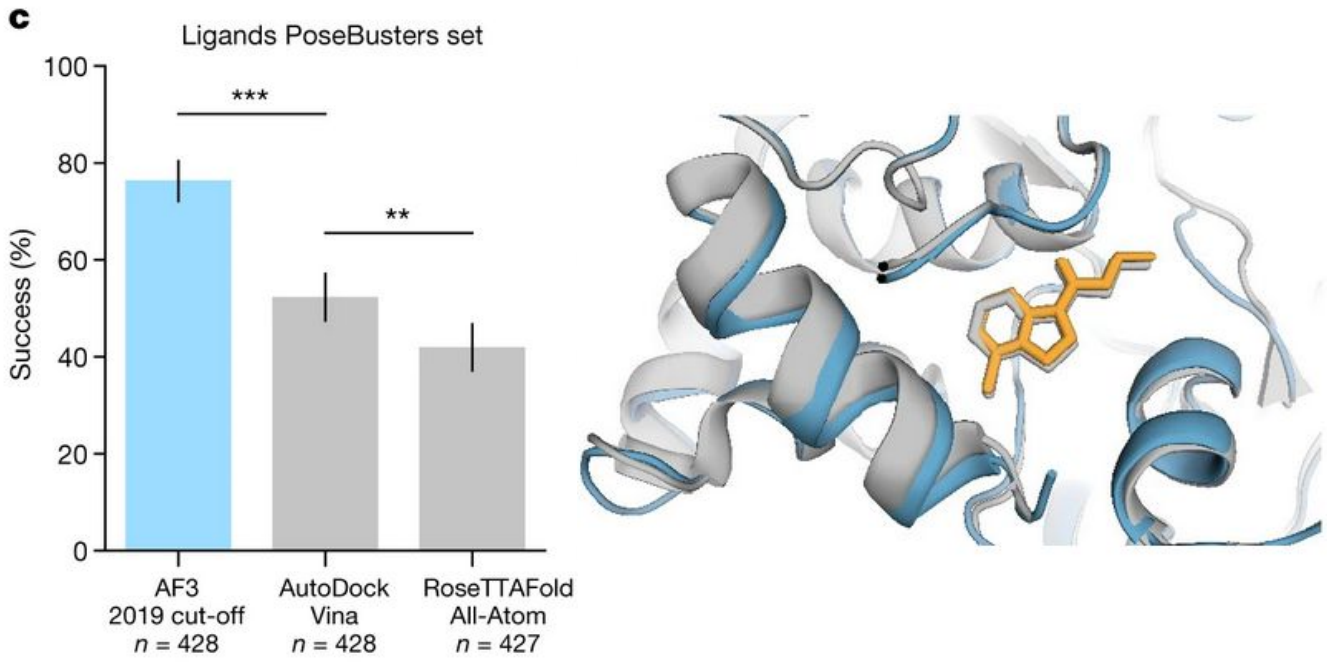
Top panels: crystal structure of the semaglutide peptide backbone (gray) in complex with its target, GLP-1 receptor (golden surfaces).

Bottom panels (from left to right): immunostaining of monkey pancreas, human pancreas, monkey muscle, and connectivity map of mice brain.

Quest of the course: to make drug discovery efficient and sustainable with mathematics and informatics



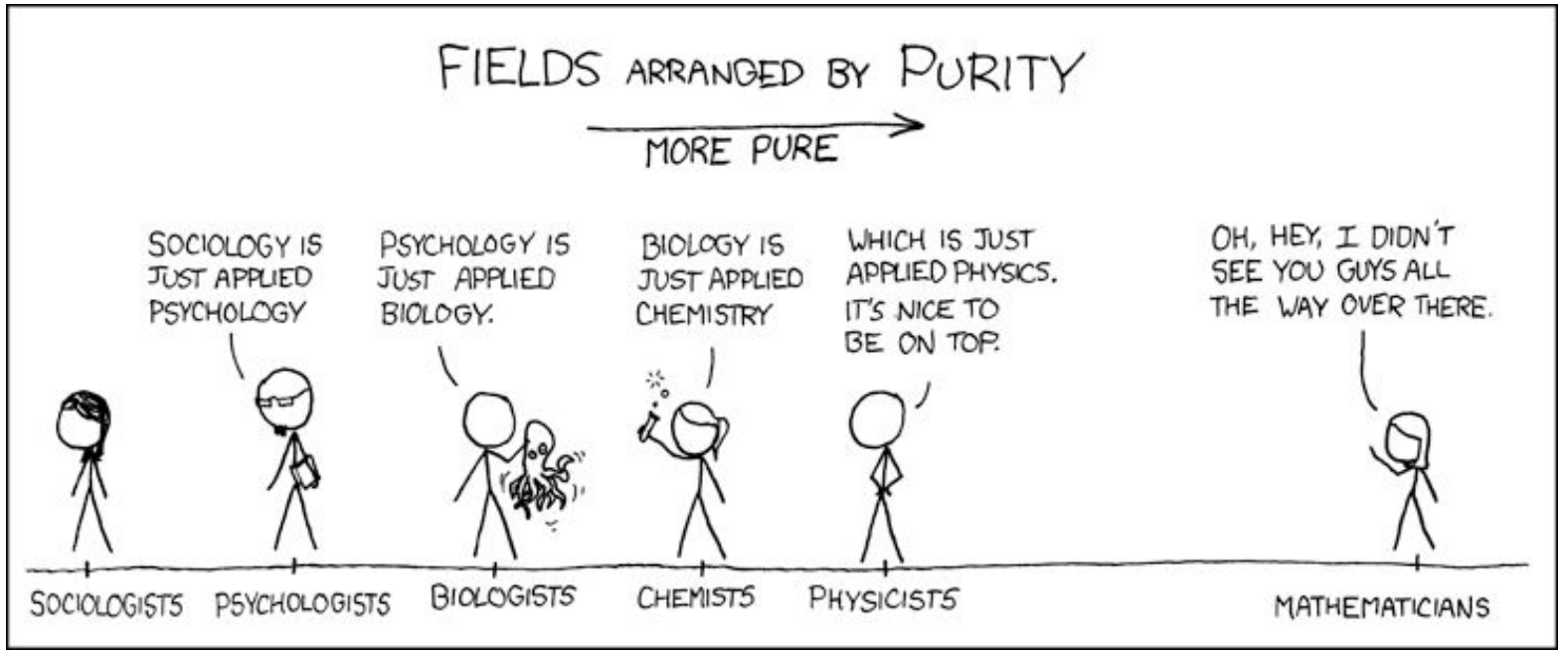
R&D efficiency of leading pharma companies, 1999-2018 (Schumacher *et al.*, 2021)



Accurate structure prediction of biomolecular interactions with *AlphaFold3* (Abramson *et al.*, 2024). The *PoseBuster* set: 428 protein-ligand released to PDB after 2021. Success: pocket-aligned ligand Root Mean Square Deviation (RMSD) of atomic positions <= 2Å. Right: AF3 prediction for which docking tools *Vina* and *Gold* were less accurate (Human Notum bound to inhibitor ARUK3004556)

Purity

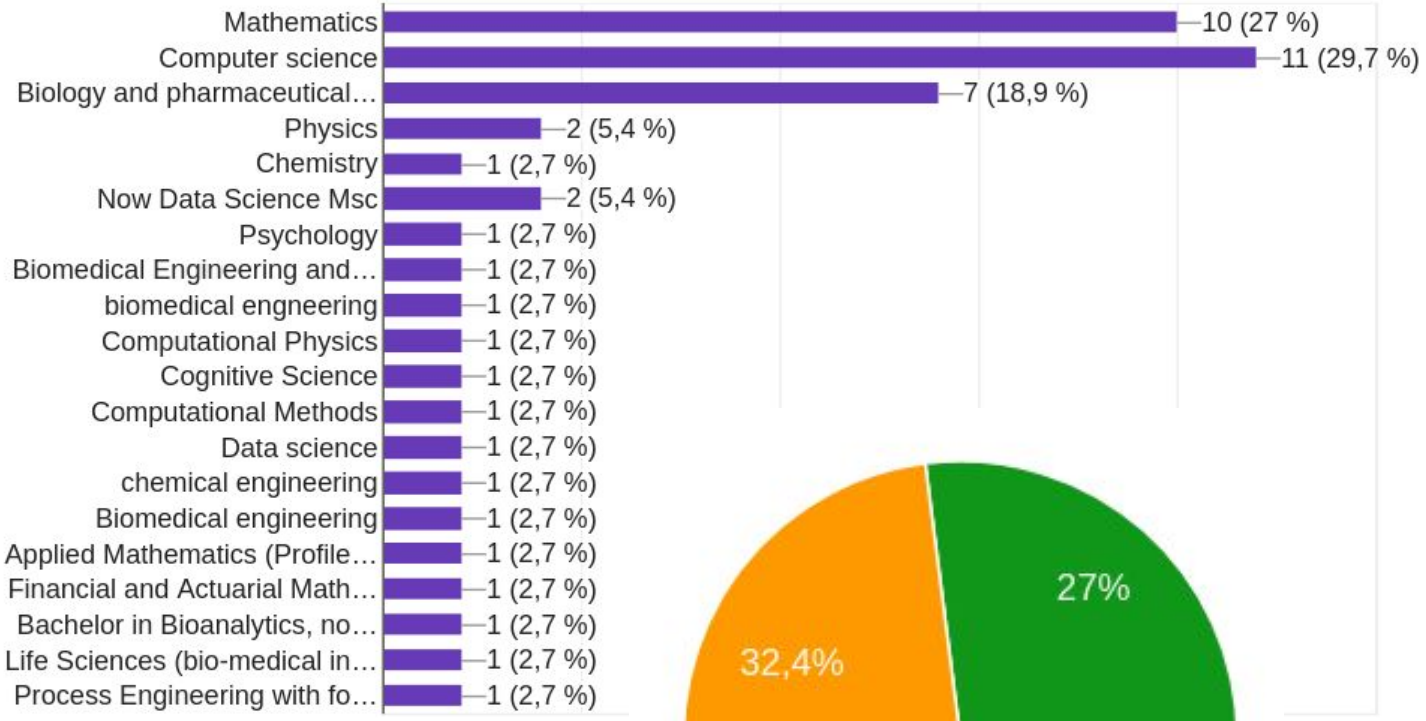
https://xkcd.com/435/



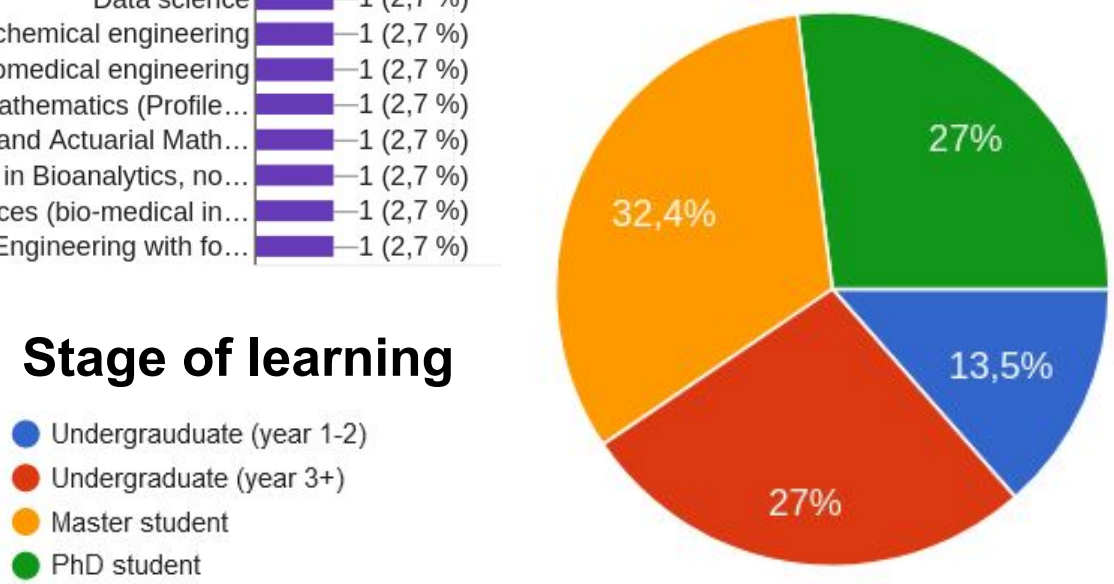
This course aims to bring people together and to promote interdisciplinary research

Our strength: we have a diverse classroom!

Background



Stage of learning



Selected Motivations

- How mathematics can be applied to different research fields.
- I would like to shape my career towards drug discovery
- As biologist, most of the time I work with a trial-and-error approach. I am curious to get to know how mathematics and informatics can guide a more informed approach to drug discovery/design.
- Interesting and different course that gives credits
- State of the art, prevalence in industry
- To learn something new
- learn about advanced computational methods AND their application

Course information for AMIDD 2024

- Lecturer: Jitao David Zhang (jitao-david.zhang@unibas.ch)
- Website: AMIDD.ch
- Thirteen lectures this semester
  - Introduction(1 session)
  - Mechanistic, statistical, and causal models (2 sessions)
  - Molecular level modelling (2 sessions)
  - Omics- & cellular models (2 sessions)
  - Organ- and system models (2 sessions)
  - Population modelling (2 sessions)
  - Invited guest speakers (1 session)
  - A collaboration challenge (1 session)

- Fridays 12:15-14:00
- Slides, exercises, pre-reading and post-reading articles are shared on the course's website http://www.amidd.ch. Unfortunately we do not provide recordings.
- The final note is given by participation including quizzes (30%), offline activities (40%), and a collaboration challenge in the final session (30%). The topic of collaboration challenge will be announced in the last session.
- **Questions**?

I am glad to share my experience in drug discovery, and to learn from you!

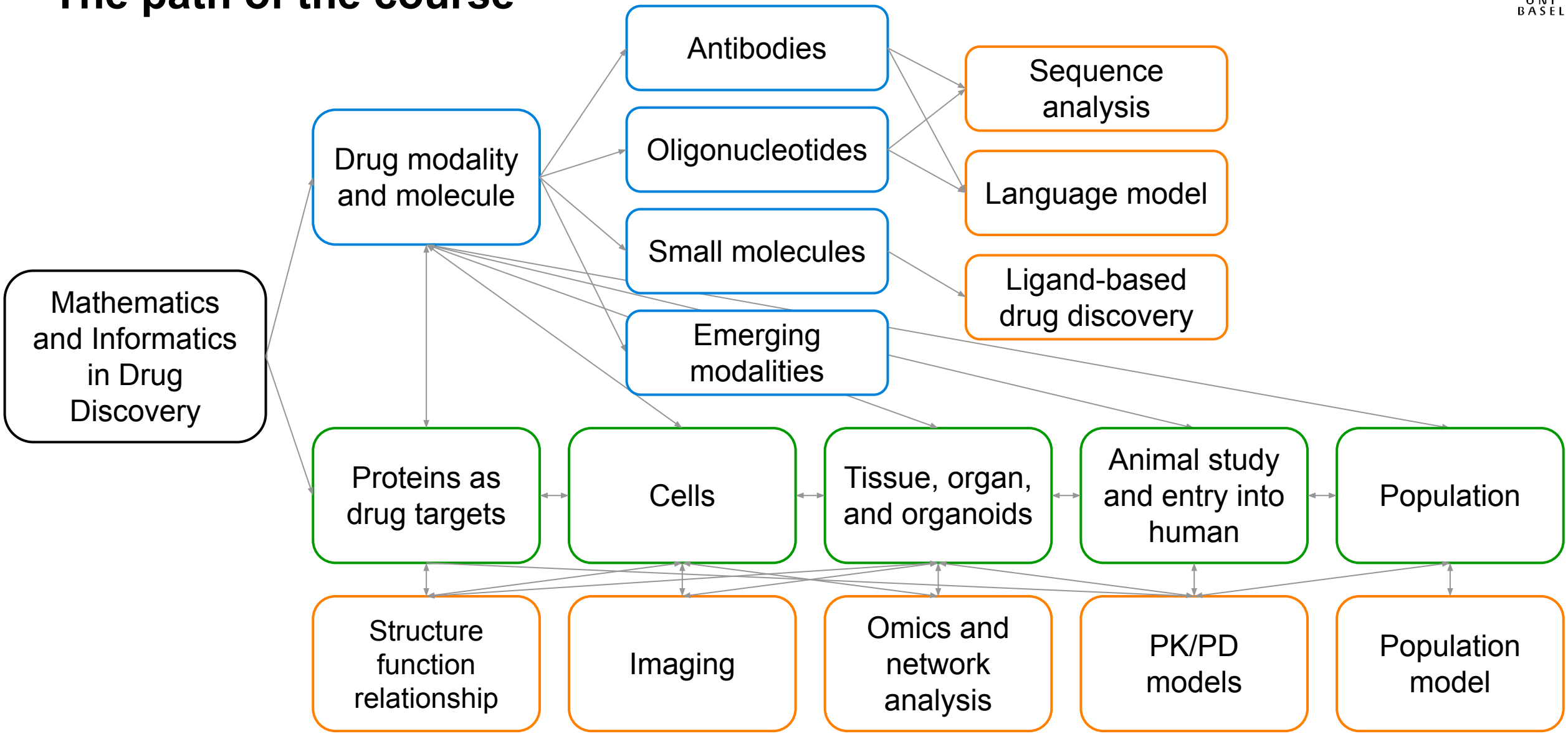
Disclaimer

Teaching is my personal engagement. My opinions and views do not necessarily reflect those by F. Hoffmann-La Roche, my employer.

Please be aware of my biases and limitations.

- I am a computational biologist working in drug discovery, with limited understanding of mathematics, computer science, biology, chemistry, pharmacology, toxicology, and medicine.
- I see my task is to share with you the mathematical concepts and computational approaches used in drug discovery that I find beautiful and useful.
- I look forward to learning from you mathematics and other expertise that I did not know.

The path of the course



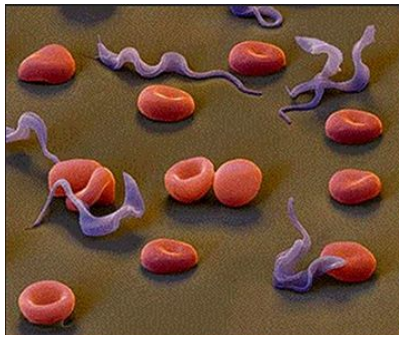
Conclusions and perspectives

- Drug discovery and development is a complex, time-consuming, and labour-intensive process.
- Mathematical models and computational tools integrate data and knowledge collected spanning scales to inform drug discovery and prioritization.
- In the AMIDD course, we will learn some basic concepts and tools we use to model interactions between biological systems and drugs at multiple levels.

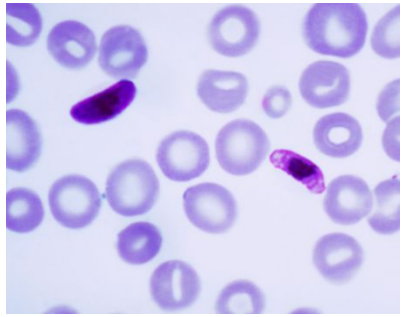
Contact the author

Backup slides

The history of *Homo sapiens* is a history of living with, understanding, and fighting diseases



Trypanosomes



Plasmodium

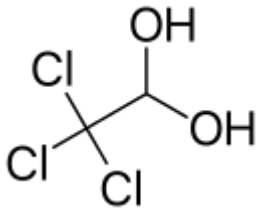
Tropical diseases

~500,000 years ago

A young patient of smallpox,
the first eradicated infectious disease

Hygiene, vaccination, and antibiotics

~250 years ago



Chloral hydrate,
the first synthesized drug

Pharmaceutical drugs

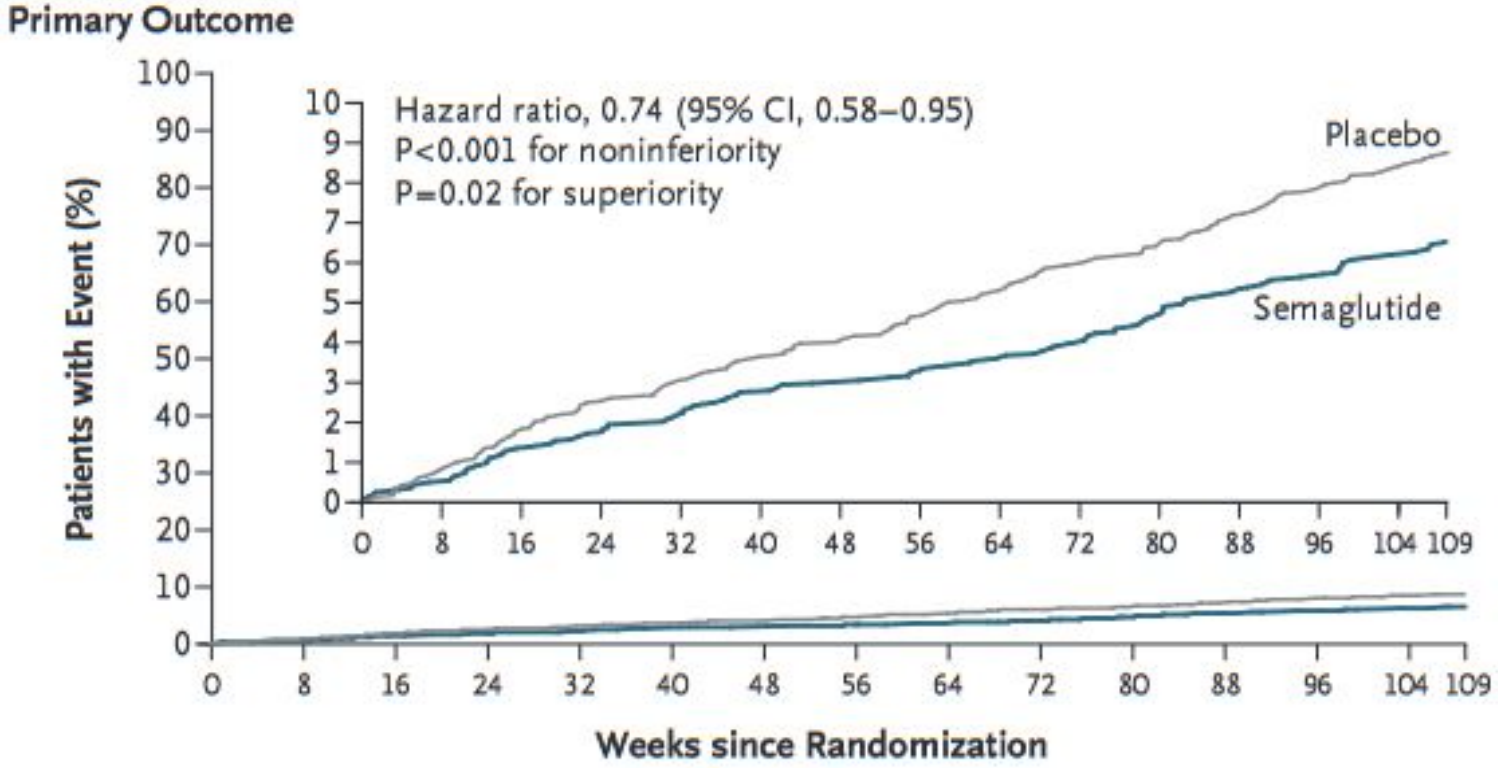
~150 years ago

Nobel prize laureates 2018,
immune checkpoints,
and drugs targeting the pathways

Personalized precise healthcare

~20 years ago

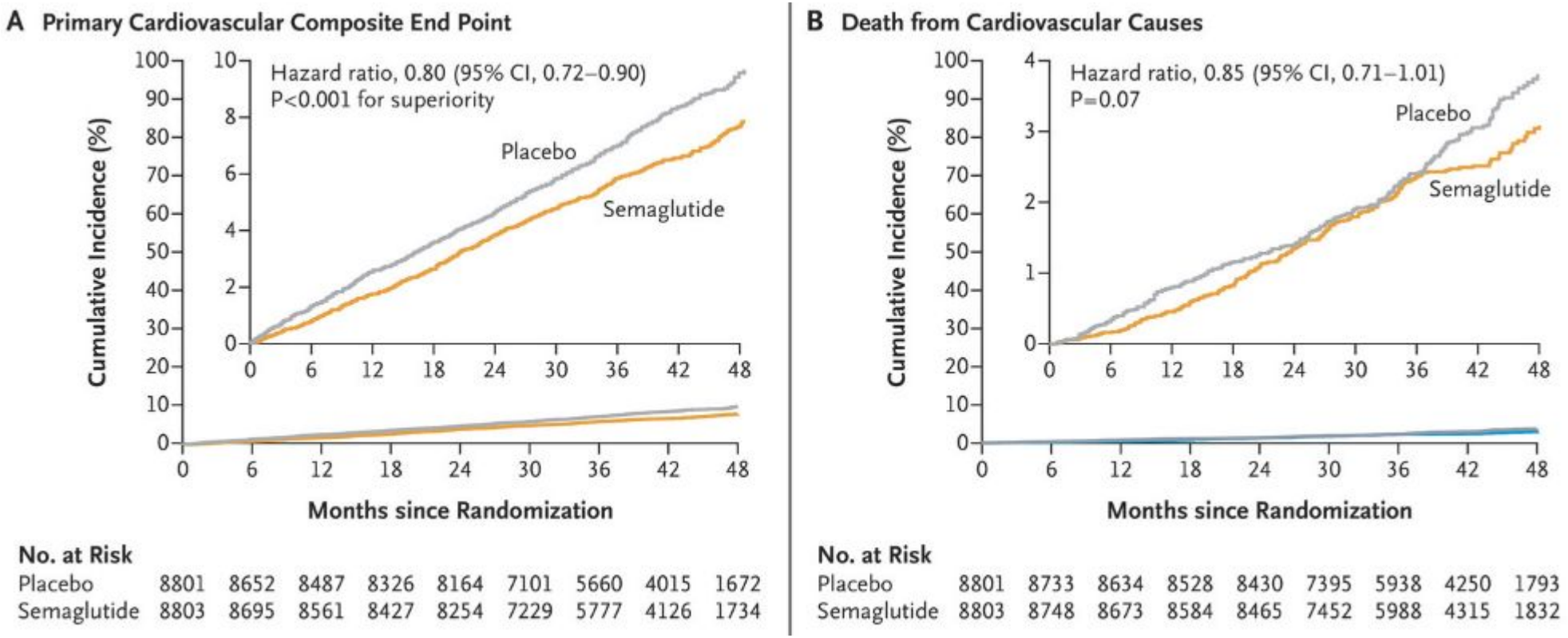
SUSTAIN-6: a non-inferiority clinical trial of semaglutide for cardiovascular safety in patients with type-2 diabetes



No. at Risk							
Placebo	1649	1616	1586	1567	1534	1508	1479
Semaglutide	1648	1619	1601	1584	1568	1543	1524

[Semaglutide and Cardiovascular Outcomes in Patients with Type 2 Diabetes](), Marso *et al.*, 2023

SELECT Trial: Semaglutide Effects on Cardiovascular Outcomes in People with Overweight or Obesity



Semaglutide and Cardiovascular Outcomes in Obesity without Diabetes, Lincoff *et al.*, 2023