

AI and the Oncology Drug Development Pipeline

James P. Long

UTRGV STEM Research Conference

April 24, 2026

AI Generated Fight Scene



Quality of AI generated fight scene is easy to assess.

If it **looks like** a good fight scene, it **is** a good fight scene.

AI Generated Websites

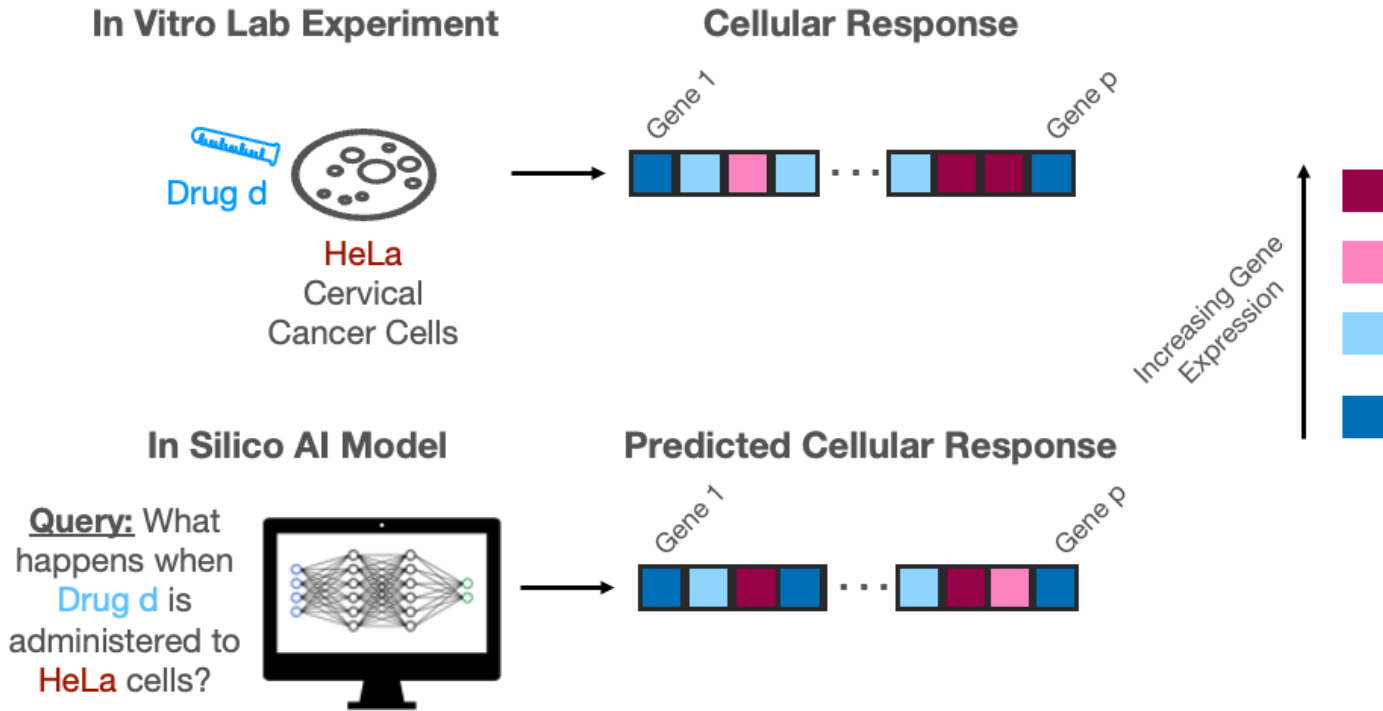


Claude Code Prompt:
“Create a website with a blue background”



- Website looks like what we want.
- But website is not just appearances:
 - Do data entry forms work?
 - Are there security vulnerabilities?
- If it **looks like** a good website, it **might** be a good website.

AI Generated Results in STEM



nature methods

[Explore content](#) [About the journal](#) [Publish with us](#)

[nature](#) > [nature methods](#) > [articles](#) > [article](#)

Article | Published: 26 February 2024

scGPT: toward building a foundation model for single-cell multi-omics using generative AI

[Haotian Cui](#), [Chloe Wang](#), [Hassaan Maan](#), [Kuan Pang](#), [Fengning Luo](#), [Nan Duan](#) & [Bo Wang](#) [✉](#)

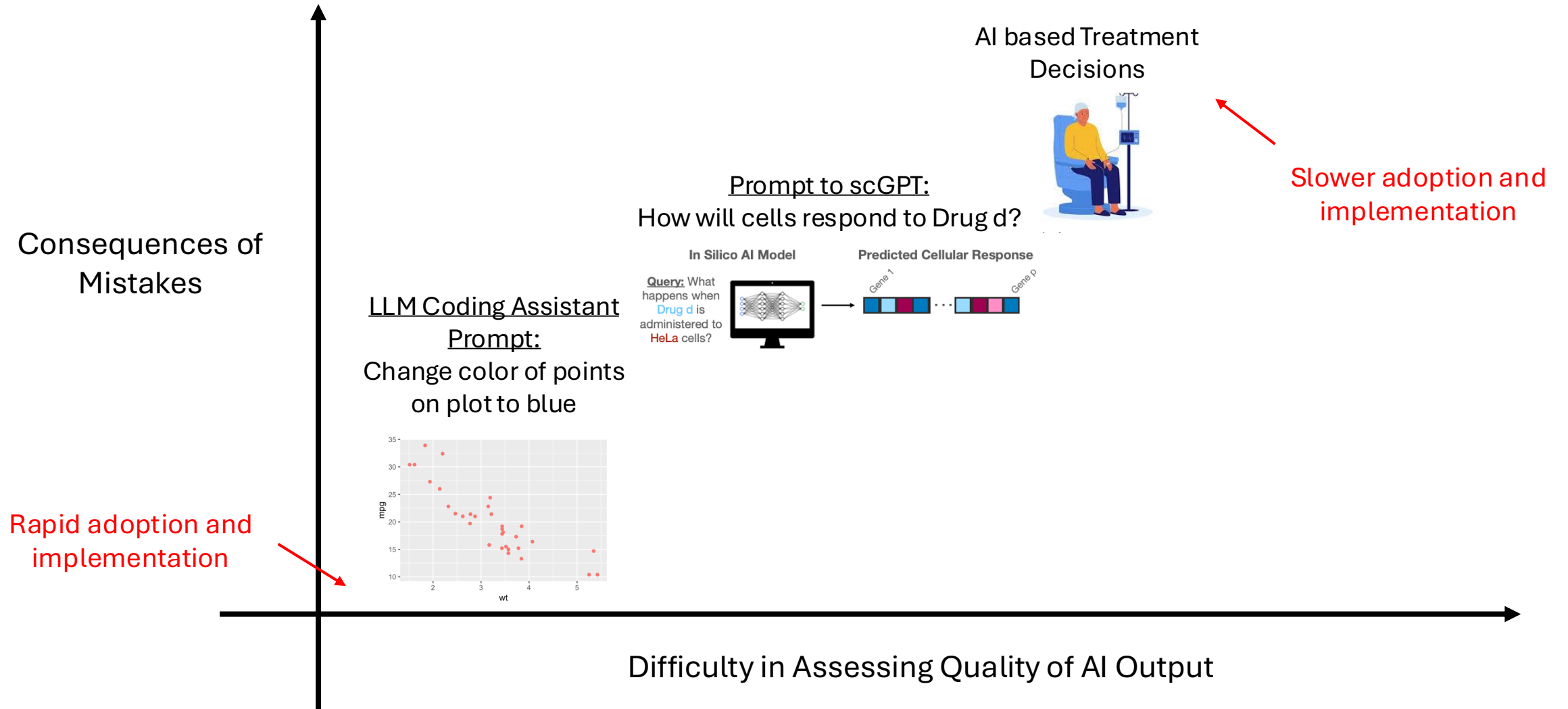
Nature Methods **21**, 1470–1480 (2024) | [Cite this article](#)

190k Accesses | 1069 Citations | 227 Altmetric | [Metrics](#)

Can AI model predictions replace (some) lab experiments?

The **look** of the results does not tell us about their quality.

Low Risk and High Risk Applications of AI in STEM



Overview of Talk

- Background on UT MD Anderson
- Overview the Oncology Drug Development Pipeline
- Discuss how AI is changing pipeline
 - Accelerating development
 - Creating bottlenecks
 - Generating new research opportunities

UT MD Anderson

Cancer Hospital

Patients receive cancer treatment



- Largest cancer center in world by patient volume
- Run more clinical trials in oncology than any other institution

Research Institute

Develop new cancer therapies and prevention strategies



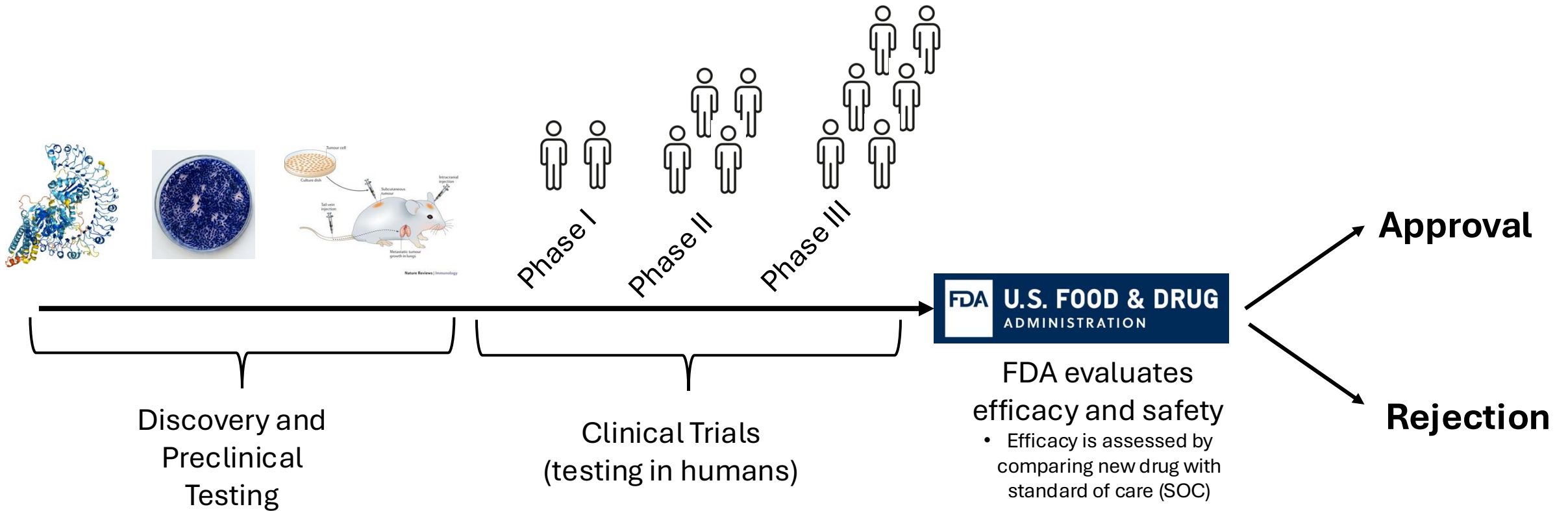
THE UNIVERSITY OF TEXAS
MDAnderson Cancer Center | UTHealth Houston
Graduate School of Biomedical Sciences
PhD and MS Programs



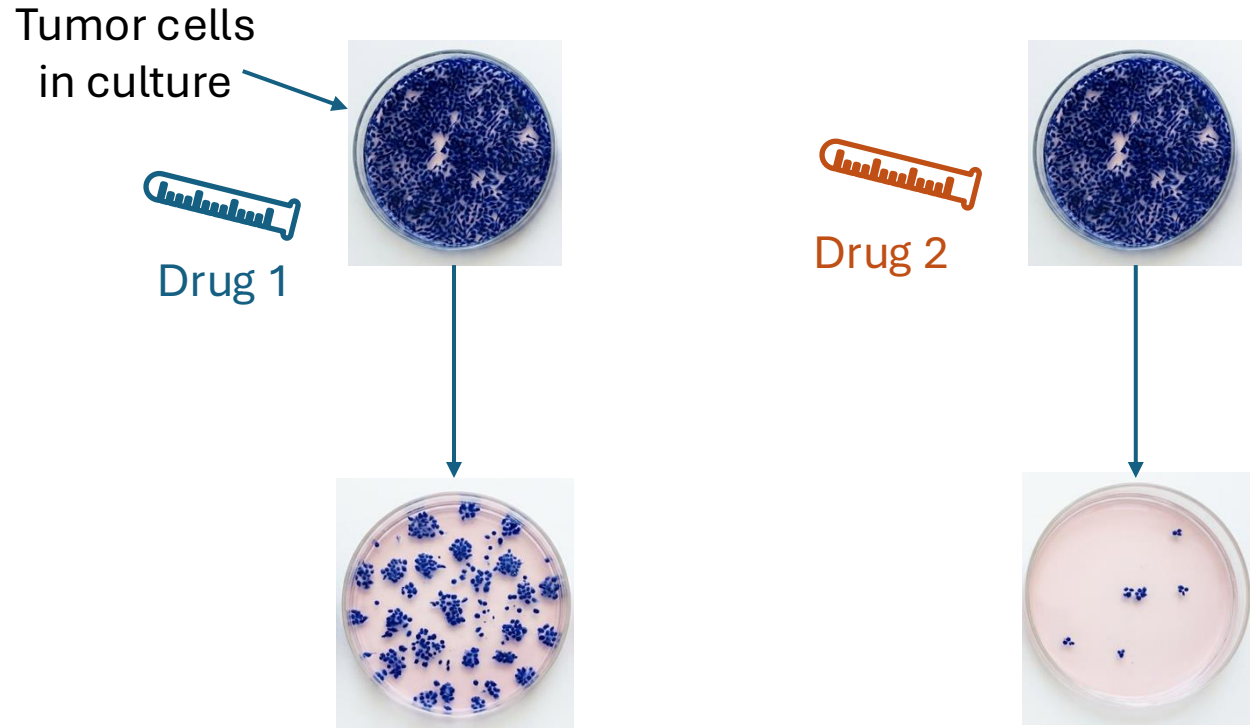
Division of Discovery Science

- Bioinformatics and Computation Biology
- **Biostatistics**
 - 24 Faculty
 - 30 Staff Biostatisticians
- Cancer Biology
- Comparative Medicine
- Genetics
- Immunology
- Systems Biology
- Veterinary Medicine and Surgery
- Epigenetics and Molecular Carcinogenesis

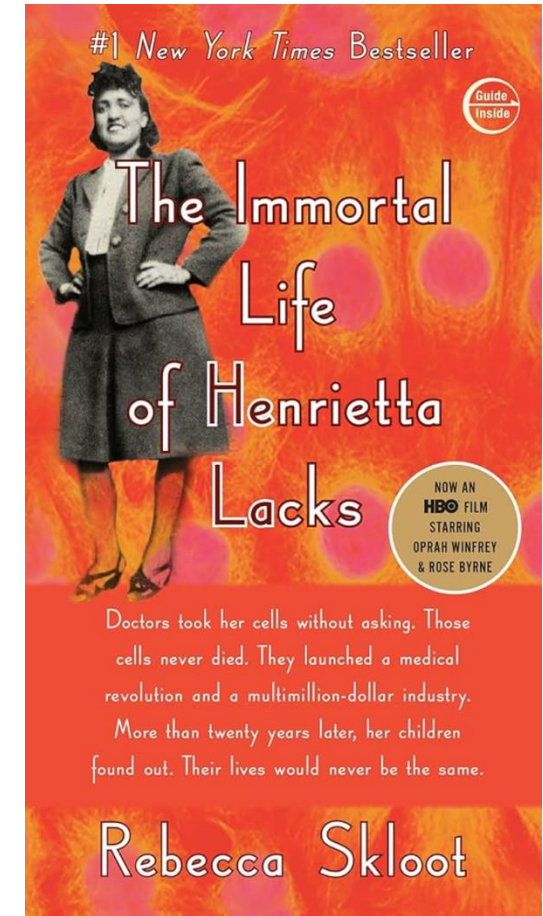
Oncology Drug Development Pipeline



Preclinical Cell Line Studies

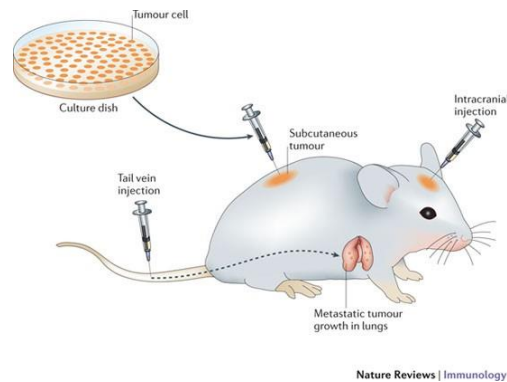


Drug 2 is more effective at killing tumor cells than drug 1.



- Many tumor cell lines (Breast, Brain, Ovarian, . . .)
- First immortalized human cell line HeLa taken from Henrietta Lacks in 1951

Preclinical Mouse Studies



Treatment

10 mice / group

Vehicle

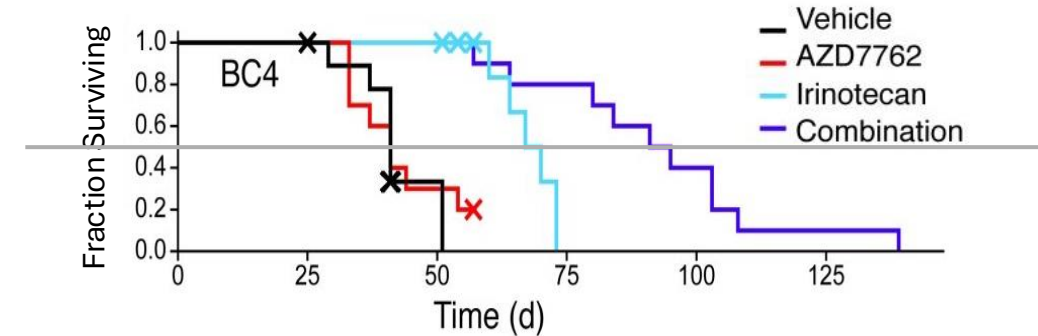
AZD7762 (A)

Irinotecan (I)

Combination (A+I)

- Important to analyze experimental results from multiple perspectives (survival and tumor growth).
- Tumor growth curves support and explain mouse survival data.

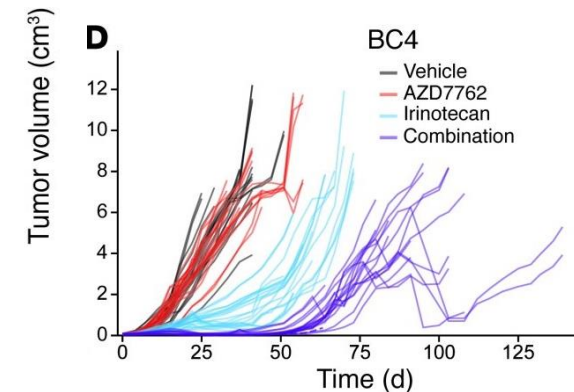
Kaplan-Meier Survival Curves



Median survival:

- Vehicle and AZD ~ 40 days
- Irinotecan ~ 70 days
- Combination ~ 90 days

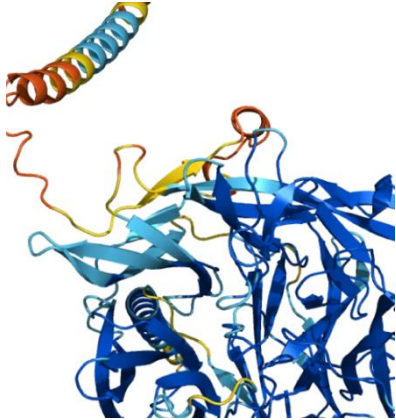
Tumor Growth Curves



- Dranoff, G., 2012. Experimental mouse tumour models: what can be learnt about human cancer immunology?. *Nature Reviews Immunology*, 12(1), pp.61-66.
- Ma, C.X., Cai, S., Li, S., Ryan, C.E., Guo, Z., Schaiff, W.T., Lin, L., Hoog, J., Goiffon, R.J., Prat, A. and Aft, R.L., 2012. Targeting Chk1 in p53-deficient triple-negative breast cancer is therapeutically beneficial in human-in-mouse tumor models. *The Journal of clinical investigation*, 122(4), pp.1541-1552.

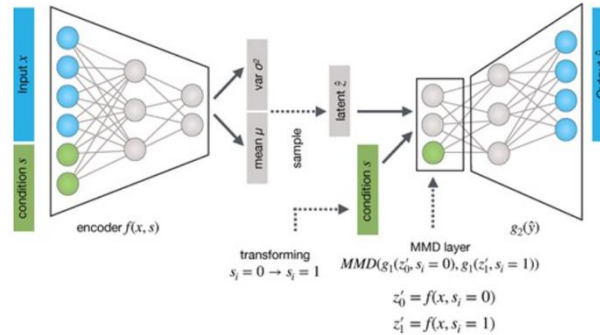
AI In Preclinical Drug Testing

AlphaFold predicts protein 3D structure
(Drug Discovery)



Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A. and Bridgland, A., 2021. Highly accurate protein structure prediction with AlphaFold. *nature*, 596(7873), pp.583-589

In Silico Perturbation Models Predict How Cells Respond to Treatment

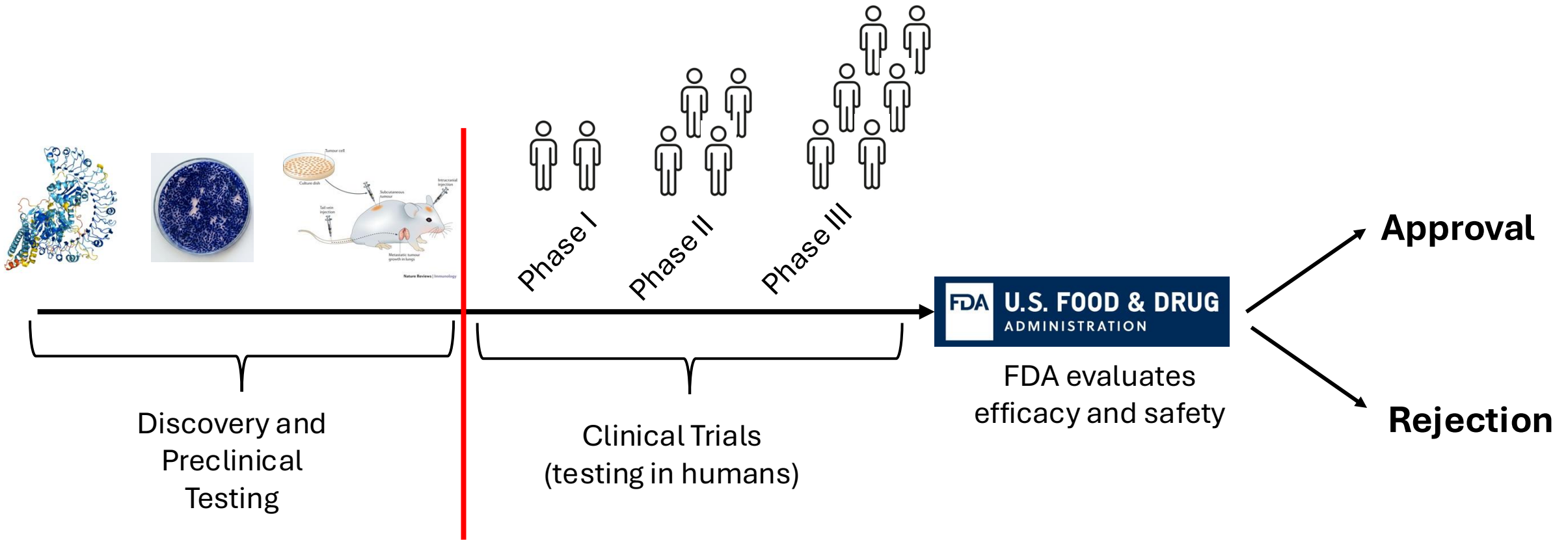


Lotfollahi, M., Naghipourfar, M., Theis, F.J. and Wolf, F.A., 2019. Conditional out-of-sample generation for unpaired data using trVAE. *arXiv preprint arXiv:1910.01791*.

Companies Focused on AI Driven Drug Discovery



Preclinical work is lower risk and AI model failures are manageable.

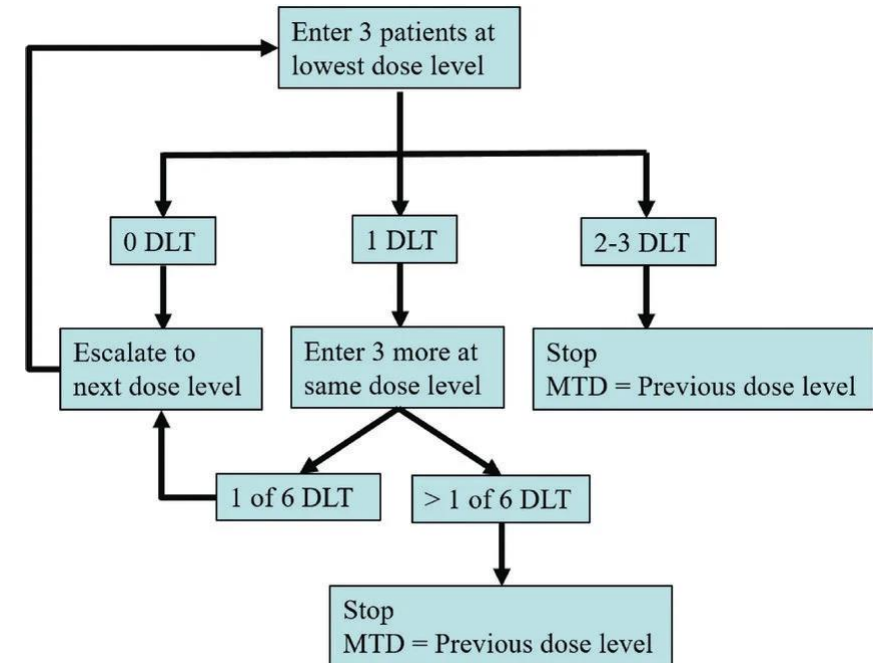


If preclinical results are promising, testing in humans begins.

Phase I Trials

- Often first in human studies
- Goal:
 - Assess safety
 - Identify the maximum tolerated dose (MTD)
- Enroll ~ 20 participants
- All participants receive study drug (no controls)
- Select set of doses to test:
 - DL0, DL1, DL2, . . . , DLX
 - Start enrolling patients at DL0
- DL increased / decreased based on dose limiting toxicities (DLT)

3+3: Algorithm



Strengths:

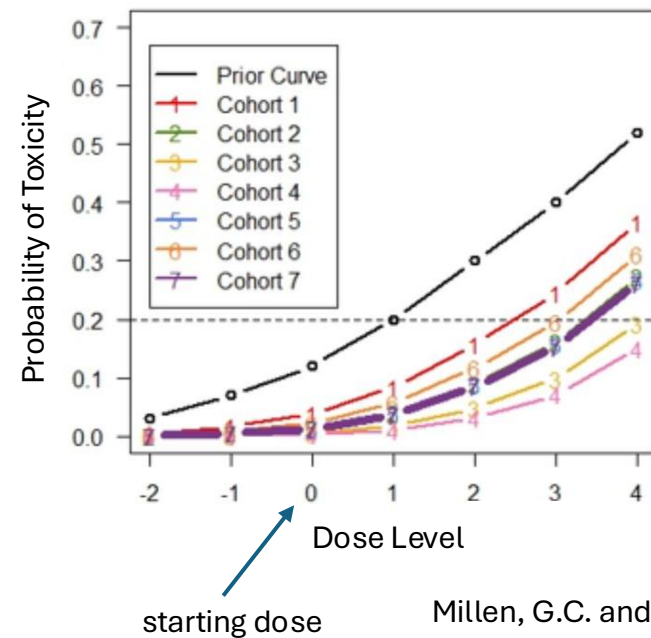
- Easy to use / understand

Weaknesses:

- Targets 33% toxicity rate
- Often under / overdoses

Phase I Design: Continual Reassessment Method (CRM)

- Model relationship between dose and probability of toxicity
- Update model and select DL dynamically as data accrues
- Uses Bayesian statistical methods to update toxicity-dose model



Cohort	Dose	#DLT/#patients
1	0	0/3
2	1	0/3
3	2	0/2
	3	0/1
4	3	0/2
5	3	1/3
6	3	1/3
7	3	0/4

Millen, G.C. and Yap, C., 2021. Adaptive trial designs: what is the continual reassessment method?. *Archives of Disease in Childhood-Education and Practice*, 106(3), pp.175-177.

Statisticians developed CRM and many other Phase I designs to select dose levels in Phase I trials.

Phase II Trials

- Objectives:
 - Preliminary Efficacy Assessment
 - Safety assessment
- All participant receive study drug (no controls)
 - “Single arm study”
- N ~ 50 participants
- Efficacy assessment is often tumor shrinkage
- At conclusion of trial, go/no go decision for Phase III trial

Enroll patients on trial.

Inclusion criteria: Locally advanced lung cancer



Patient outcomes from standard of care treatments derived from medical literature



	Experimental Therapy (n=50)	Standard of Care (Control)
Tumor Shrinkage	48% (24/50)	25%
Adverse Events (Side Effects)	28% (14/50)	10%

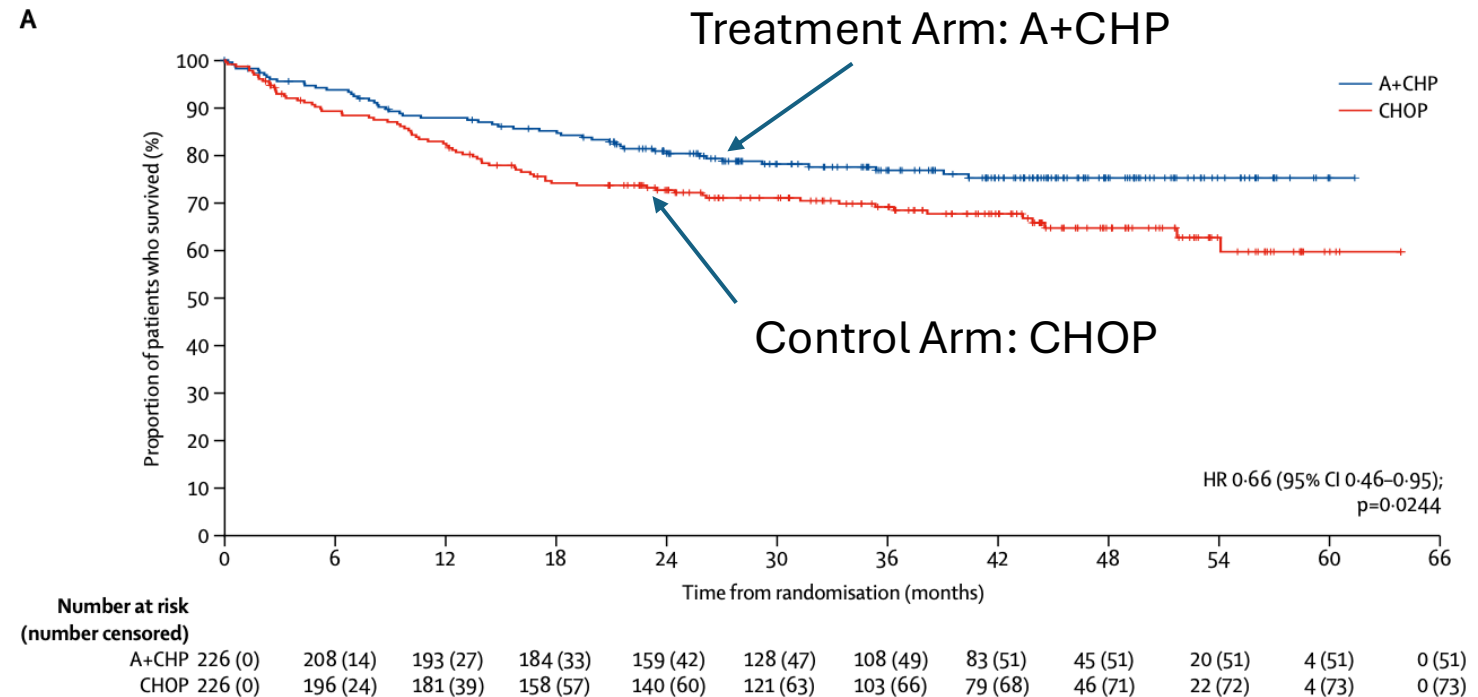
Conclusion: New treatment appears more effective but more toxic.

Phase III Trials

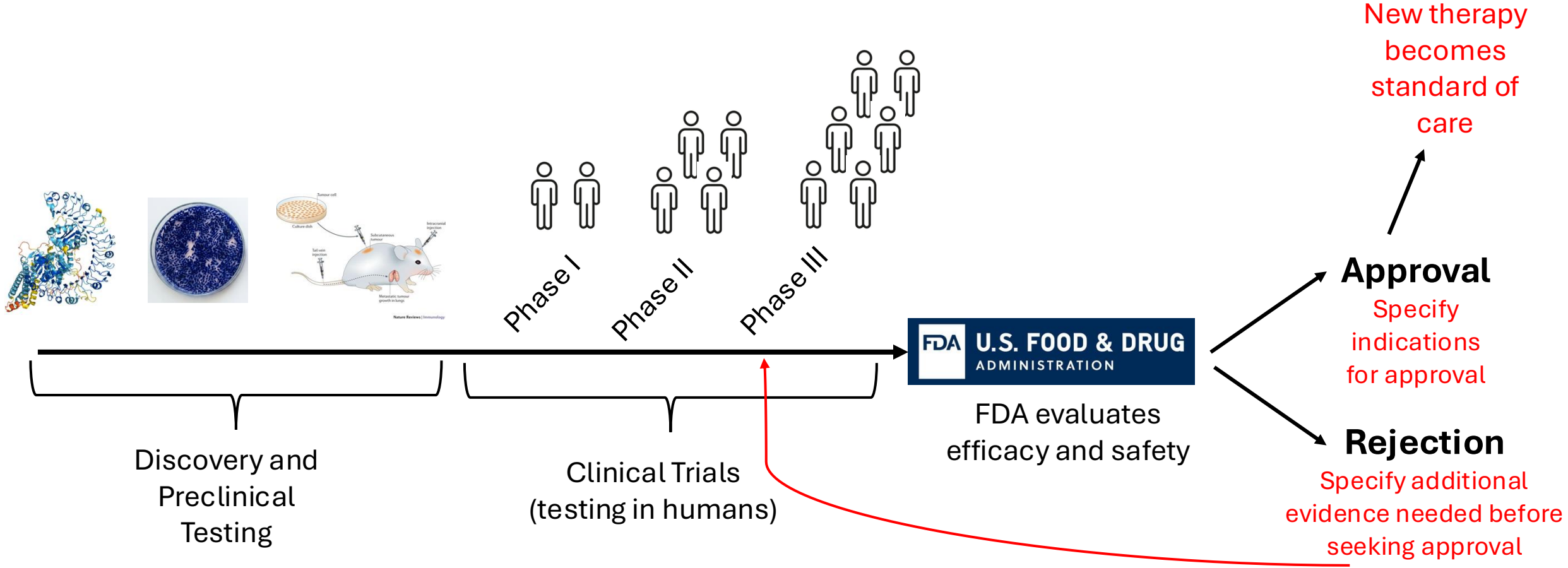
- Randomized trials comparing standard of care (SOC) to new therapy
- n ~ 100s – 1000s
- Multi-institutional
- Often 5+ years
- Goal: Rigorous demonstration of efficacy and change of clinical practice

ECHELON-2 study

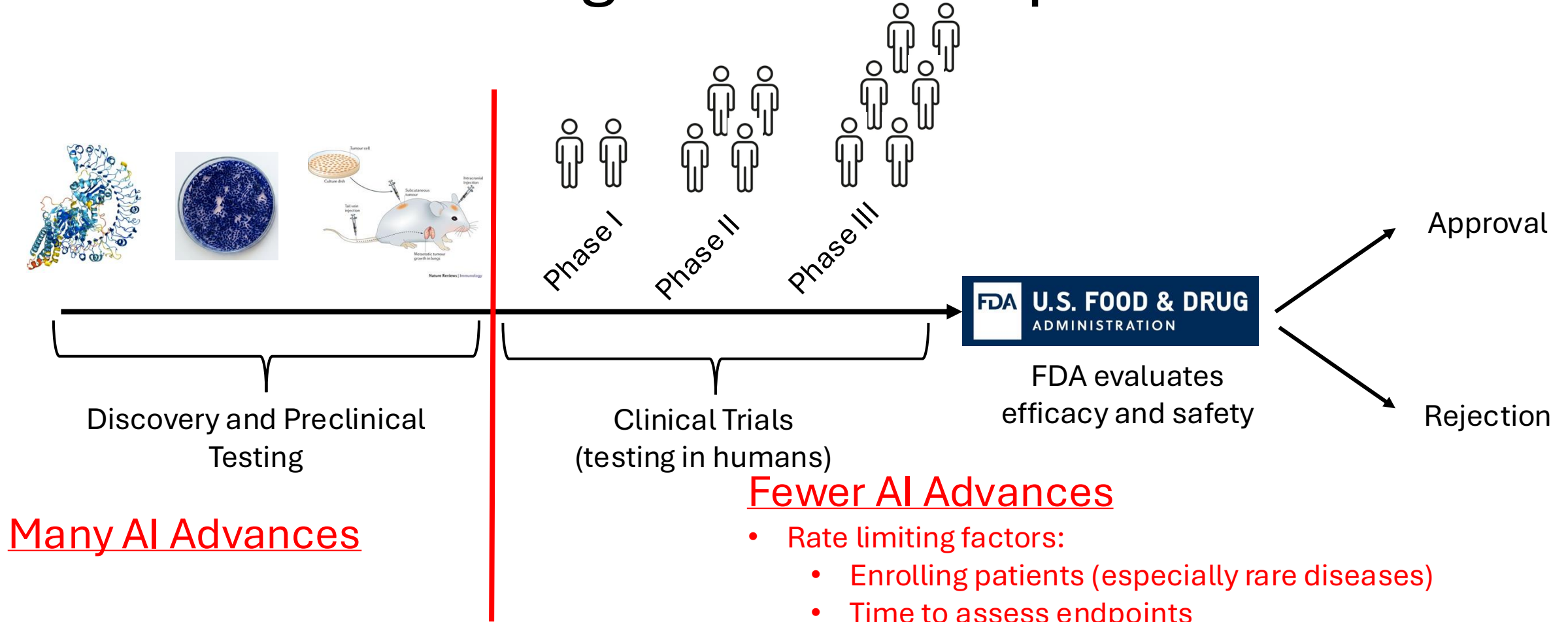
- Enrolled ~400 patients with CD30+ peripheral T-cell lymphoma (PTCL)
 - N=200 treated with A+CHP (new drug)
 - N=200 treated with CHOP (SOC)
- ~ 3.5 years of patient enrollment



- This trial changed standard of care.
- After trial, PTCL patients prescribed mostly A+CHP



AI Is Accelerating Preclinical Pipeline



- Many good candidate therapeutics
- Not enough clinical resources to test them

Fewer AI Advances

- Rate limiting factors:
 - Enrolling patients (especially rare diseases)
 - Time to assess endpoints (e.g. % of patients surviving 5 years)
 - Many regulations

Possible Solutions to Bottleneck

- Relax RCTs requirements (no Phase III trials)
 - Alternative: Use results from Phase II studies to assess efficacy.
 - Compare responses of Phase II trial participants to historical response rates.
- Use Surrogate Endpoints
 - Clinical efficacy endpoint: 5-year survival
 - Surrogate endpoints: tumor shrinkage
 - Can be determined earlier in trial
 - Less directly relevant to patients

Debates over RCTs and Surrogate Endpoints

Advocates for more flexibility
“deregulatory approach”



Dr. Marty Makary
FDA Commissioner

THE WALL STREET JOURNAL. FDA: How We’re Approving More Drugs to Treat Rare Diseases

Our application process was designed for common ailments, but roughly 30 million Americans suffer from uncommon ones.

By Marty Makary

Feb. 23, 2026 11:15 am ET



Favors RCTs,
clinical endpoints



Dr. Vinay Prasad
Director of Center for Biologics Evaluation and Research (CBER) at FDA

THE WALL STREET JOURNAL. Who’s in Charge at the FDA?

Commissioner **Marty Makary** talks about accelerated cures while biologics chief **Vinay Prasad** quietly scuttles them.

By The Editorial Board [Follow](#)

Feb. 24, 2026 5:25 pm ET



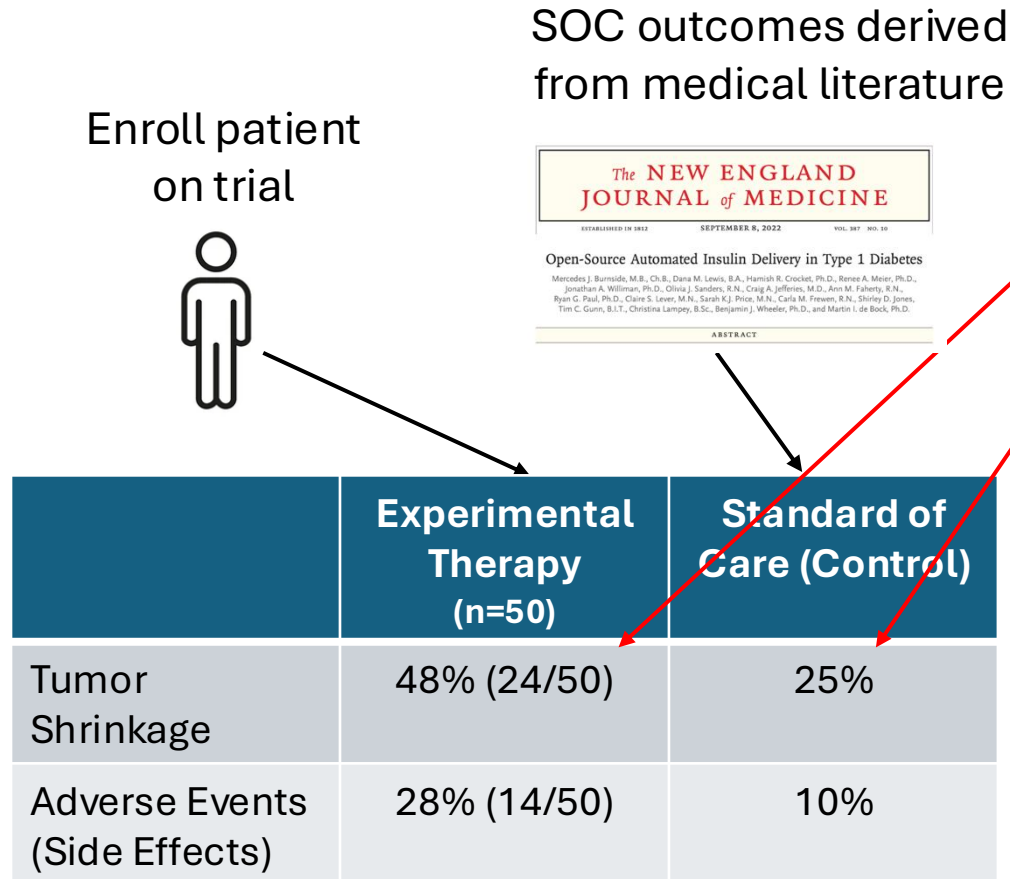
This is a Research Opportunity

1. Identify limitations with **single arm trials** and surrogate endpoints
2. Develop and validate methods to overcome these limitations
3. Implement new methods in clinical trials

Where is AI creating bottlenecks in your area of STEM?

Are there research opportunities to overcome the bottlenecks?

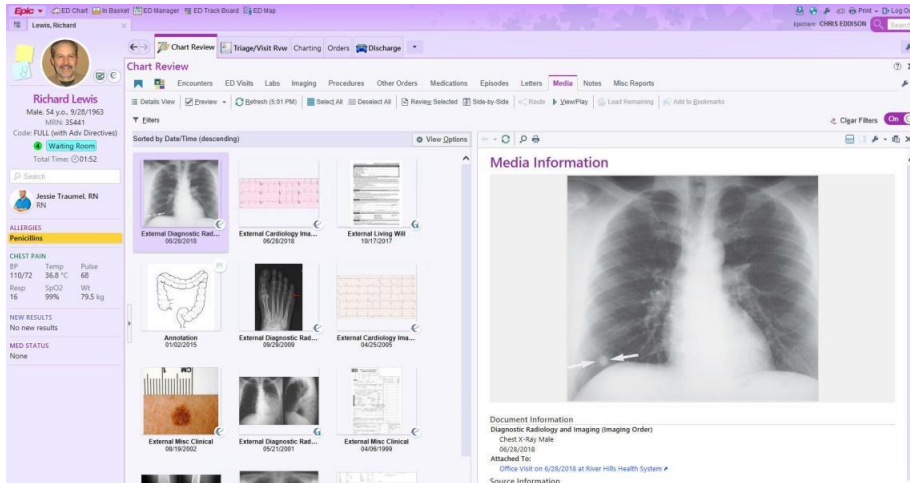
Limitations of Single Arm Trials (Phase II)



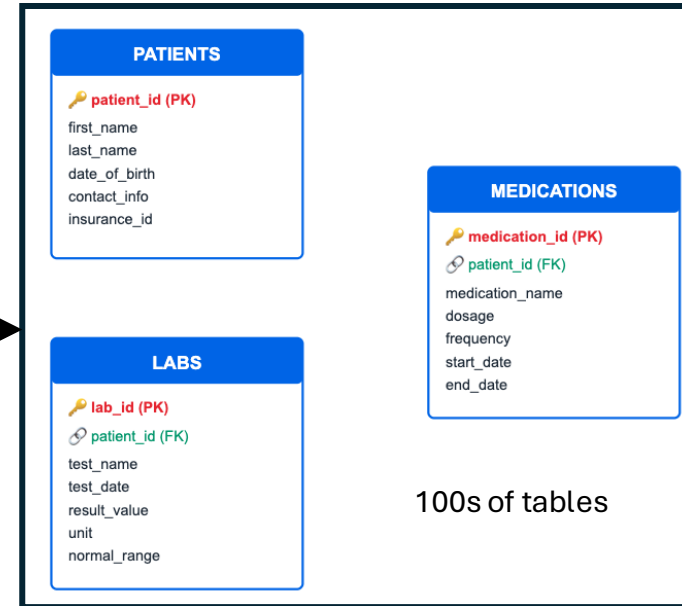
- Observe differences between outcomes for patients treated on trial and SOC response rates taken from the literature.
- Explanation for difference:
 1. New Treatment is better
 2. **Confounders:** Differences in patient characteristics between trial patients and controls
Examples: Age, # prior lines of treatment, disease subtype
- Randomization in Phase III trials eliminates confounding

One Solution: Synthetic Control Cohorts

Epic Electronic Health Record



Databases of Millions of Medical Records



1) Identify control patients from databases
Example: If testing new drug in patients with metastatic breast cancer, identify controls with this diagnosis.

2) Compare Efficacy:
Statistical methods used to compare outcomes between cohorts, accounting for confounding variables.

Age	Race	ECOG	HER2 Expression	# Prior Lines Tx	Time to Progression	Time to Death	Last FU
25	Asian	0	10%	NA	1.2 yr	2.5 yr	2.5 yr
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
50	Black	1	40%	1	2.2 yr	-	4 yr
70	White	Sarcoma (MPNST)	5%	0	3.0 yr	4.8 yr	4.8 yr
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

“Synthetic controls” derived from database, treated with SOC

Patients treated on trial with experimental therapy



Do Synthetic Control Arms Work?

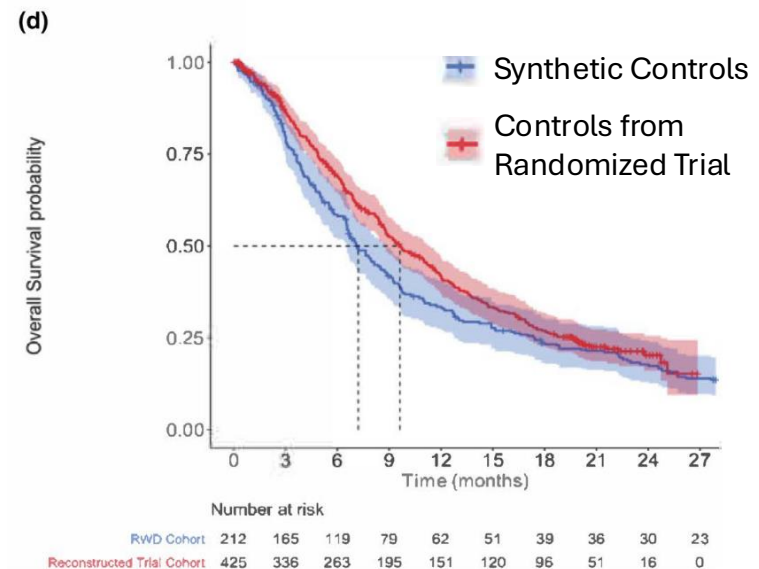


Validation Strategy: Use synthetic cohorts to reproduce results of control arms from completed Phase III trials (where ground truth is known).

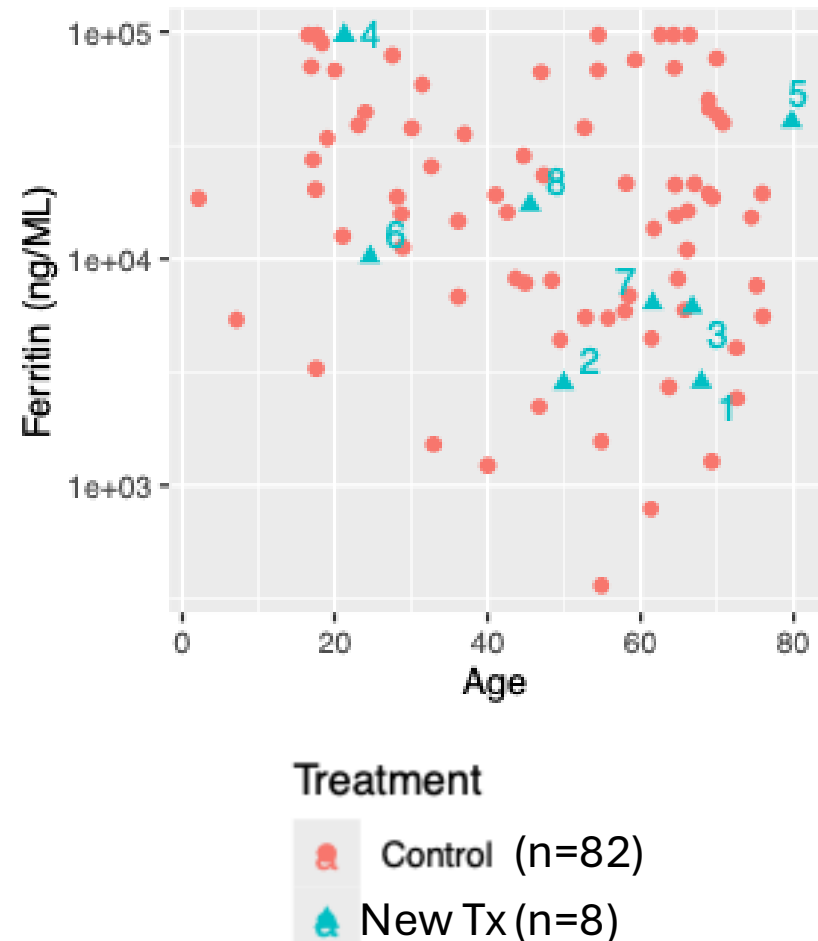
ARTICLE

Emulating Control Arms for Cancer Clinical Trials Using External Cohorts Created From Electronic Health Record-Derived Real-World Data

Katherine Tan^{1*}, Jonathan Bryan¹, Brian Segal¹, Lawrence Bellomo¹ , Nate Nussbaum¹, Melisa Tucker¹, Aracelis Z. Torres¹, Carrie Bennette¹, William Capra², Melissa Curtis¹ and Rebecca A. Miksad¹ 



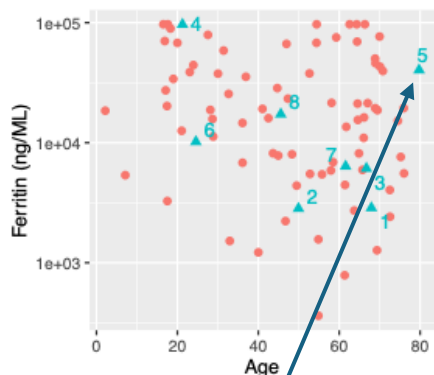
Piloting Synthetic Controls at MD Anderson





- HLH is extremely rare disease with poor prognosis
 - Median survival time ~ 2 months with current therapies
- Treated n=8 patients with new experimental therapeutic on trial
 - No control arm
- Used MDA Electronic Health record databases to identify synthetic controls
 - Age and Ferritin level possible confounders / prognostic variables
- For each of the 8 trial patients, we identified 10 controls with closest Age and Ferritin

Preliminary Results

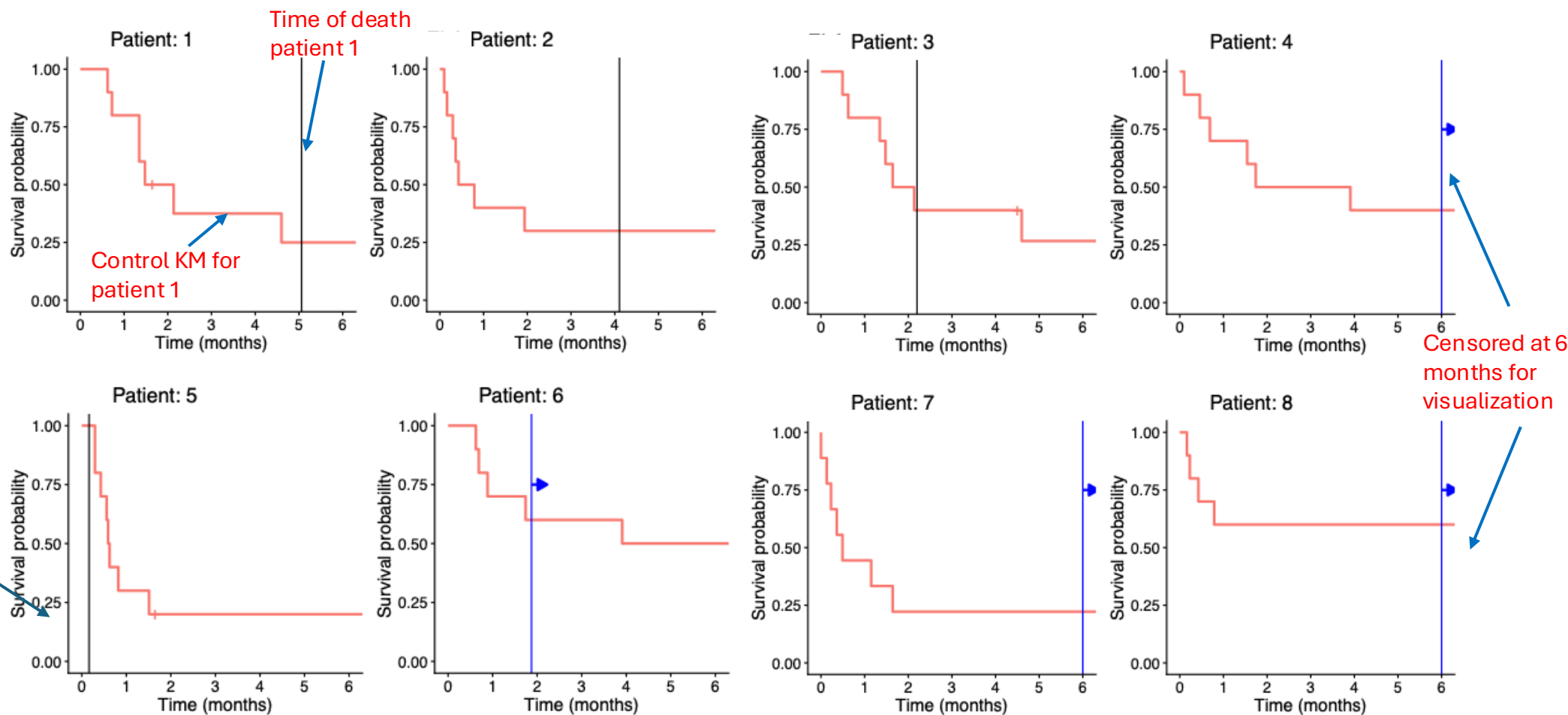
Using the 10 matched controls, created subject-specific KM curve representing expected outcomes under SOC



Treatment

-  Control (n=82)
-  New Tx (n=8)

- Patient 5 has very poor survival.
- Matched controls have poor survival as well with 2-month OS < 25%.



Message: By matching participants to controls we can explain some of the observed variation in outcomes.

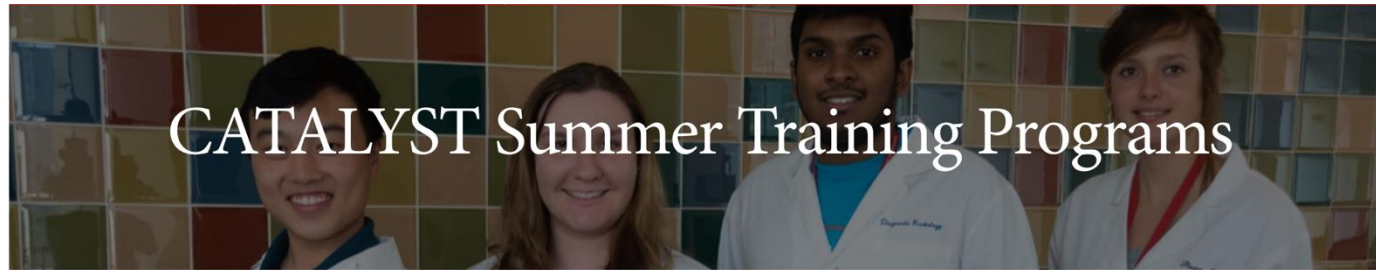
Longer Term Goal

- Routine generation of Synthetic Control cohorts for Phase II trials at MD Anderson
- More informed go / no go decisions for Phase III trials
- Adjust inclusion criteria for Phase III trials based on results
 - Example: Synthetic controls suggest treatment only effective in subgroup
- Facilitate regulatory approval of therapeutics

Closing Thoughts

- AI induced bottlenecks create research opportunities
- AI Usage Suggestion
 - Learn how to solve problems yourself first
 - Then use AI to reduce workloads
 - Learn long division in elementary school before using calculator
 - Learn to code before relying on AI code generators
- Model quality assessment is becoming a whole research area in itself
 - Two of my most recent papers involved testing/validating AI models
 - Good way to become involved in AI research for skeptics
- Focus on important problems and use whatever methods are effective to solve them.

MD Anderson Catalyst Program



Research Training > Early Career Pathway Programs > CATALYST Summer Training Program

CATALYST Summer Training Program



Programs & Eligibility

How to Apply & Application
FAQ

Activities & Events

MD Anderson CATALYST Summer Training Programs

Cancer Advanced Training and Learning for Young Scientists

- Summer programs for high school, undergraduates, graduate students in science and medicine
- Google “MD Anderson Catalyst Program”
 - <https://www.mdanderson.org/education-training/research-training/early-career-pathway-programs/summer-research-programs.html>

Thank you!
Questions?