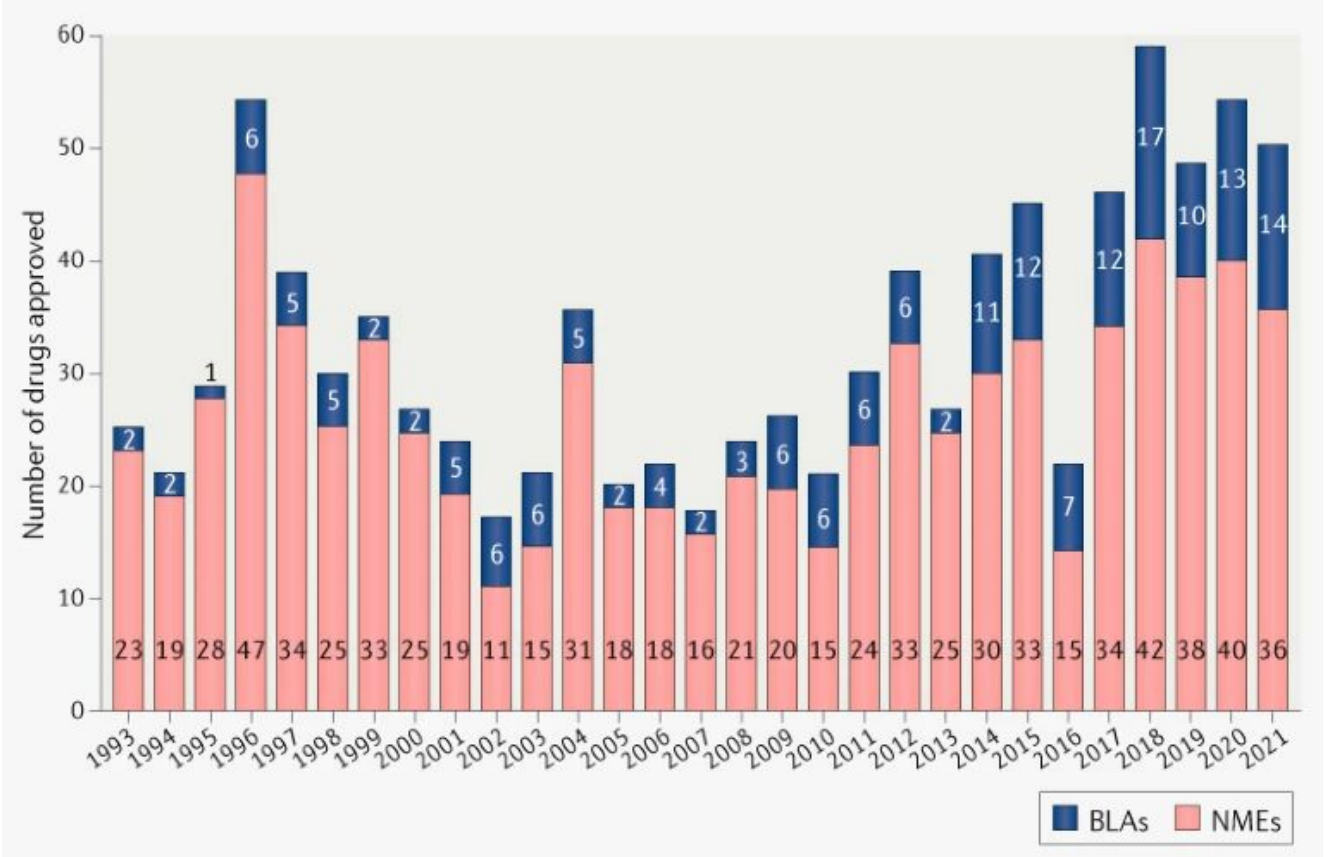
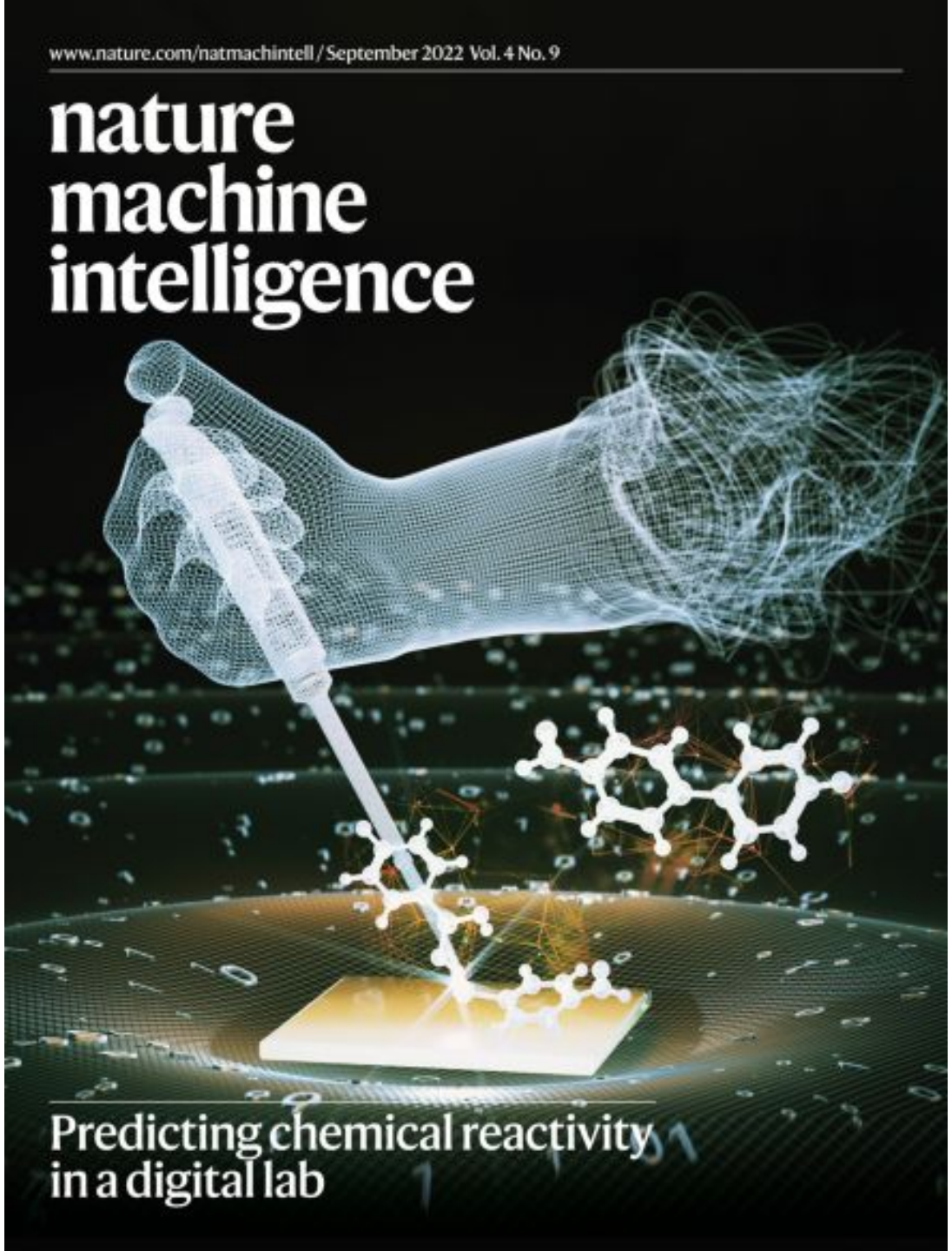
Applied Mathematics and Informatics In Drug Discovery (2022)



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



Research

Accelerated rational PROTAC design via deep learning and molecular simulations

Shuangjia Zheng, Youhai Tan, Zhenyu Wang, Chengtao Li, Zhiqing Zhang *et al.*

Article | 15 September 2022 *Nature Machine Intelligence* 4 | doi:10.1038/s42256-022-00527-y

Deep learning-based robust positioning for all-weather autonomous driving

Yasin Almalioglu, Mehmet Turan, Niki Trigoni & Andrew Markham

Article | 08 September 2022 *Nature Machine Intelligence* 4 | doi:10.1038/s42256-022-00520-5

Deep neural networks with controlled variable selection for the identification of putative causal genetic variants

Peyman H. Kassani, Fred Lu, Yann Le Guen, Michael E. Belloy & Zihuai He

Article | 15 September 2022 *Nature Machine Intelligence* 4 | doi:10.1038/s42256-022-00525-0

A generalized-template-based graph neural network for accurate organic reactivity prediction

Shuan Chen & Yousung Jung

Article | 15 September 2022 *Nature Machine Intelligence* 4 | doi:10.1038/s42256-022-00526-z

Recovery of continuous 3D refractive index maps from discrete intensity-only measurements using neural fields

Renhao Liu, Yu Sun, Jiabei Zhu, Lei Tian & Ulugbek S. Kamilov

Article | 16 September 2022 *Nature Machine Intelligence* 4 | doi:10.1038/s42256-022-00530-3

Interpretable meta-score for model performance

Alicja Gosiewska, Katarzyna Woźnica & Przemysław Biecek

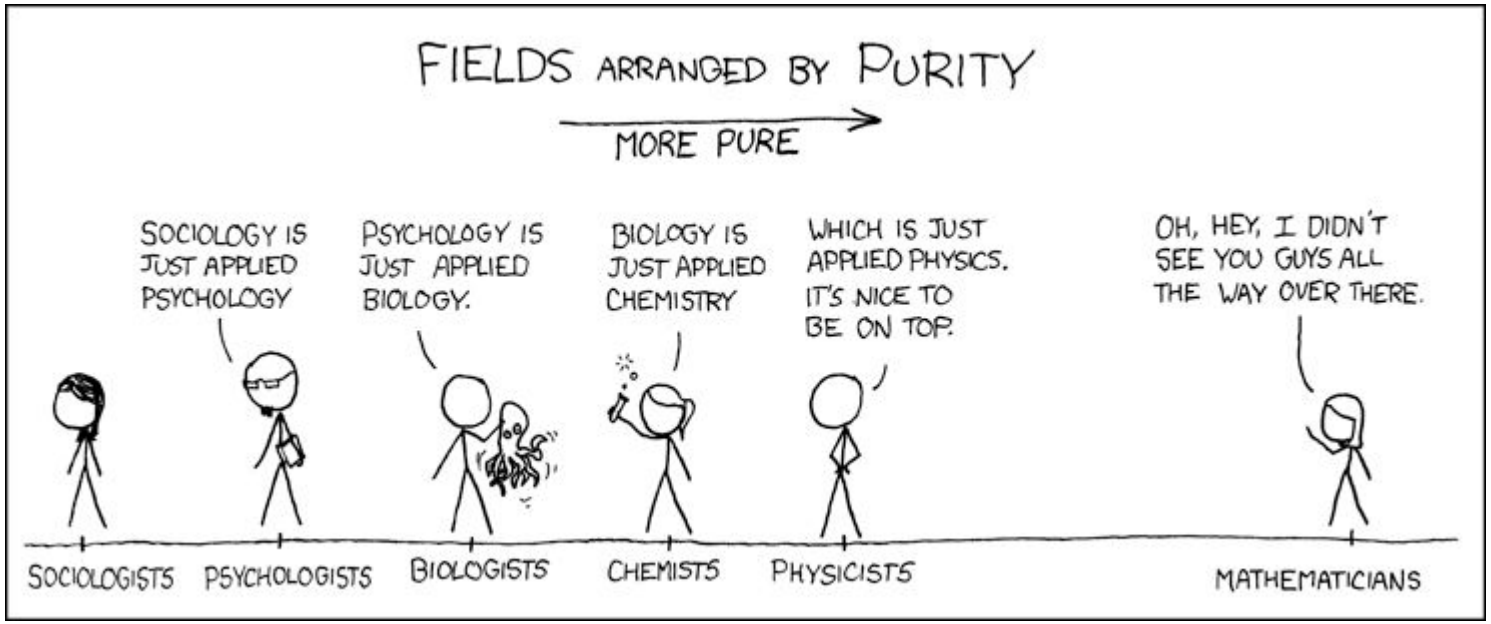
Article | 22 September 2022 *Nature Machine Intelligence* 4 | doi:10.1038/s42256-022-00531-2

Today's goals

- **Introduction to the course**
- **Why mathematics and informatics matter for drug discovery?**
- **Two views of the drug discovery and development process**
  - The linear view
  - The multiscale modelling view

Purity

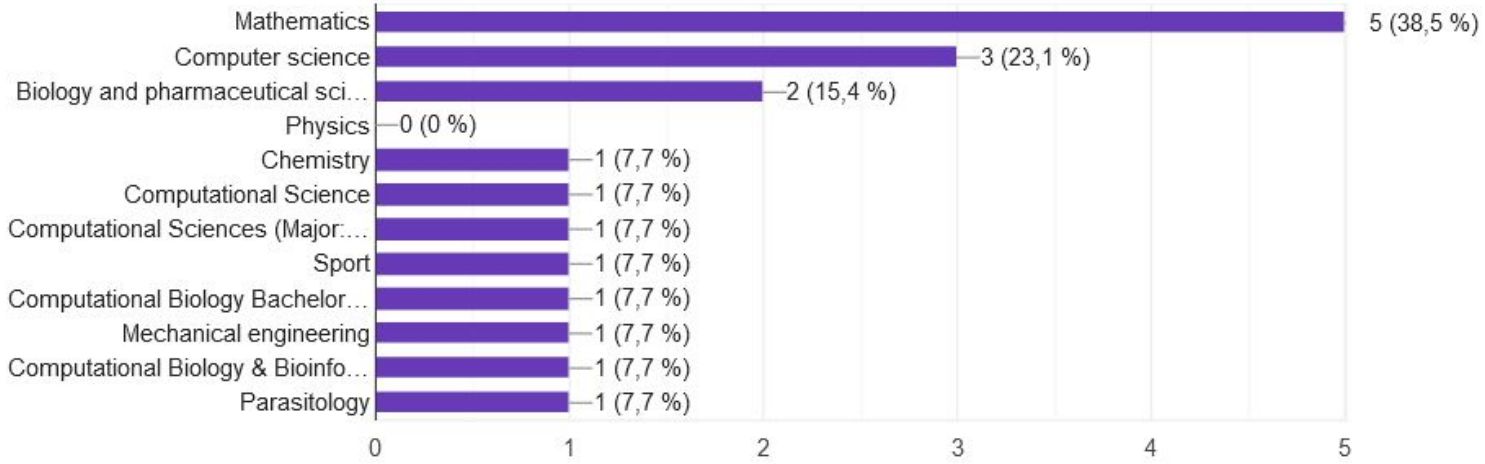
https://xkcd.com/435/



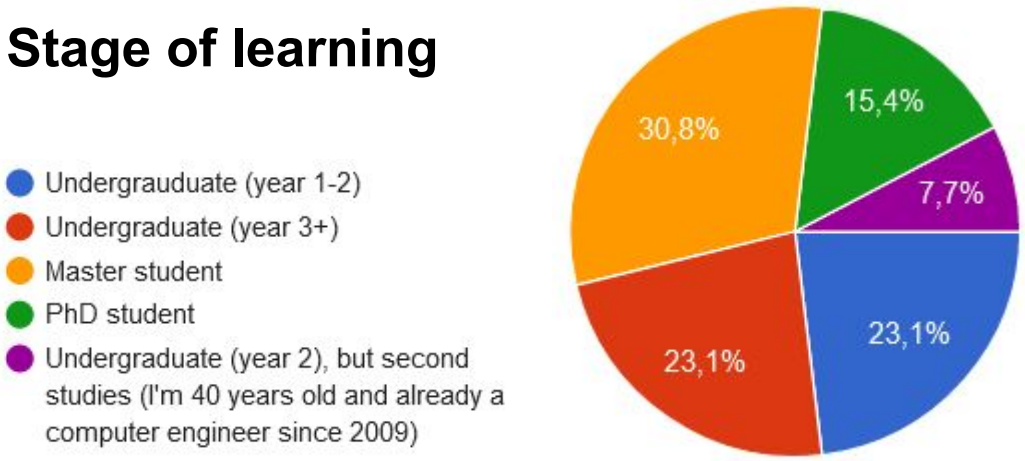
This course aims to bring people together and to promote interdisciplinary research

Our strength: we have a diverse classroom!

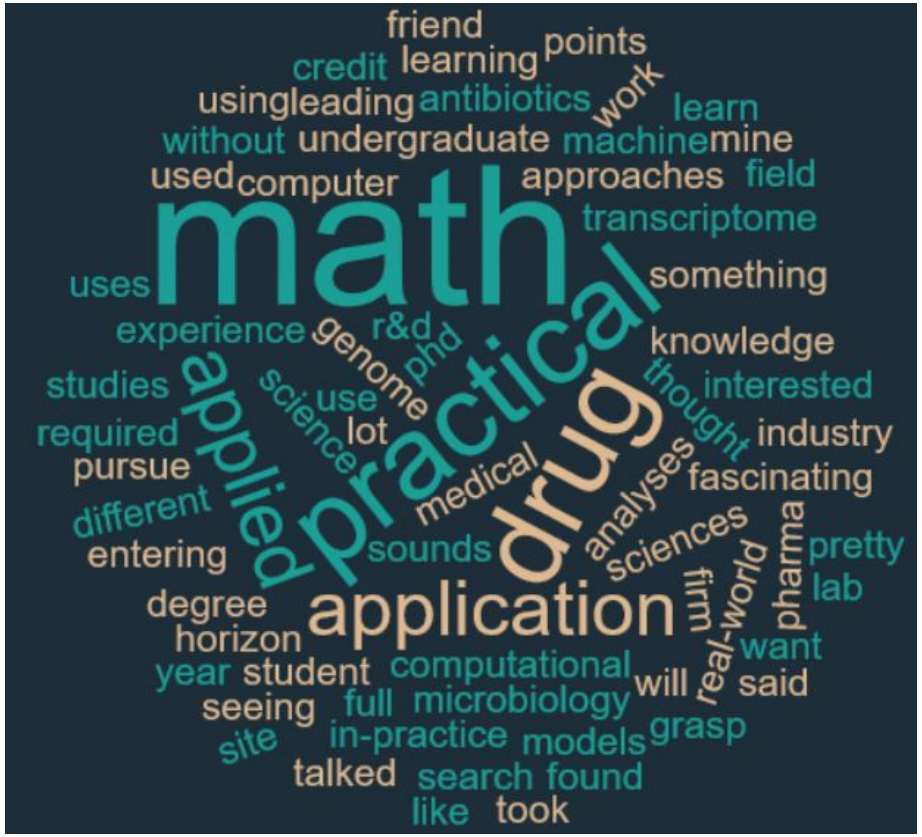
Background



Stage of learning



Motivations



Course information

- Lecturer: Jitao David Zhang
  - jitao-david.zhang@unibas.ch
- Website: AMIDD.ch
- Thirteen lectures this semester
  - Introduction to drug discovery (1 session)
  - Molecular level modelling (2 sessions)
  - Omics- and cellular modelling (2 sessions)
  - Organ- and system modelling (2 sessions)
  - Population modelling and reverse translation (2 sessions)
  - *Dies Academicus - Ask Me Anything (optional)*
  - Invited guest speakers (1 sessions)
  - Near-end-term presentations (2 sessions)

- 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.
- No exercise hours. One-to-one virtual sessions are possible upon request and reservation.
- The final note is given by participation (40%), presentation (30%), and project work (30%).
- The project work will be about concepts that we learned together and their applications in practice. Details will follow.
- **Questions**?

I am glad to share my expertise 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.

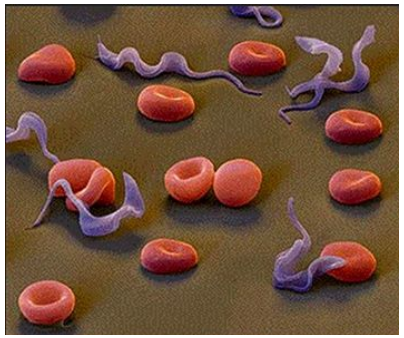
Please be aware of my biases and limitations.

- I am a computational biologist working in drug discovery, with limited understanding of mathematics, computer science, biology, 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.

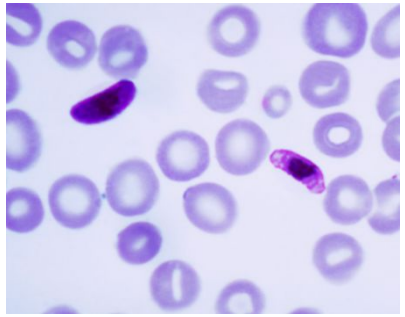
Why applied mathematics and informatics in drug discovery, why now?

- Now is the best time in human history to fight diseases
- Applied mathematics and informatics approaches are indispensable to modern drug discovery
- Applied mathematics and informatics will join interdisciplinary efforts to transform drug discovery in the coming decades

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



Trypanosomes



Plasmodium

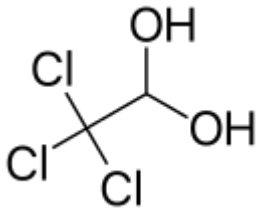
Tropical diseases

~500,000 years ago

A young patient of smallpox,
the first eradicated infectious disease

Hygiene, vaccination, and antibiotics

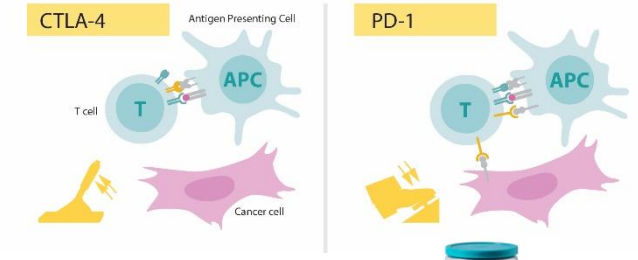
~250 years ago



Chloral hydrate,
the first synthesized drug

Pharmaceutical drugs

~150 years ago

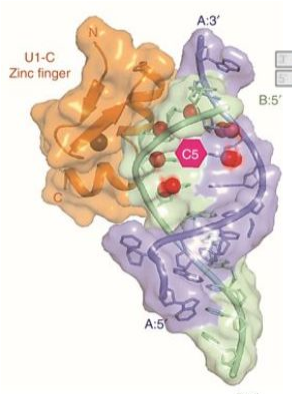


Nobel prize laureates 2018,
immune checkpoints,
and drugs targeting the pathways

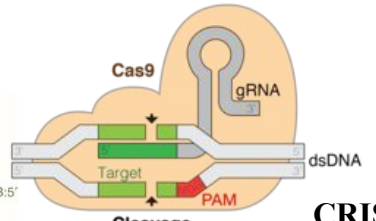
Personalized precise healthcare

~20 years ago

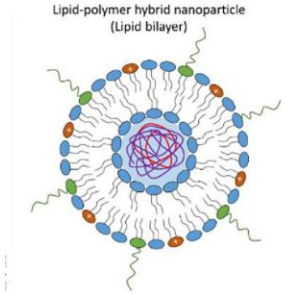
Now is the best time in human history to fight diseases



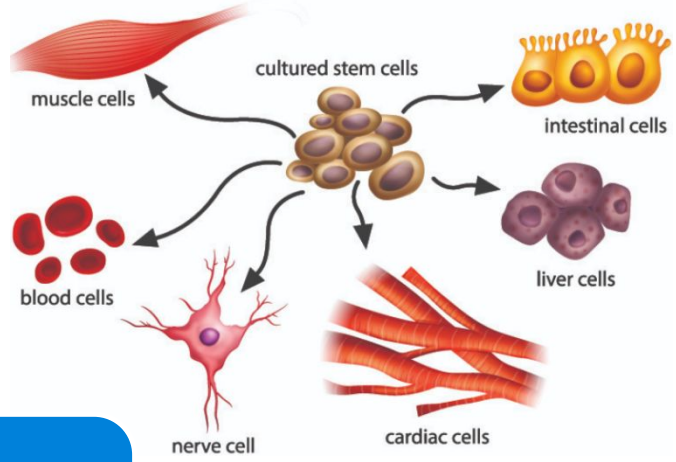
Risdiplam, a SMN2 splicing modifier



CRISPR-CAS9 gene editing system



mRNA delivery with nano lipid particles



Stem cells

More biological, chemical, and medicinal knowledge

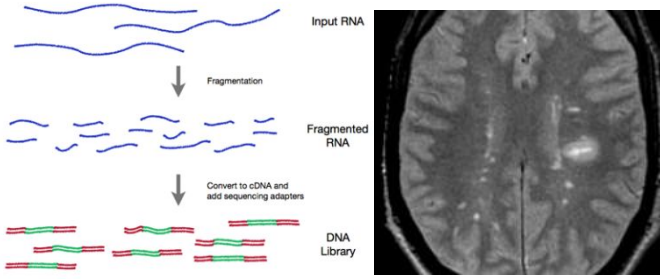
New disease-modelling systems

Digitalization of molecular mechanisms in living organisms

Better algorithms, models, and more computing resources

New therapeutic modalities

Comprehensive Sensing

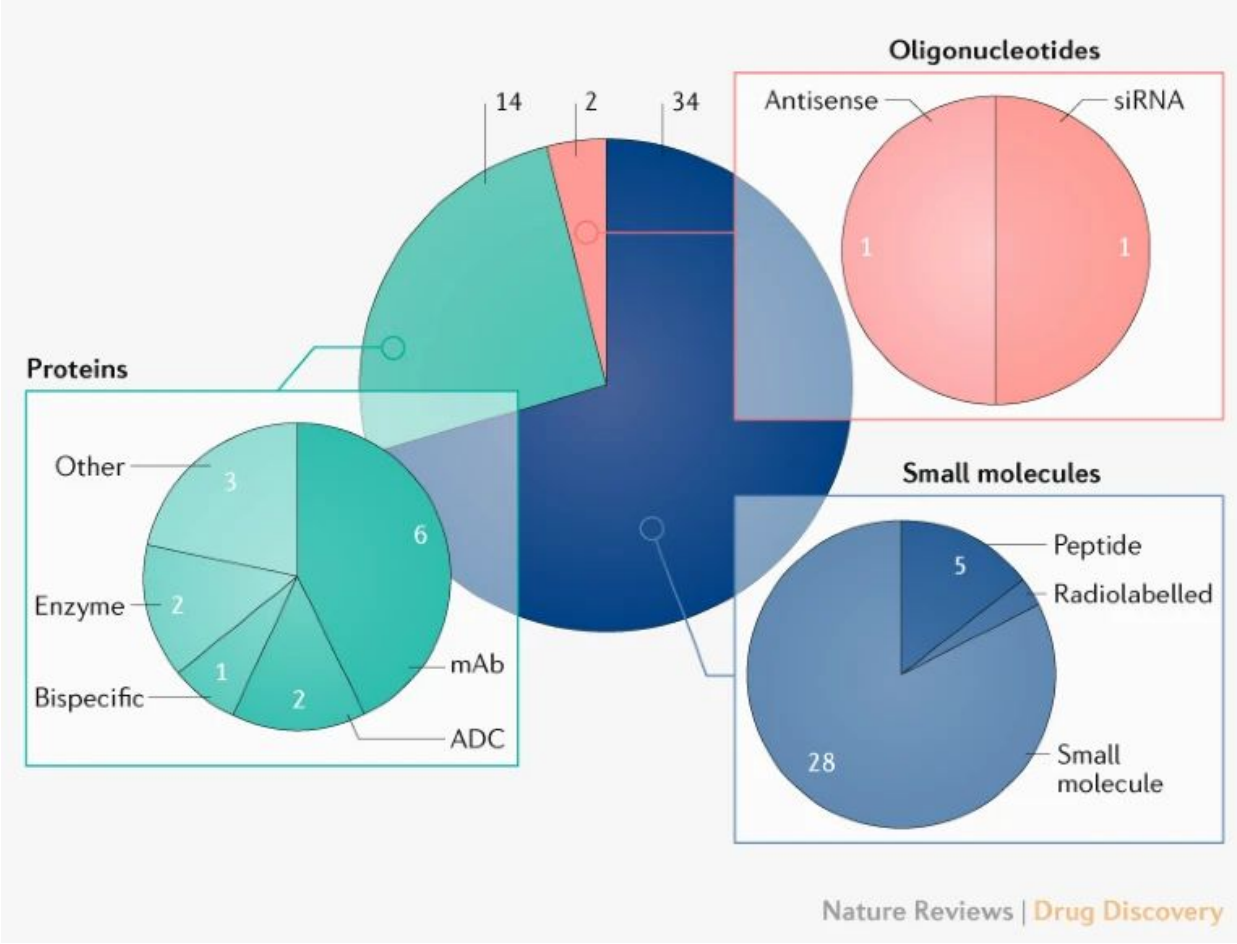


Single-cell biology, multi-modal omics profiling, and imaging

Novel drugs approved by the FDA's Center for Drug Evaluation and Research (CDER) in 2021

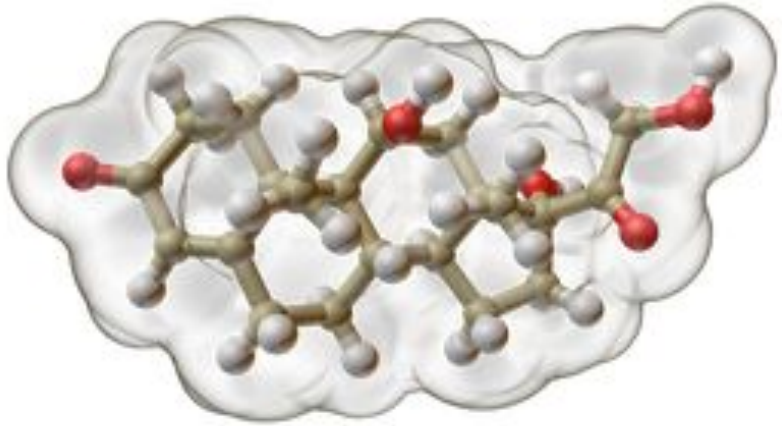
Top three modalities by approval in 2021

- Small molecules (molecular weight under 900 daltons)
- Proteins
  - mAb: monoclonal antibody
  - ADC: antibody-drug conjugate
- Oligonucleotides

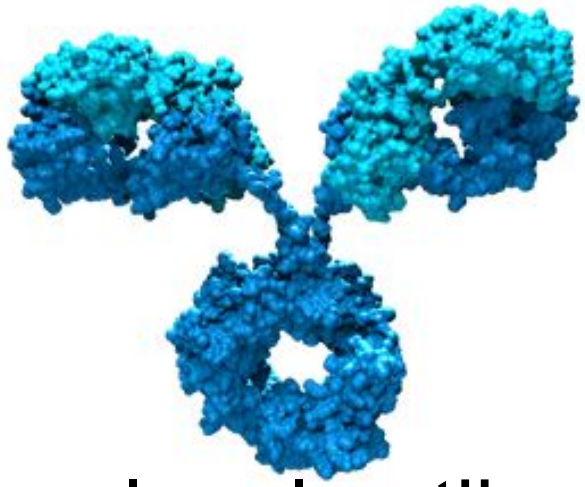


2021 FDA drug approvals, Asher Mullard, Nature Reviews Drug Discovery, https://www.nature.com/articles/d41573-022-00001-9

A zoo of modalities

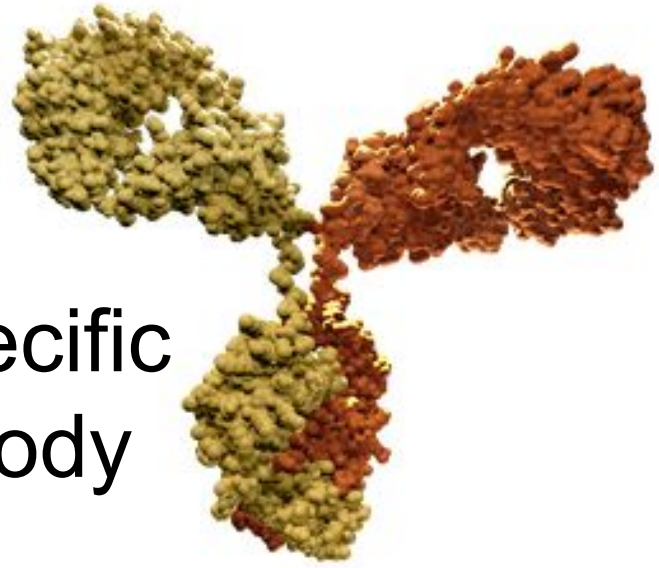


Small molecule

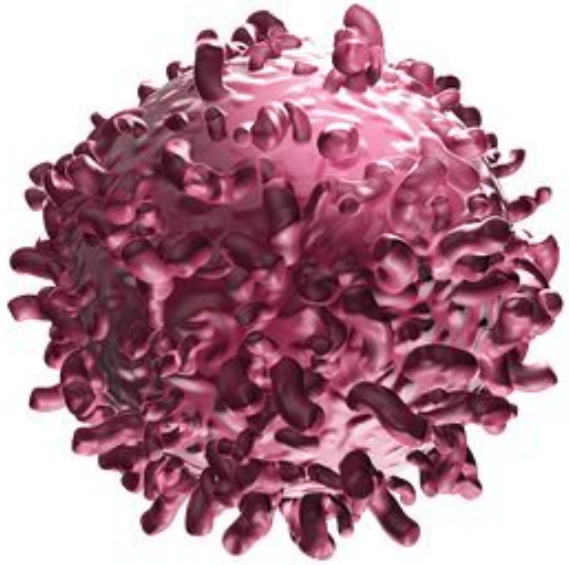


Monoclonal antibody

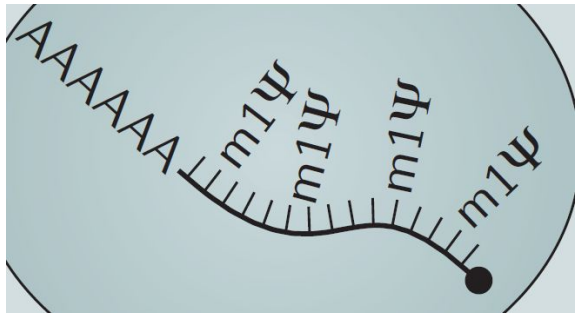
Oligonucleotides



Bispecific antibody

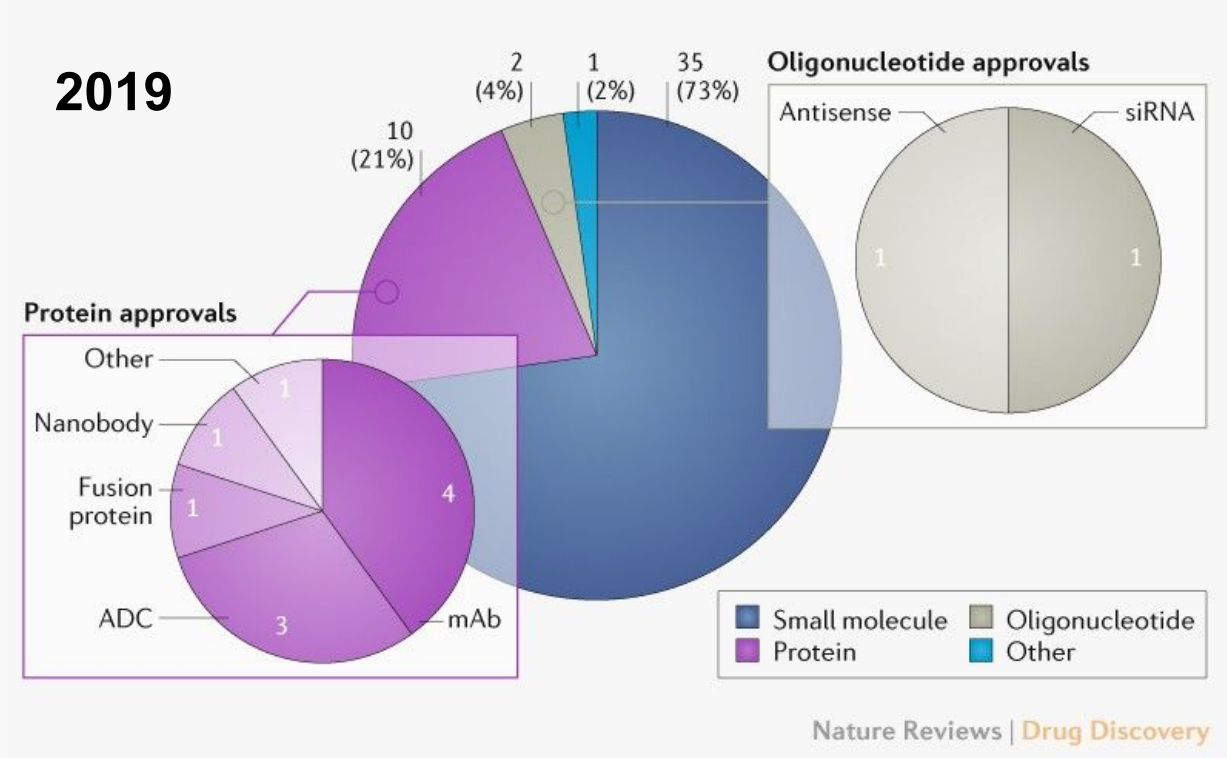


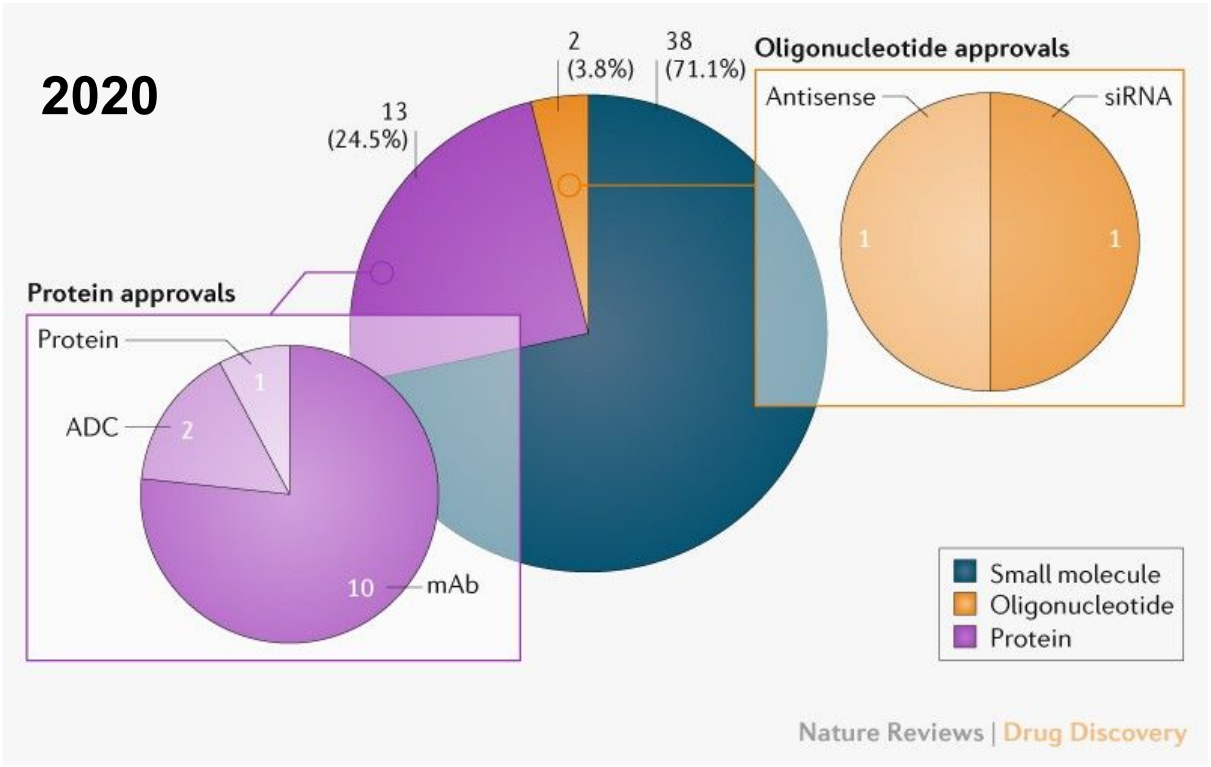
Chimeric Antigen Receptor (CAR) T-cells



mRNA vaccines

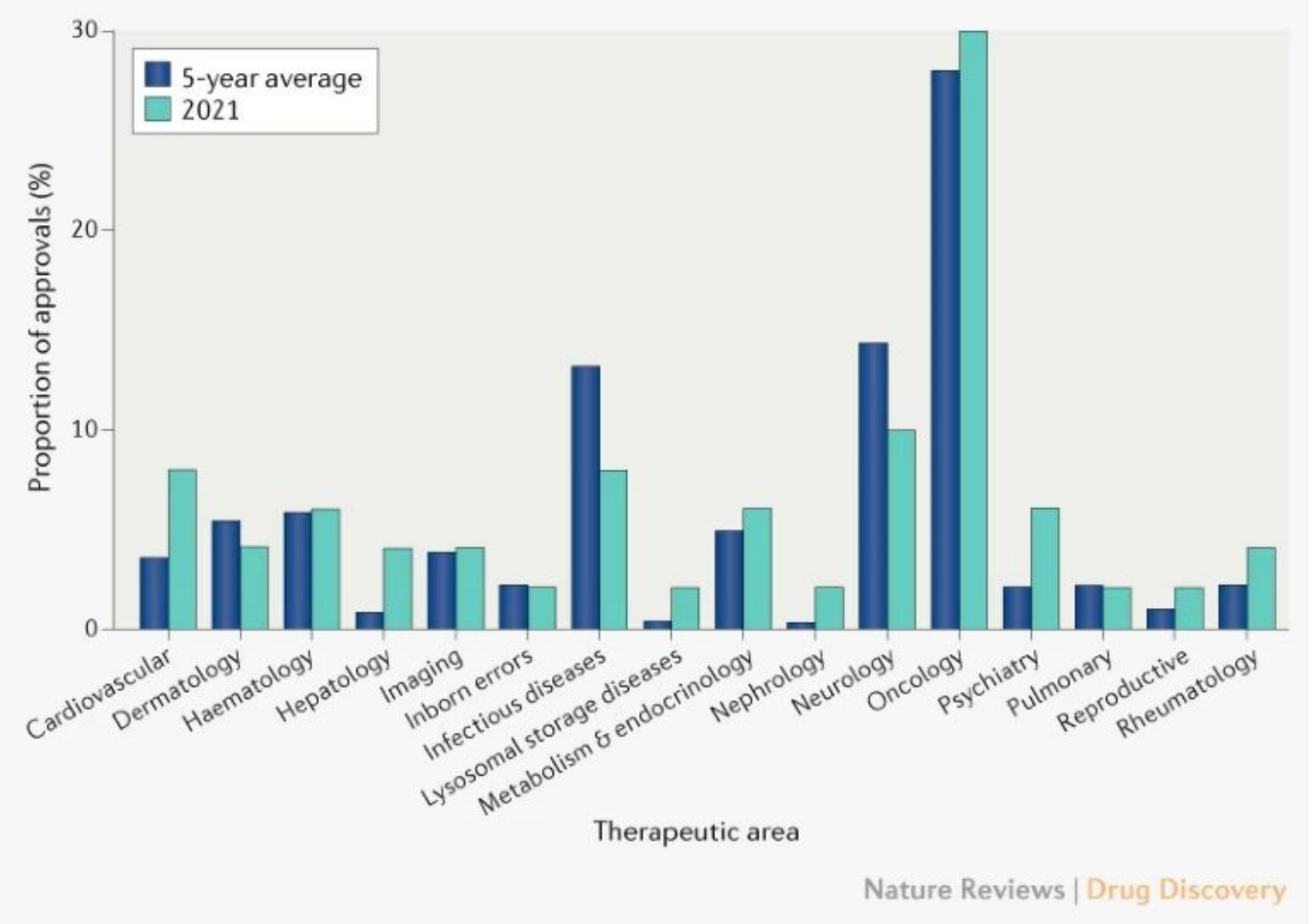
Relative contributions of modalities remain constant in the past three years





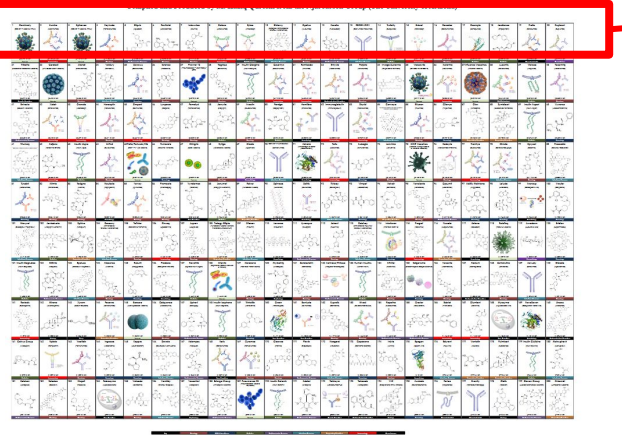
2019 FDA drug approvals, Asher Mullard, Nature Reviews Drug Discovery, https://www.nature.com/articles/d41573-020-00001-7 mAb: monoclonal antibodies; ADC: antibody-drug conjugate.
2020 FDA drug approvals, Asher Mullard, Nature Reviews Drug Discovery, https://www.nature.com/articles/d41573-021-00002-0

New drug approvals vary between disease areas



1 Comirnaty (COVID-19 Vaccine, mRNA) $36.781 Billion Vaccine	2 Humira (Adalimumab) $21.874 Billion Immunology	3 Spikevax (COVID-19 Vaccine, mRNA) $17.675 Billion Vaccine	4 Keytruda (Pembrolizumab) $17.186 Billion Oncology	5 Eliquis (Apixaban) $16.732 Billion Cardiovascular Diseases	6 Revlimid (Lenalidomide) $12.898 Billion Oncology	7 Imbruvica (Ibrutinib) $9.777 Billion Oncology	8 Stelara (Ustekinumab) $9.54 Billion Immunology
9 Eylea (Aflibercept) $9.005 Billion Ophthalmology	10 Biktarvy (Bictegravir/Emtricitabine/ Tenofovir Alafenamide) $8.624 Billion Infectious Diseases	11 Opdivo (Nivolumab) $8.431 Billion Oncology	12 Xarelto (Rivaroxaban) $7.651 Billion Cardiovascular Diseases	13 REGEN-COV (Casirivimab/Imdevimab) $7.582 Billion Infectious Diseases	14 Trulicity (Dulaglutide) $6.757 Billion Diabetes	15 Enbrel (Etanercept) $6.211 Billion Immunology	16 Darzalex (Daratumumab) $6.023 Billion Oncology

Top 200 Pharmaceuticals by Retail Sales in 2021



Top pharmaceuticals by retail sails in 2021

Poster compiled by the Jon Njardarson group/ University of Arizona. Source: https://njardarson.lab.arizona.edu, accessed on 22.09.2022. Citation: J. Chem. Ed. 2010, 87, 1348

Questions:

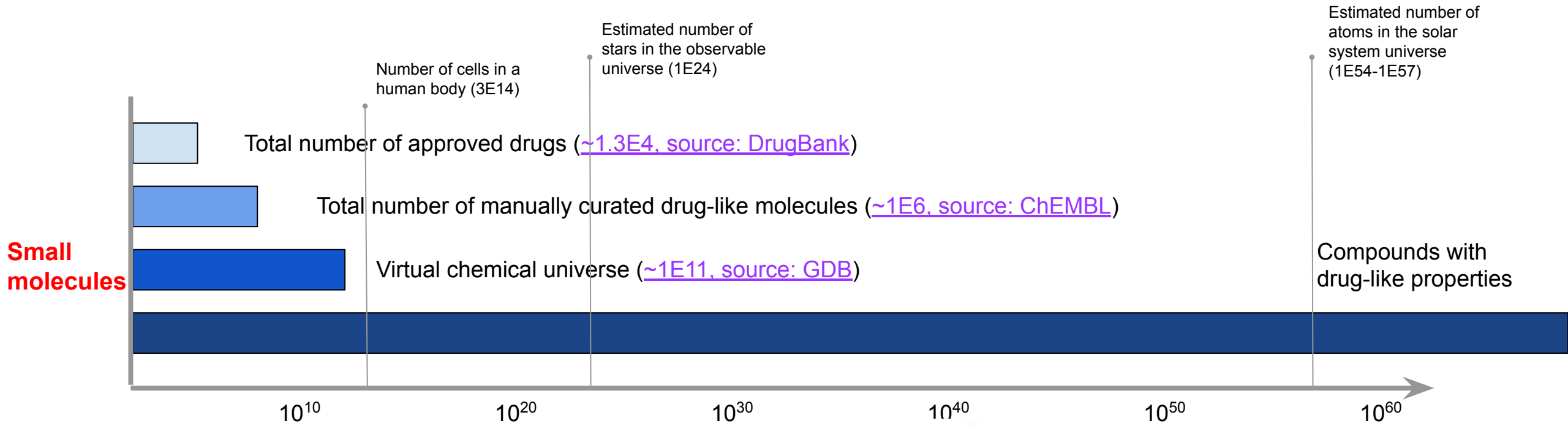
1. How many are small molecules, proteins, and oligonucleotides each? What other modalities are there?
2. What patterns do you observe? Do you have explanations for these patterns?

How Do You Make A Drug?

It sounds simple, but…

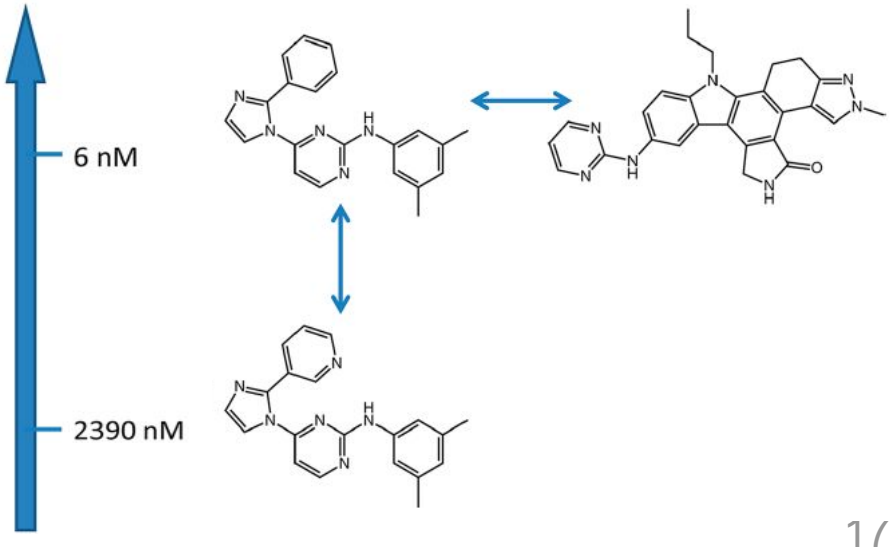
Source and copyright: roche.com, assessed on 1.2.2019

Why drug *discovery*?



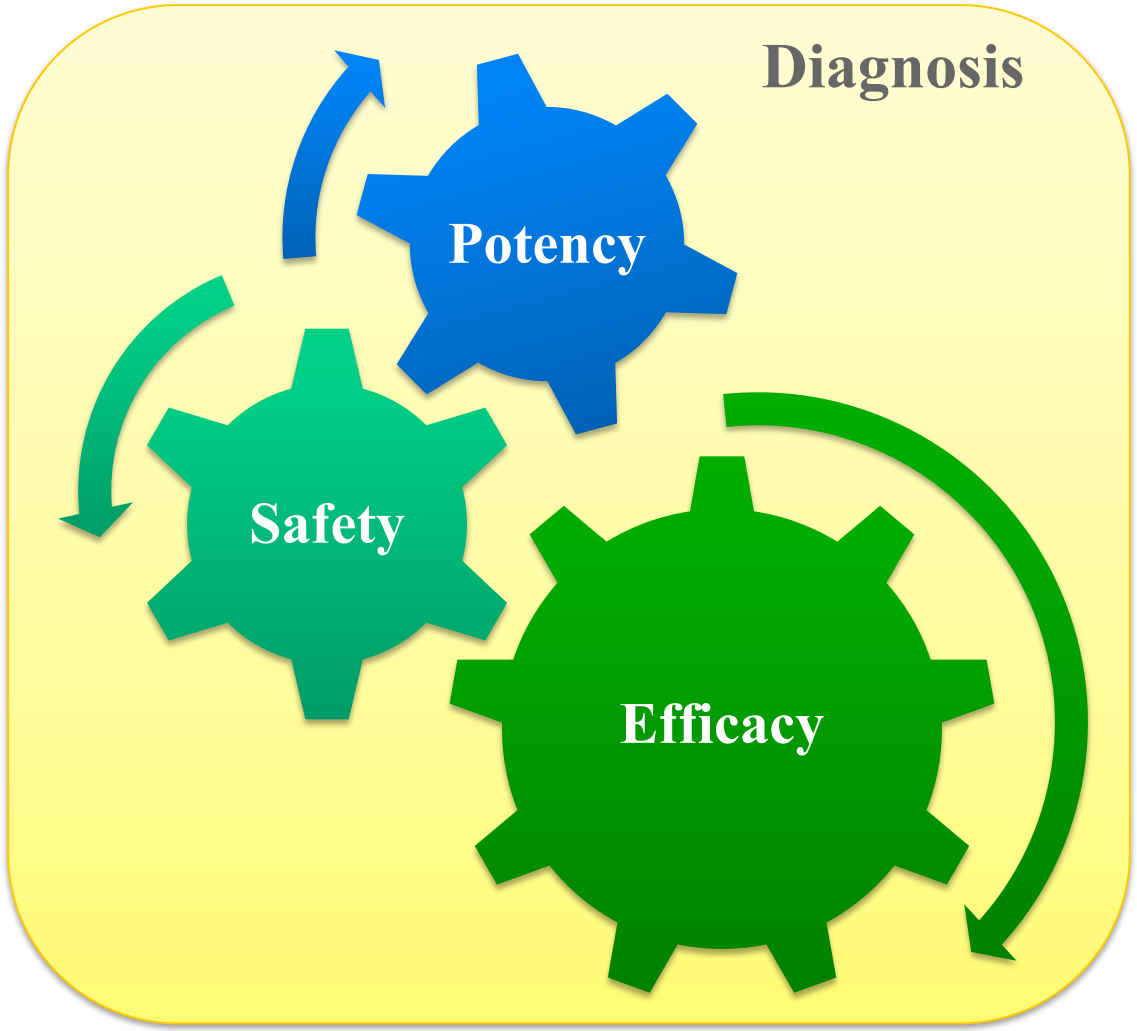
Human protein interactome (BioGrid)

Experiment Type	Raw Interactions	Non-Redundant Interactions	Unique Genes	Unique Publications
PHYSICAL	631,623	480,099	25,528	32,051
GENETIC	9,122	8,996	3,732	349
COMBINED	640,745	488,387	25,941	32,191



Even if we understood everything, the search space of drug hunting is huge

Prerequisites to make a good drug that works

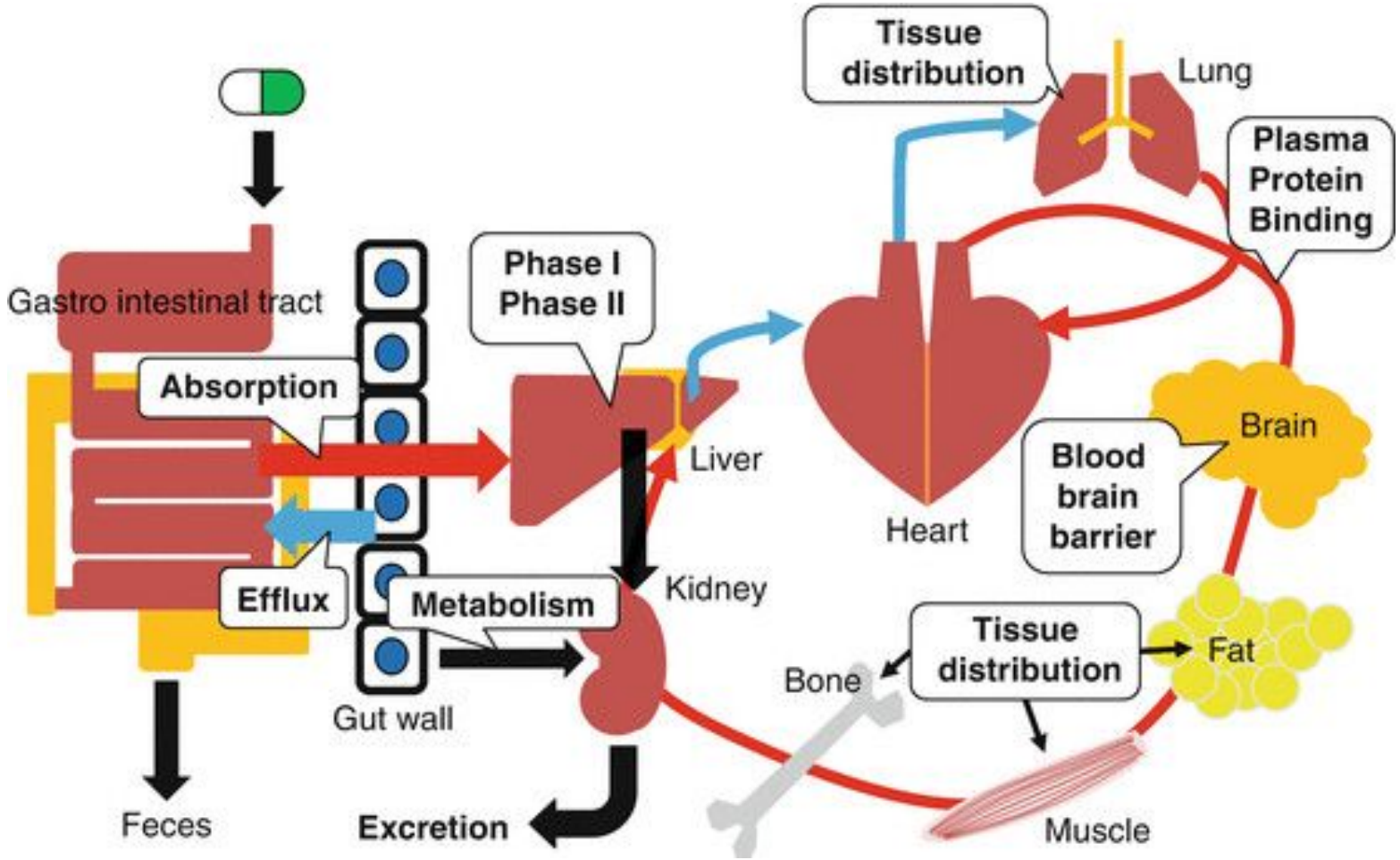


- **Potency**
- **Safety**
- **Efficacy**
- **Diagnosis**: doctors' judgement + *biomarkers*
  - *Biomarkers* are informative features derived from measurements of patient or patient material, *e.g.* blood chemistry, genetic make-up, imaging, *etc.*
- Other criteria: commercial rationale, development ability, intellectual property, *etc.*

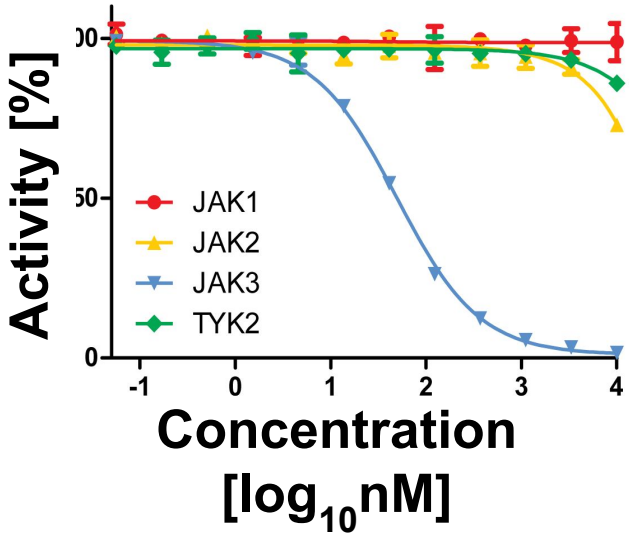
Success in drug discovery is determined by potent, safe, efficacious drugs and accurate diagnosis

Factors that affect efficacy and safety profiles

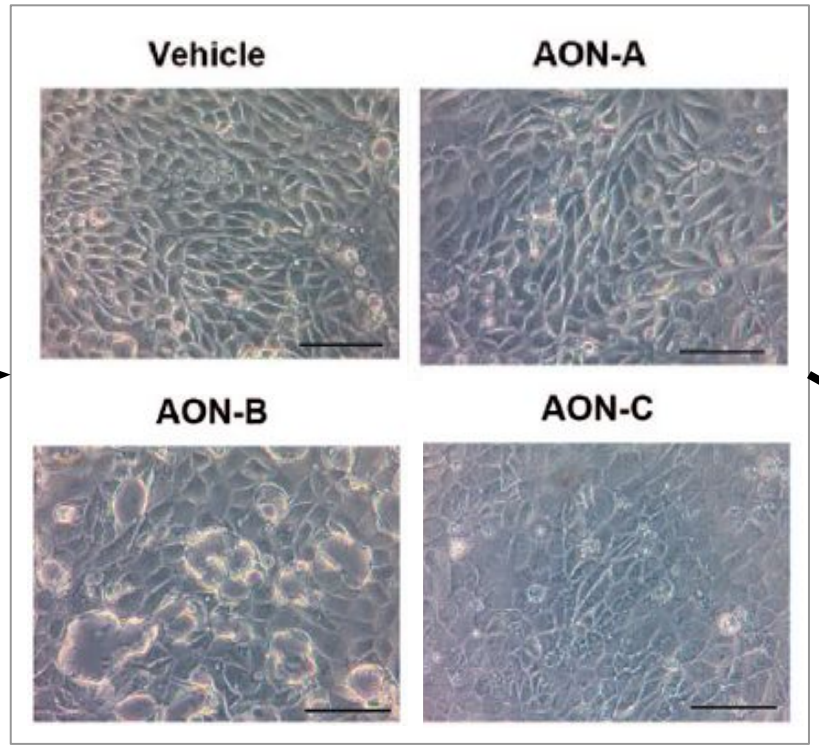
- Absorption
- Distribution
- **Pharmacology**
- **Toxicology**
- Metabolism
- Excretion



Classical workflow of efficacy and toxicity assessment

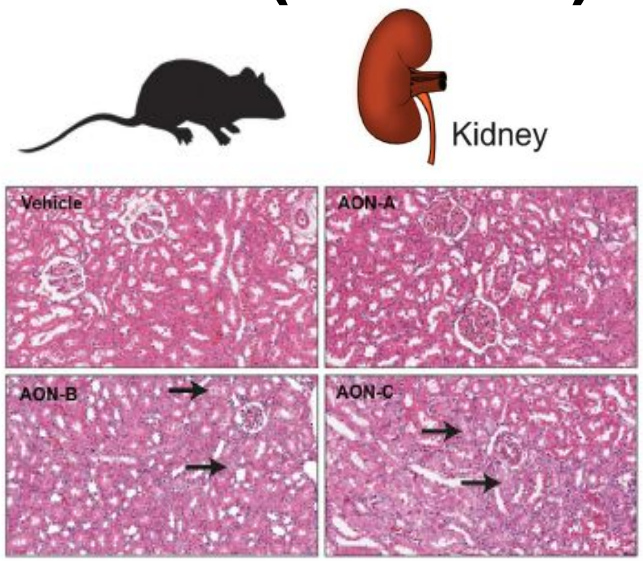


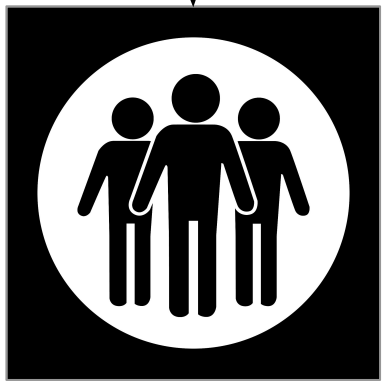
Biochemical & biophylical assays



Cellular assays (*in vitro*)

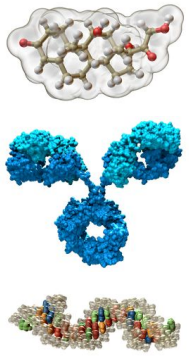
Animal experiments (*in vivo*)

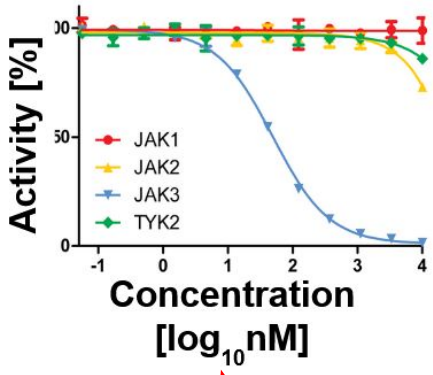


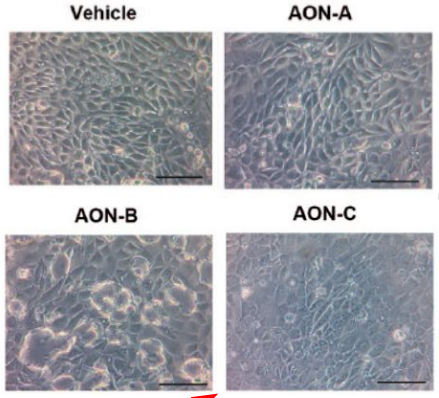


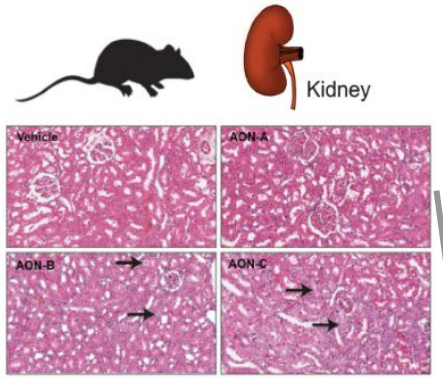
Clinical trials

Computational methods empower efficacy and toxicity assessment



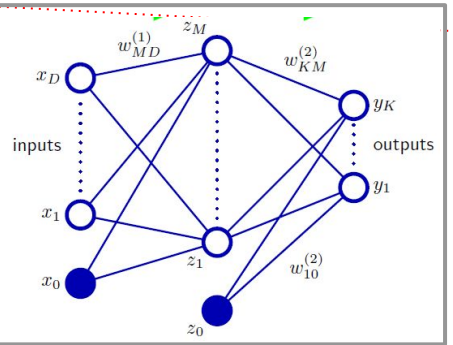






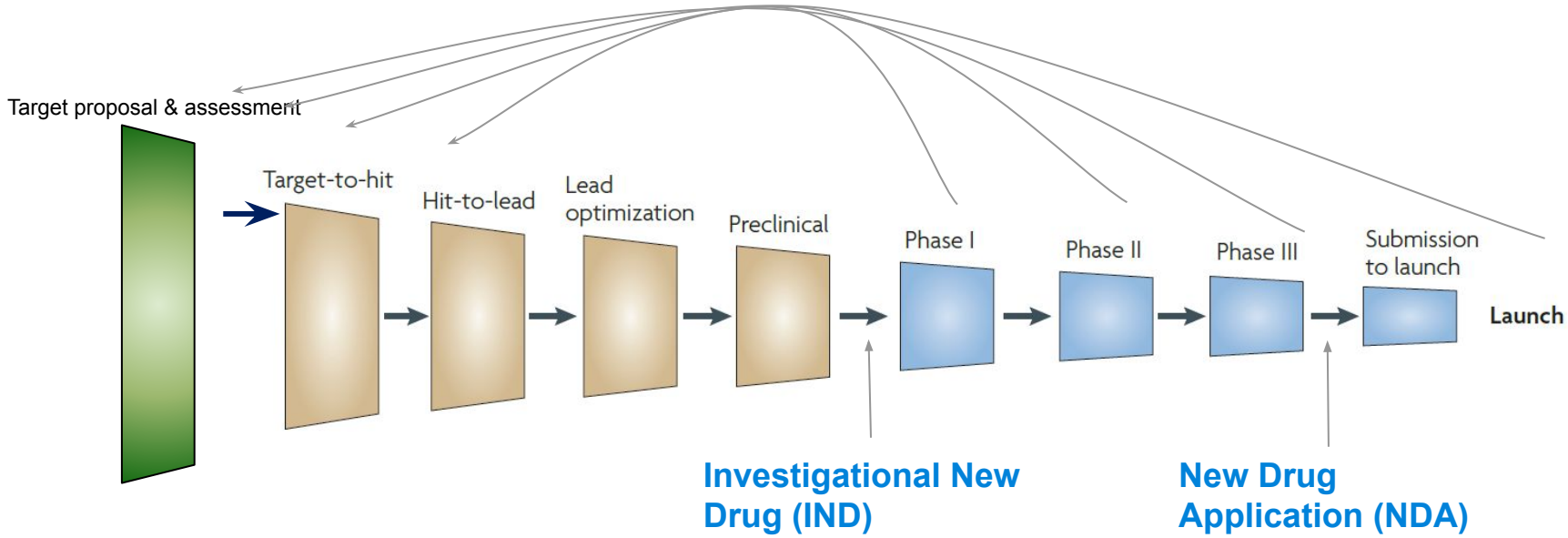
High-throughput technologies (omics, microscopy, etc.)

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



Mechanistic, causal, and statistical models

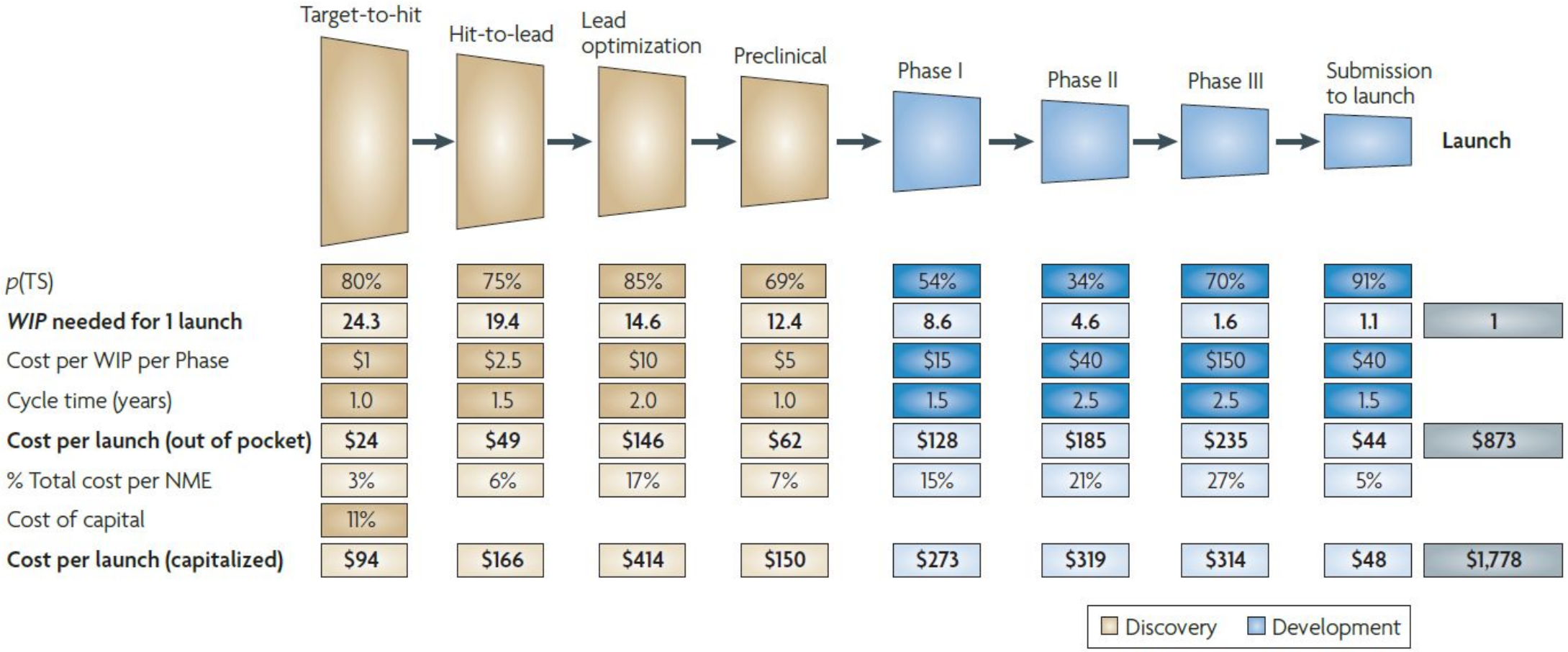
The linear view of drug discovery



Acronym	TI/TA/TV	Screening	LI	LO
Description	Target identification/ target assessment/ target validation		Lead identification	Lead optimization

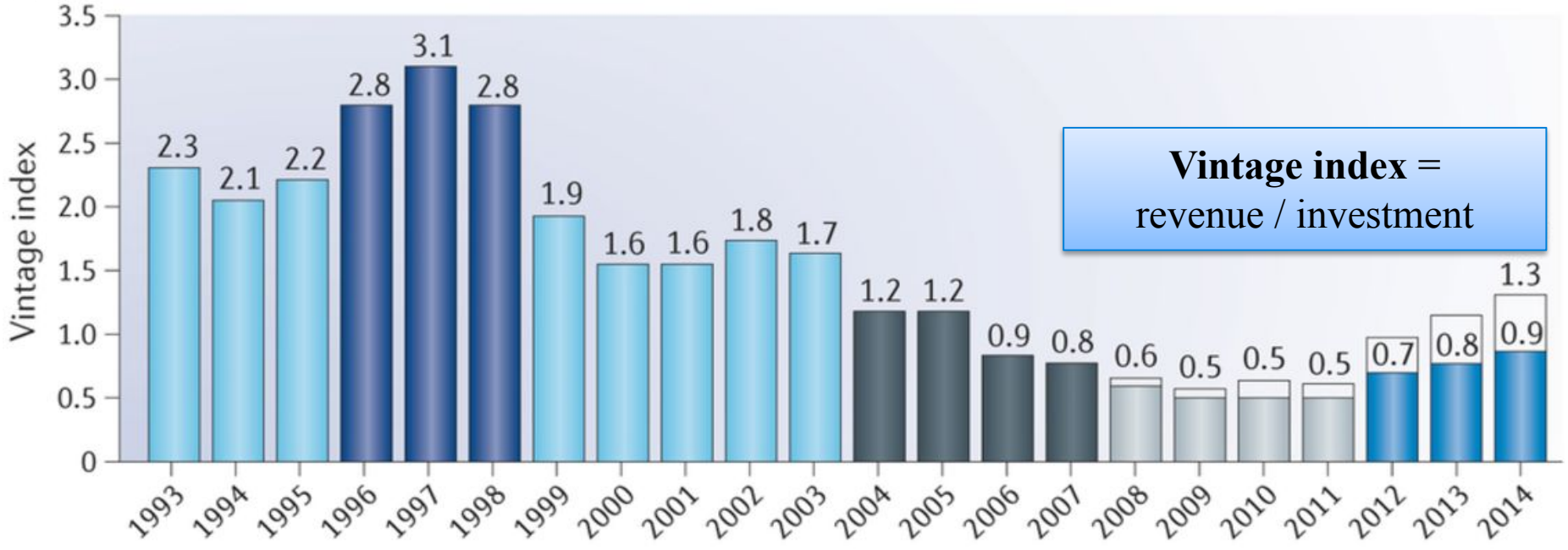
Adapted from Paul *et al.* "How to Improve R&D Productivity: The Pharmaceutical Industry's Grand Challenge." Nature Reviews Drug Discovery, 2010

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.

Increasing cost and decreasing return of investment in drug discovery



Modified from Smietana *et al.* "Improving R&D Productivity." Nature Reviews Drug Discovery, 2015

Finding new drugs has become more challenging and expensive

危机

~~*n.* crisis~~

Danger + Opportunity

Applied mathematics empowers drug discovery in many ways

Applied mathematics *in drug discovery*
is not a definable scientific field
but a human attitude.

Richard Courant
(1888-1972)

Quantitative critical thinking in communication and decision making

Statistics, Data Mining and Machine Learning	Applied Combinatorics and Graph Theory	Dynamical Systems	Causal inference
Ordinary / Partial/ Stochastic Differential Equations	Network Analysis	Molecular, Quantum, and Continuum Mechanics	Multiscale modelling

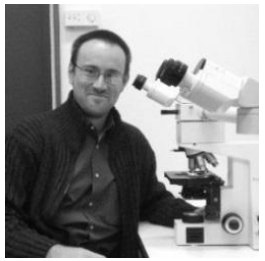
Acknowledgements





F. Hoffmann-La Roche Ltd	
Clemens Broger†	**Faye Drawnel**
Martin Ebeling	Markus Britschgi
Manfred Kansy	Roland Schmucki
Fabian Birzele	Martin Stahl
Kurt Amrein	Isabelle Wells
Annie Moisan	**Lu Gao**
Luca Piali	Lue Dai
John Young	Ravi Jagasia
Lisa Sach-Peltason	**Marco Prunotto**
Mark Burcin	**John Moffat**
Christoph Patsch	Gang Mu
Michael Reutlinger	Jianxun Jack Xie
Matthias Nettekoven	Filip Roudnicky
Andreas Dieckmann	Holger Fischer
Klas Hatje	Iakov Davydov
Laura Badi	**Ulrich Certa**
Tony Kam-Thong	Detlef Wolf
Corinne Solier	Ken Wang
Thomas Singer	Nikolaos Berntenis





External to Roche
Stefan Wiemann
Wolfgang Huber
Ozgür Sahin
Agnes Hovrat
Katharina Zweig
Sally Cowley
Alexandros Stamatakis
Michael Prummer
Mark D. Robinson
Michael Hennig
Philipp Mekler
Jung Kyu Canci
Verdon Taylor
Maria Anisimova
Lorenzo Gatti
Erhard van der Vries
Ab Osterhaus
Nevan Krogan
Oliv Eidam





Conclusions and perspectives

- **It is now probably the best time in human history to join the fight against diseases.**
- **We learned about modalities and the drug discovery and development process.**
- **Interdisciplinary research**, especially applying mathematical approaches and tools to biological, chemical and medicinal questions, **is imperative** to fill the knowledge gaps and to make potent, safe, and efficacious drugs and to perform accurate diagnosis.
- **Mathematics and informatics will continue transforming drug discovery by integrating human intelligence and creativity, data, and machine intelligence.**
  - From correlation to causation
  - From qualitative description to quantitative prediction
  - From trial-and-error to systematic understanding
  - From population inference to individual prediction and continuous intervention
  - From descriptive studies of biology to predictive and counterfactual models
  - From observations to engineering and synthesis of the biological system
- **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 (*multiscale-modeling of drug mechanism and safety)*. The Mathematical and Computational Biology in Drug Discovery (MCBDD) course in spring semesters builds upon the basics in AMIDD and introduces advanced topics.**

Contact the author

Offline Activities

OA1: Questions on the video on Herceptin by Susan Desmond-Hellmann

Link to the video

Questions for the video

1. What is the **indication** of *Herceptin*? What is its generic (USAN, or United States Adopted Name) name?
2. What is the **gene target** of Herceptin?
3. In which year was the **target** of Herceptin described? When was Herceptin **approved**?
4. What was the **improvement** of Herceptin compared with earlier antibodies?
5. Why does a **biomarker** matter besides developing drugs?
6. In the clinical trial of *Herceptin* for **metastatic breast cancer**, how much improvement in the **median survival** did Herceptin achieve? And how much improvement is in the **adjuvant setting** (Herceptin applied directly after operation)?

Questions for further thinking

- Susan Desmond-Hellmann summarizes successful drug development in four aspects: (1) having a deep understanding of the basic science and the characteristics of the drug, (2) targeting the right patients, (3) setting a high bar in the clinic, and (4) working effectively with key regulatory decision makers. Where do you think mathematics and computer science play a crucial role?
- She emphasized the importance of collaboration. What skill sets do we need for that?
- How do you like her presentation? Anything that you can learn from her about presentation and storytelling?

OA2: Required and recommended reading

[Required]

Principles and workflow of early drug discovery:

Hughes, JP, S Rees, SB Kalindjian, and KL Philpott. 2011. "Principles of Early Drug Discovery." British Journal of Pharmacology 162 (6): 1239–49. https://doi.org/10.1111/j.1476-5381.2010.01127.x.

[Recommended]

History of drug discovery and the rise of pharmaceutical company:

Jones, Alan Wayne. 2011. "Early Drug Discovery and the Rise of Pharmaceutical Chemistry." *Drug Testing and Analysis* 3 (6): 337–44. https://doi.org/10.1002/dta.301.

Mathematics and biology:

Cohen, Joel E. 2004. "Mathematics Is Biology's Next Microscope, Only Better; Biology Is Mathematics' Next Physics, Only Better." PLOS Biology 2 (12): e439. https://doi.org/10.1371/journal.pbio.0020439.

Extensive reading about history of medicine:

- Taking the Medicine: A Short History of Medicine's Beautiful Idea, and our Difficulty Swallowing It by Druin Burch (ISBN: 1845951506, ISBN13: 9781845951504)

OA3: Preparation for Lecture 2 and 3

If you want to learn more about the **Central Dogma of Molecular Biology**:

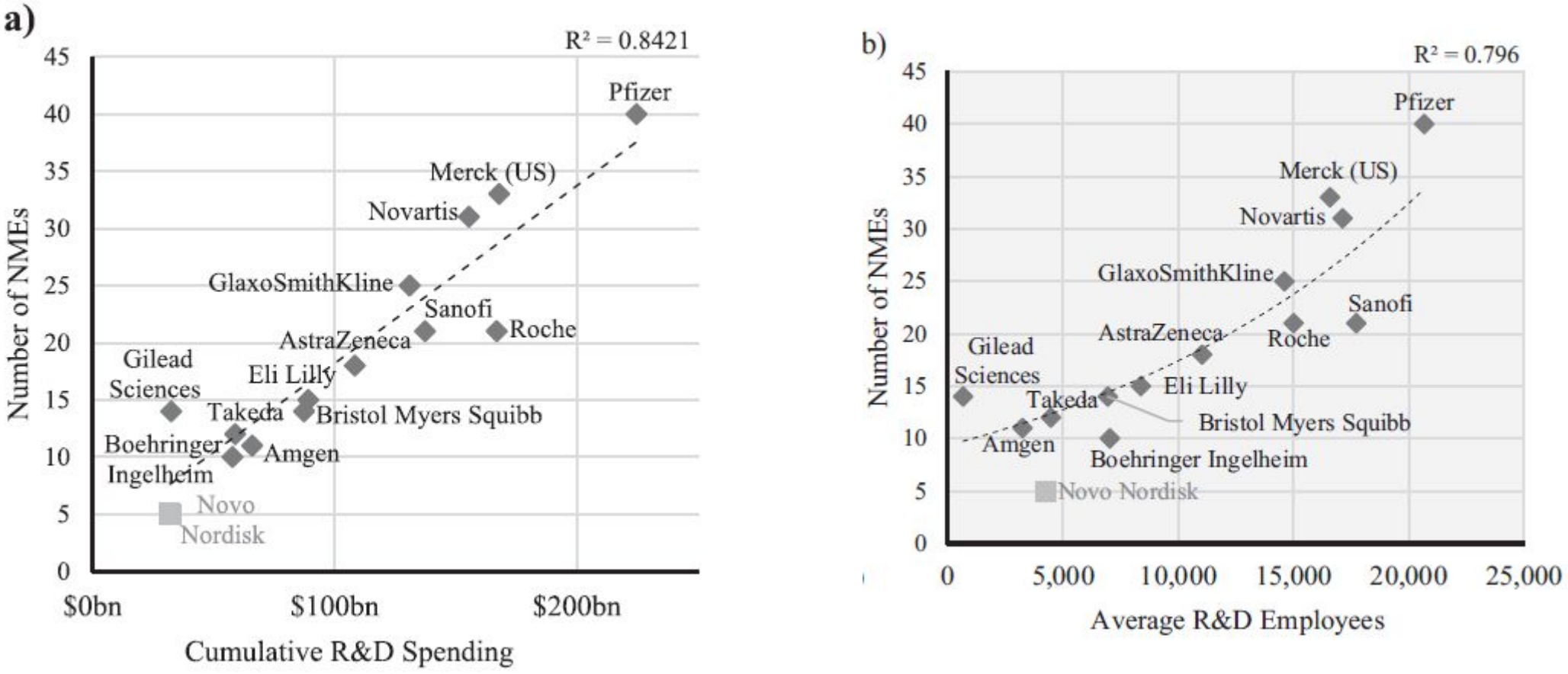
1. If you are a film person, watch this video on Youtube about the Central Dogma of Life, and this animated film.
2. If you are a reading person, read *Biology Briefs,* a six-part introductory article series run by *The Economist*, or if you prefer textbooks, read *DNA & The Central Dogma of Biology* by Prof. Henry Jakubowski in the Biology LibreTexts.

If you are already familiar with the Central Dogma and wishes to refresh relevant **mathematical knowledge**:

1. Find on Youtube or other websites introductory courses to probability (like this)
2. Read the first chapter of *Biological Sequence Analysis* by Richard Durbin *et al.*, available here

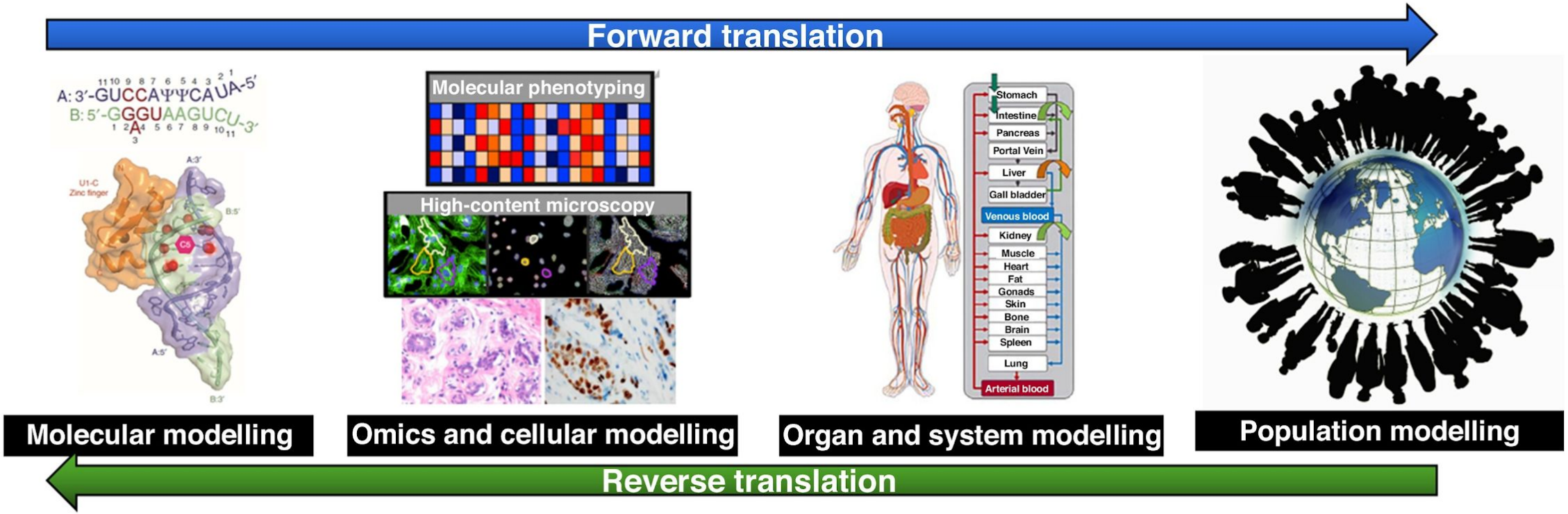
Backup slides

R&D of leading pharmaceutical companies neccesates both high spending and huge organizations



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.

The alternative, 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.

Introduction to Applied Mathematics and Informatics in Drug Discovery (*AMIDD*)

A course series at DMI, University of Basel

- **Introduction to drug discovery**
- **Molecular modelling**
  - Biological sequence analysis
  - Protein sequence and structure
  - Molecular modelling and dynamics
- **Omics and cellular modelling**
  - From drug-target interactions to networks
  - Gene expression profiling
  - Cell-based phenotypic drug discovery
- **Mathematical modelling**
  - Principles and applications of modelling in pharmacology
  - Pharmacokinetics (PK) and pharmacodynamics (PD) modelling
  - Clinical pharmacology and pharmacometrics
- **Population modelling**
  - Non-linear mixed-effect models (NLMEs)
  - Essentials of clinical trials
- **Guest lectures**
- **Your presentations**

It is hoped that AMIDD builds a bridge between students and quantitative aspects of drug discovery