Advancing Precision Medicine in Early-stage Cancer

Daniel J. Boffa, MD, MBA

Thoracic Surgeon

Disclosures

- Patent in cryptography August 2025
- Board of a nonprofit Data2

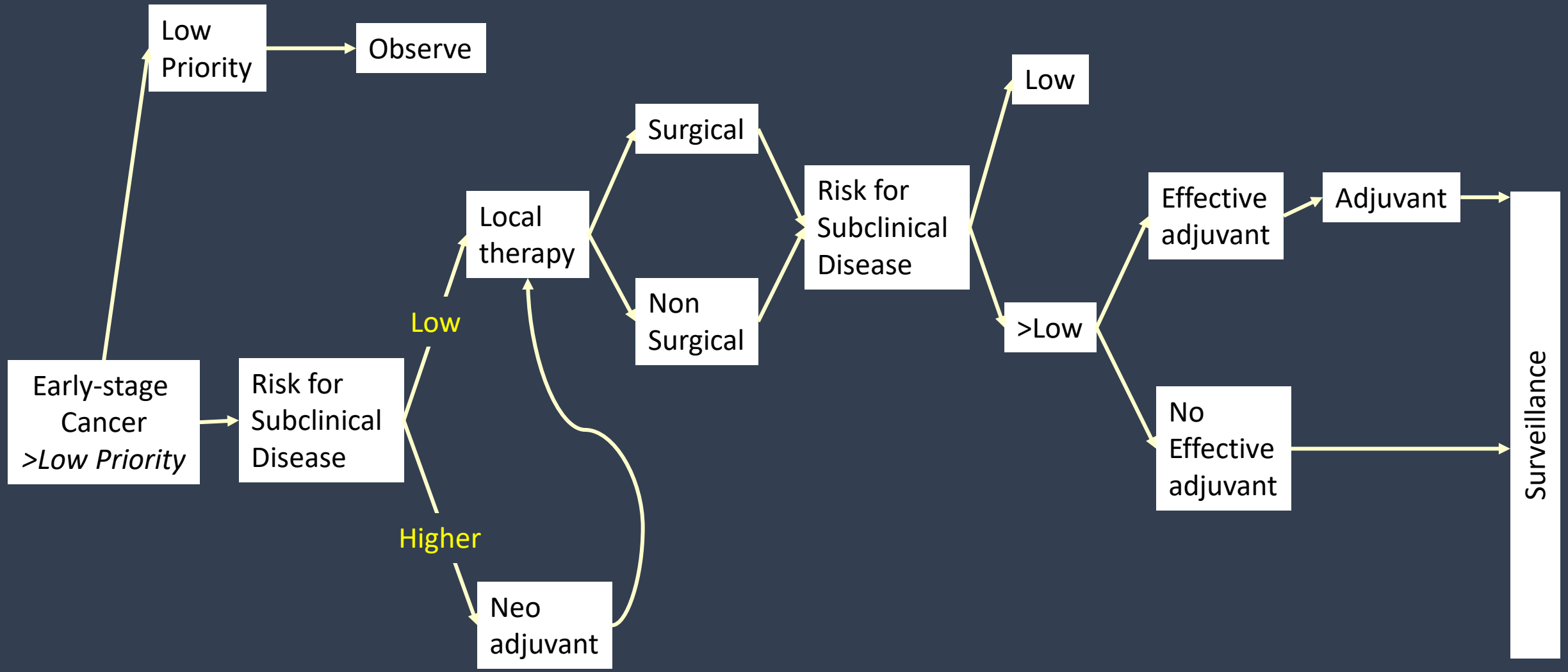
Learning objectives

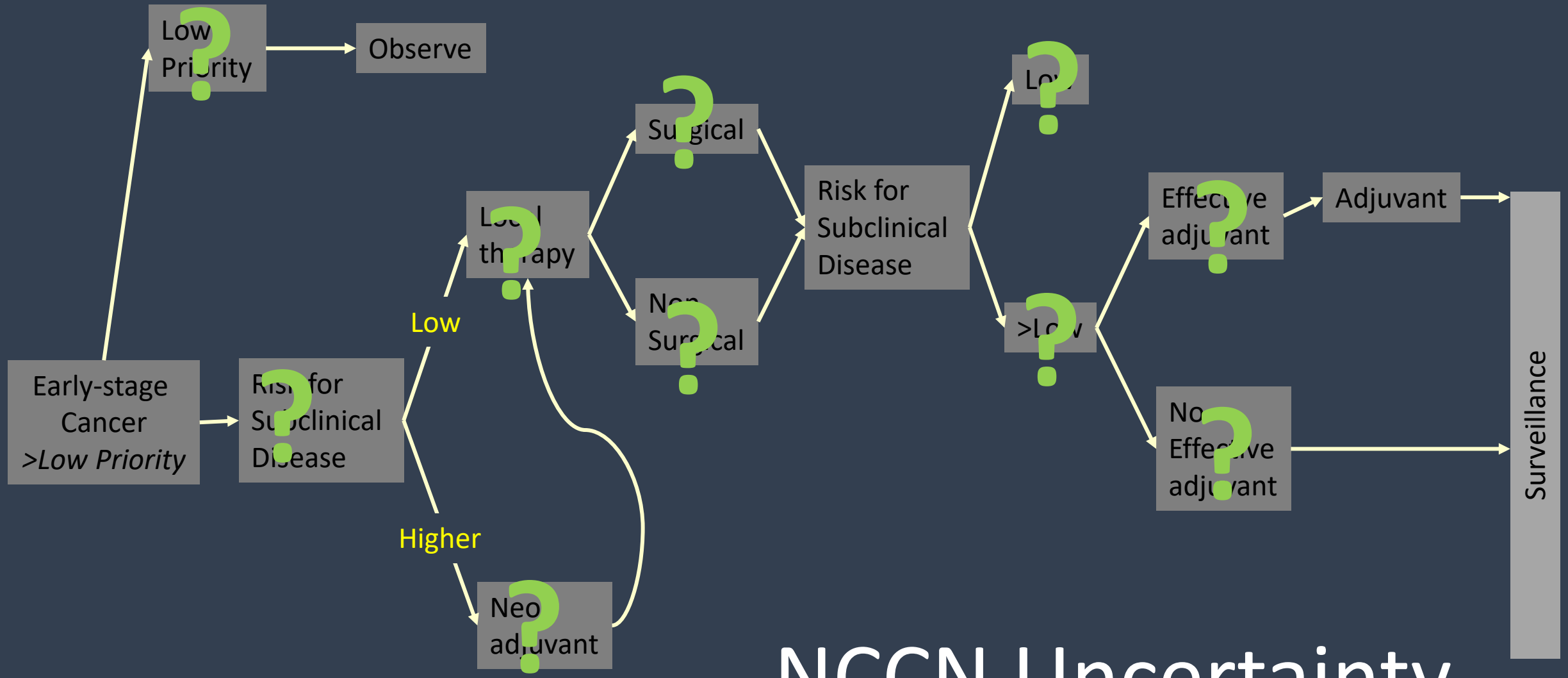
- Outline approach to estimating components of survival risk in early stage cancer
- Establish treatment priorities according to presumed survival threat
- Analyze possible ways that big data can inform big models to personalize care

Disclaimers

- Not an AI expert (but know a lot of their words)
- Frame a lot of my points around lung cancer

Early-stage Cancer
Surveillance


Low Priority
Observe
Early-stage Cancer
>Low Priority
Risk for Subclinical Disease
Low
Higher
Local therapy
Surgical
Non Surgical
Neo adjuvant
Risk for Subclinical Disease
Low
>Low
Effective adjuvant
Adjuvant
No Effective adjuvant
Surveillance



NCCN Uncertainty

Driver of Your Demise

- Locoregional cancer
- Systemic progression
- Other

Knowable?

Driver of Your Demise

- Locoregional cancer
- Systemic progression
- Other

Knowable?

Modifiable?

Driver of Demise

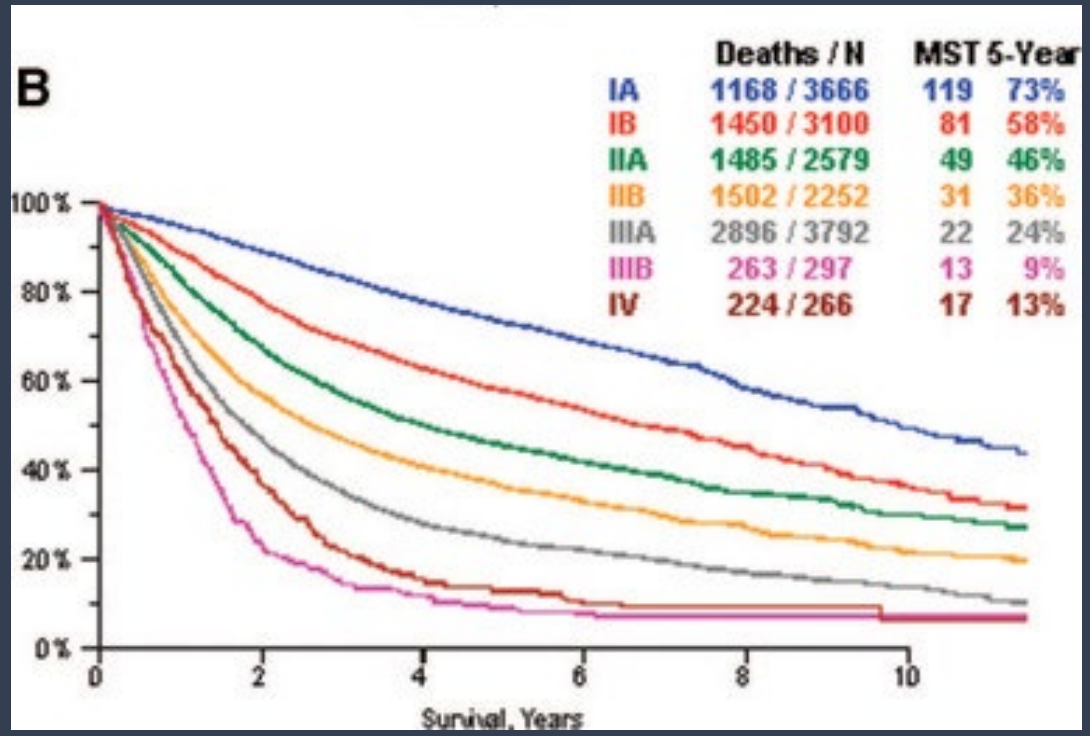
IASLC

INTERNATIONAL ASSOCIATION FOR THE STUDY OF LUNG CANCER

Lung Cancer TNM Stages–9th Edition

Stage Groups of the 9th Edition of the Tumor, Node, Metastasis (TNM) Classification of Lung Cancer

9th Edition TNM Descriptors and Stages						
				N2		
T/M	Categories and Descriptors	N0	N1	N2a	N2b	N3
T1	T1a ≤1 cm	IA1	IIA	IIB	IIIA	IIIB
	T1b >1 to ≤2 cm	IA2	IIA	IIB	IIIA	IIIB
	T1c >2 to ≤3 cm	IA3	IIA	IIB	IIIA	IIIB
T2	T2a Visceral pleura / central invasion	IB	IIB	IIIA	IIIB	IIIB
	T2a >3 to ≤4 cm	IB	IIB	IIIA	IIIB	IIIB
	T2b >4 to ≤5 cm	IIA	IIB	IIIA	IIIB	IIIB
T3	T3 >5 to ≤7 cm	IIB	IIIA	IIIA	IIIB	IIIC
	T3 Invasion	IIB	IIIA	IIIA	IIIB	IIIC
	T3 Same lobe separate tumor nodules	IIB	IIIA	IIIA	IIIB	IIIC
T4	T4 >7 cm	IIIA	IIIA	IIIB	IIIB	IIIC
	T4 Invasion	IIIA	IIIA	IIIB	IIIB	IIIC
	T4 Ipsilateral separate tumor nodules	IIIA	IIIA	IIIB	IIIB	IIIC
M1	M1a Contralateral tumor nodules	IVA	IVA	IVA	IVA	IVA
	M1a Pleural / pericardial effusion, nodules	IVA	IVA	IVA	IVA	IVA
	M1b Single extrathoracic metastasis	IVA	IVA	IVA	IVA	IVA
	M1c1 Multiple metastases in 1 organ system	IVB	IVB	IVB	IVB	IVB
	M1c2 Multiple metastases in >1 organ systems	IVB	IVB	IVB	IVB	IVB



Driver of Demise

Early stage

Driver = *local cancer*

Late stage

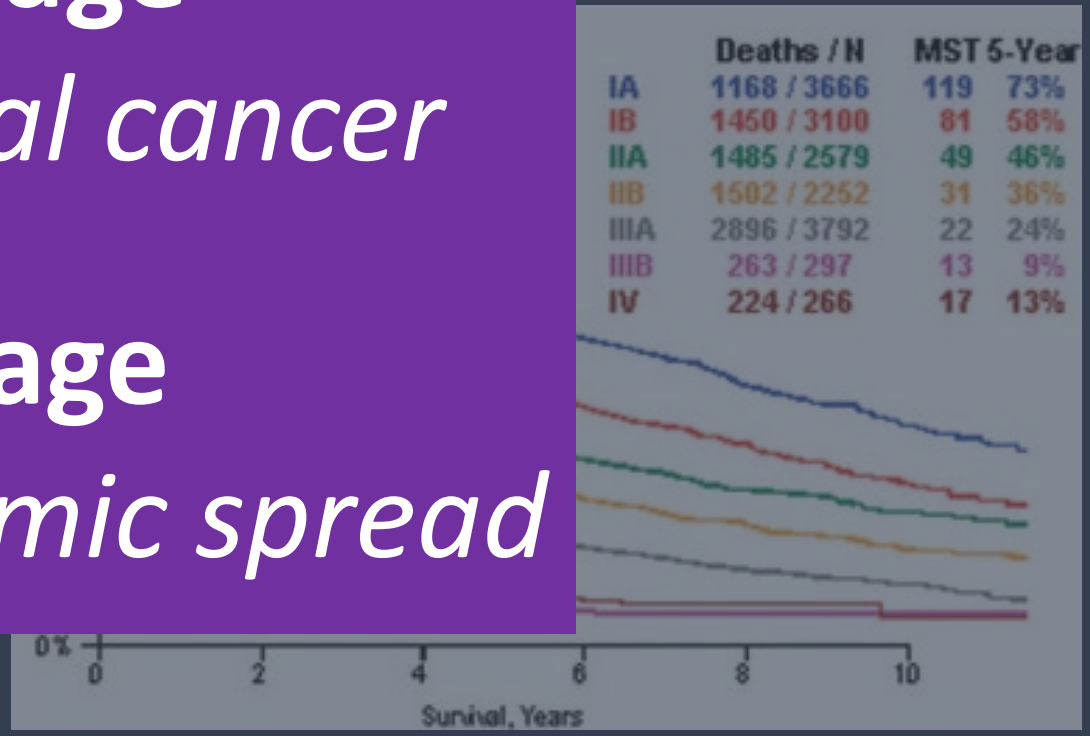
Driver = *systemic spread*

IASLC INTERNATIONAL ASSOCIATION FOR THE STUDY OF LUNG CANCER

Lung Cancer

Stage Groups of the 9th Edition of the Tum
Classification of Lung Cancer

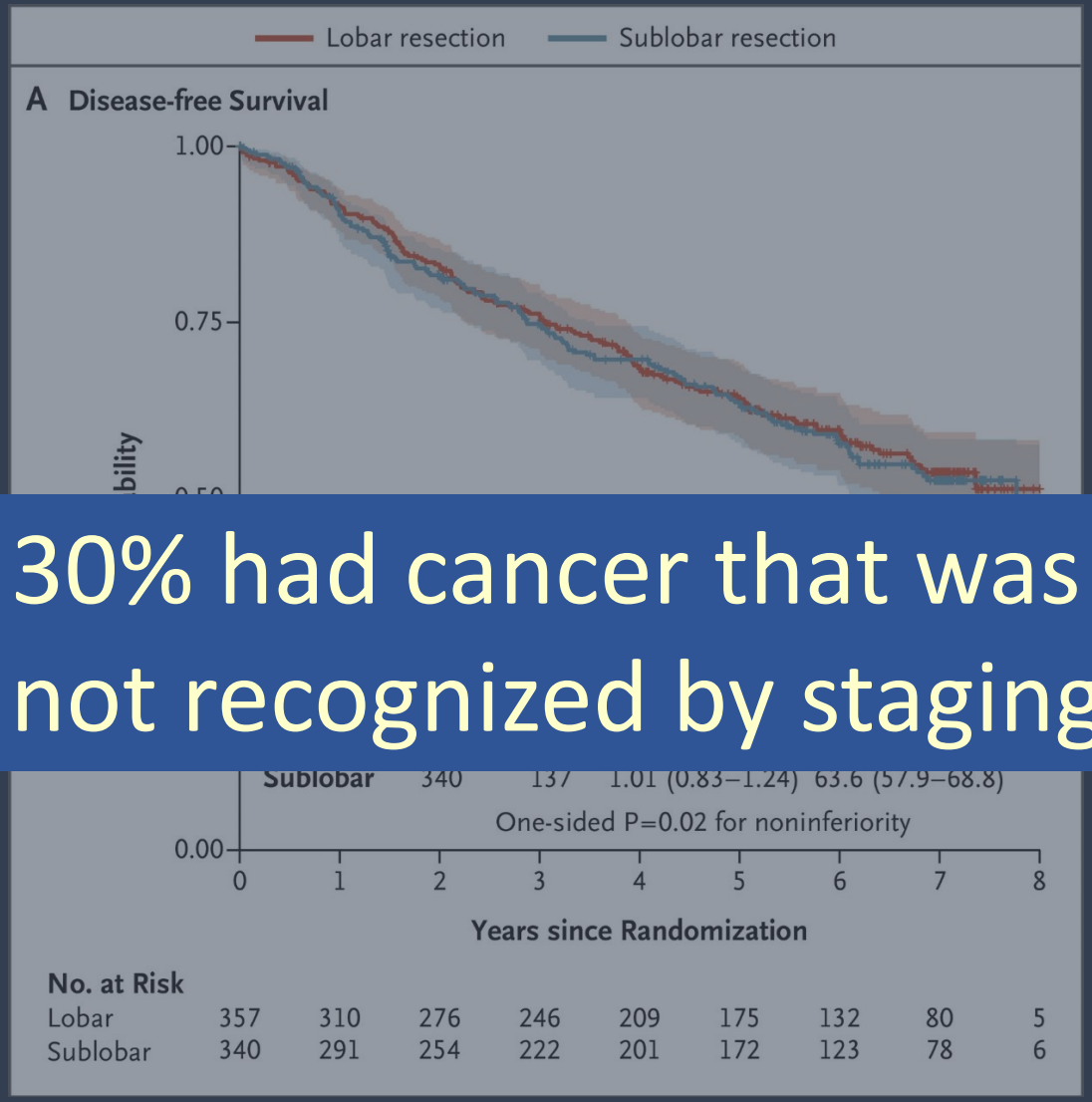
9th Edition TNM Descriptors and Stages						
T/M	Categories and Descriptors					
T1	T1a ≤1 cm					
	T1b >1 to ≤2 cm					
	T1c >2 to ≤3 cm					
T2	T2a Visceral pleura / central invasion					
	T2a >3 to ≤4 cm					
	T2b >4 to ≤5 cm					
T3	T3 >5 to ≤7 cm					
	T3 Invasion					
	T3 Same lobe separate tumor nodules					
T4	T4 >7 cm					
	T4 Invasion					
	T4 Ipsilateral separate tumor nodules					
M1	M1a Contralateral tumor nodules	IVA	IVA	IVA	IVA	IVA
	M1a Pleural / pericardial effusion, nodules	IVA	IVA	IVA	IVA	IVA
	M1b Single extrathoracic metastasis	IVA	IVA	IVA	IVA	IVA
	M1c1 Multiple metastases in 1 organ system	IVB	IVB	IVB	IVB	IVB
	M1c2 Multiple metastases in >1 organ systems	IVB	IVB	IVB	IVB	IVB



Early-Stage Lung Cancer

- Driver = local disease
- Strategy = local therapy to primary
- Strategy = look for predictors of spread*

Disease Free Survival (cure) Early Lung Ca



30% had cancer that was not recognized by staging

Altorki NEJM 2023

Driver of Your Demise

- Locoregional cancer
- Systemic progression
- Other

Knowable?

Standard 5.8

3 mediastinal nodes

1 hilar node

Driver of Your Demise

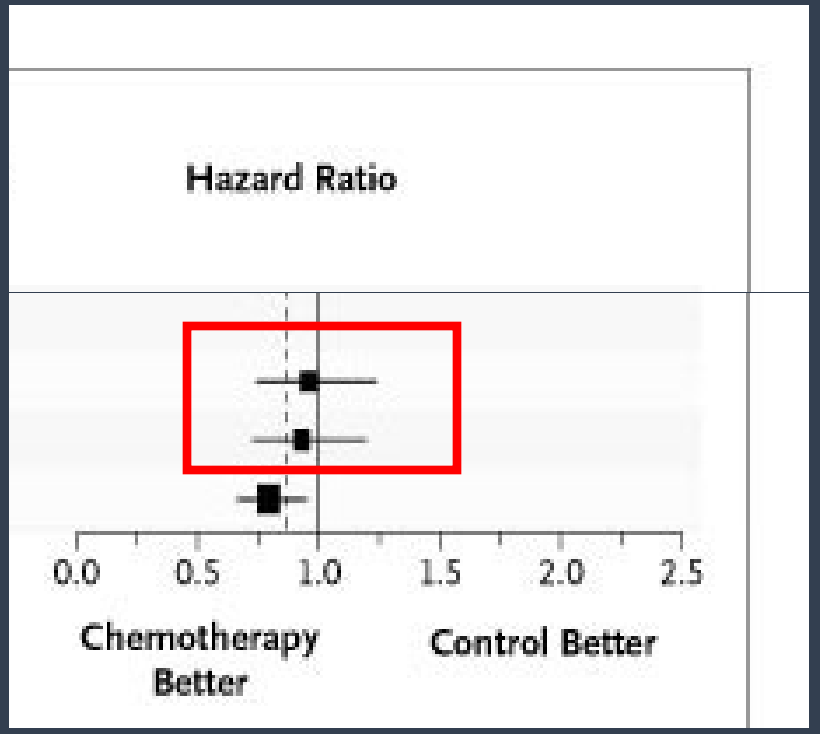
- Locoregional cancer
- Systemic progression
- Other

Knowable?

Modifiable?

Unselected adjuvant chemotherapy

Stage I
Stage II
Stage III

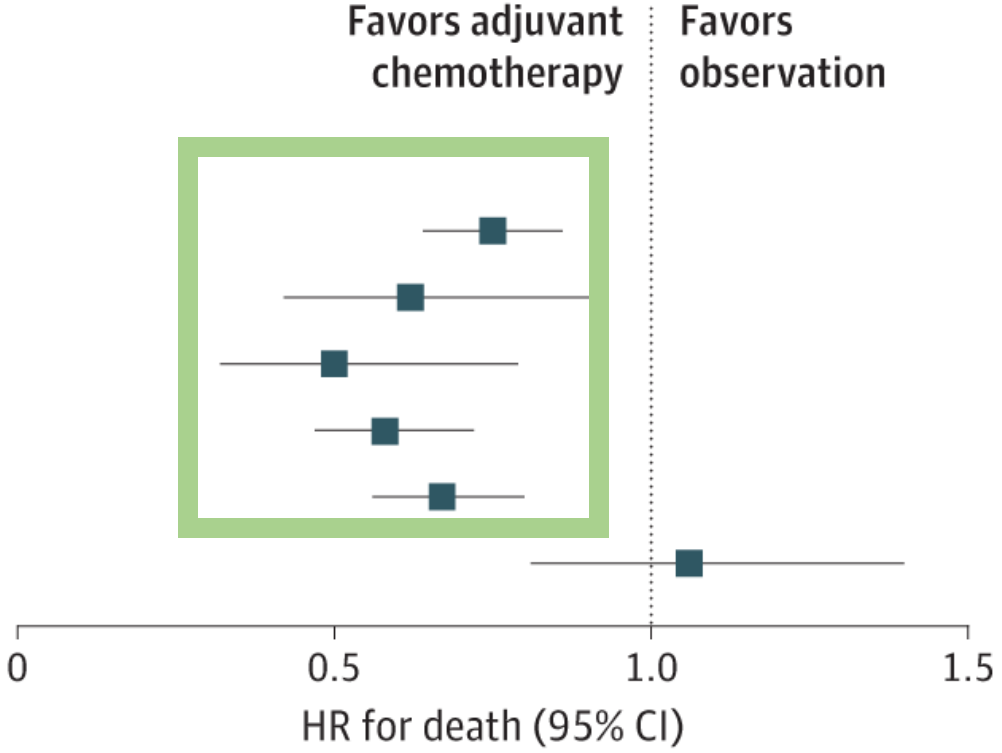


Early-stage NSCLC - NOT helped

IALT trial NEJM 2004

Addition of therapy to high-risk features

Tumor size (stage)	HR for death (95% CI)	P value
>4-5 cm (IIA)		
Overall	0.75 (0.64-0.86)	<.001
LVI	0.62 (0.42-0.91)	.005
VPI	0.50 (0.32-0.79)	<.001
High grade	0.58 (0.47-0.72)	<.001
≥1 Risk factor	0.67 (0.56-0.80)	.02
0 Risk factors[a]	1.06 (0.81-1.40)	.37

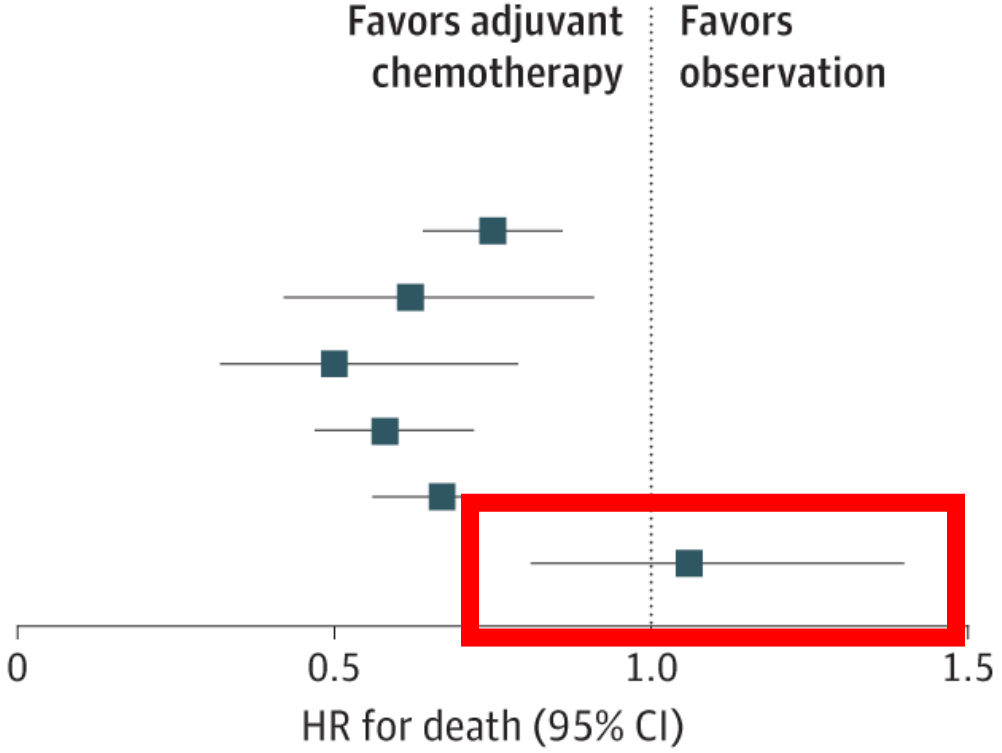


Adjuvant cytotoxic to *high-risk feature*

Pathak JAMA onc 2020

Addition of therapy to high-risk features

Tumor size (stage)	HR for death (95% CI)	P value
>4-5 cm (IIA)		
Overall	0.75 (0.64-0.86)	<.001
LVI	0.62 (0.42-0.91)	.005
VPI	0.50 (0.32-0.79)	<.001
High grade	0.58 (0.47-0.72)	<.001
≥1 Risk factor	0.67 (0.56-0.80)	.02
0 Risk factors[a]	1.06 (0.81-1.40)	.37

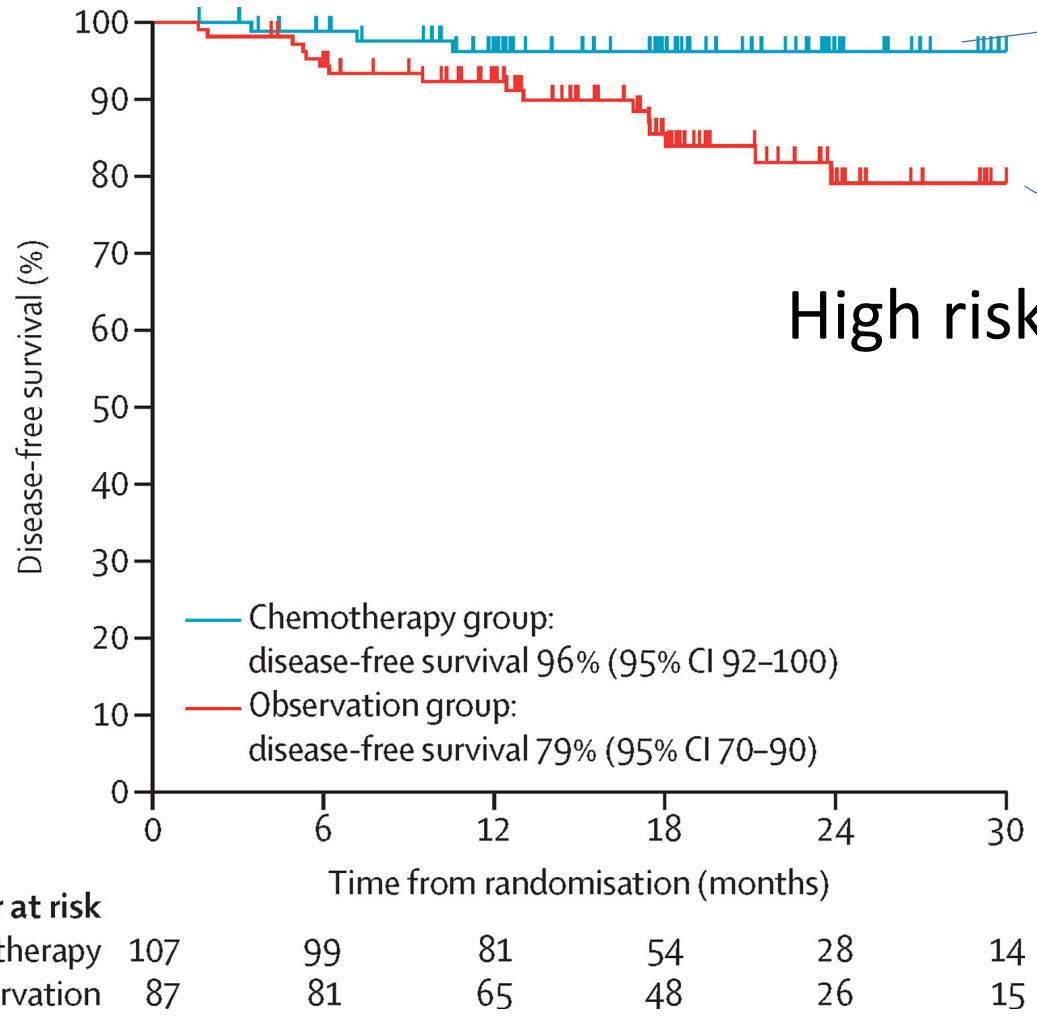


Pathak JAMA onc 2020

Aim High – 14 gene panel

High Risk, Early-Stage patients

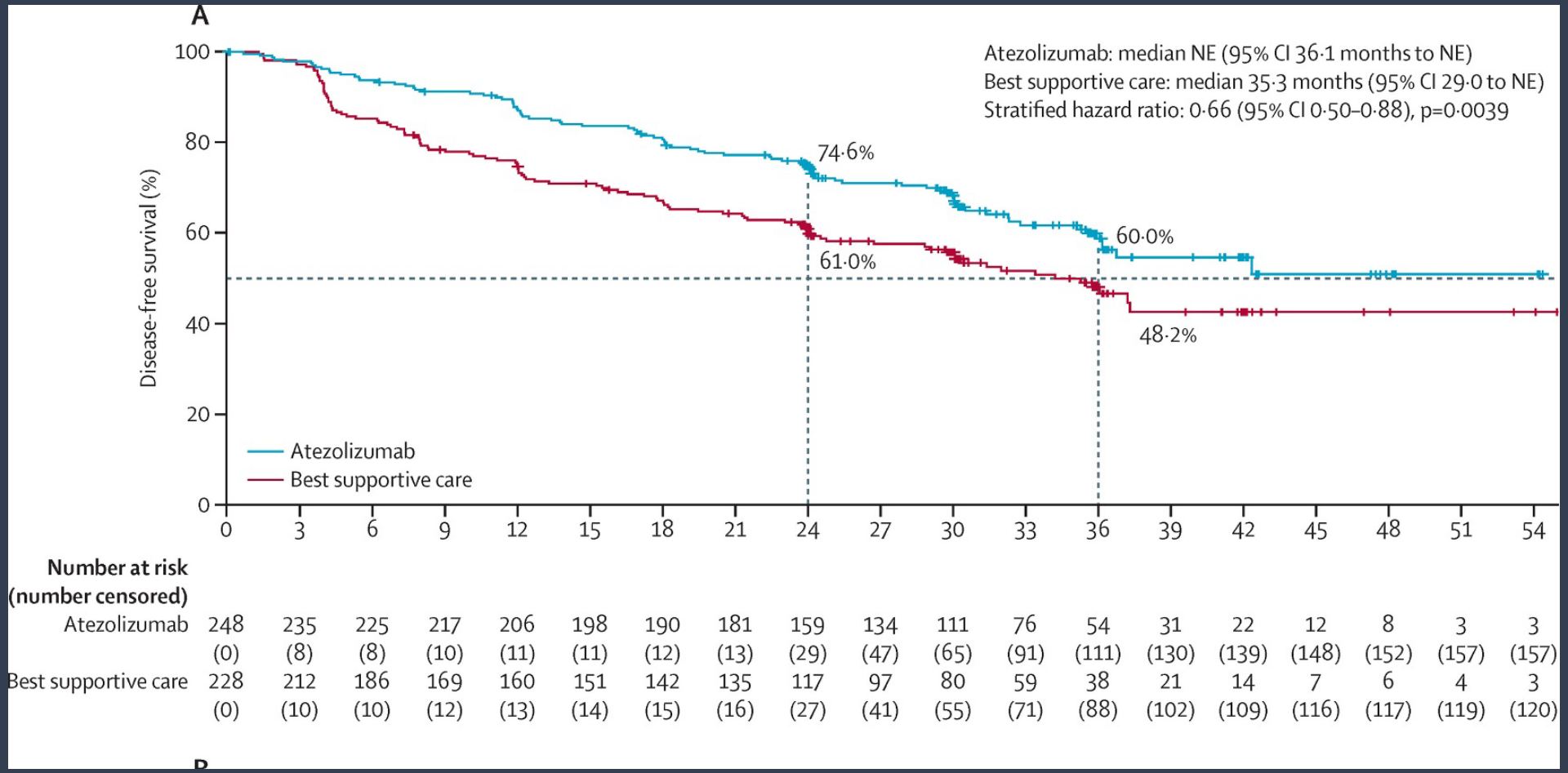
High risk - cytotoxic chemo



High risk - Observation

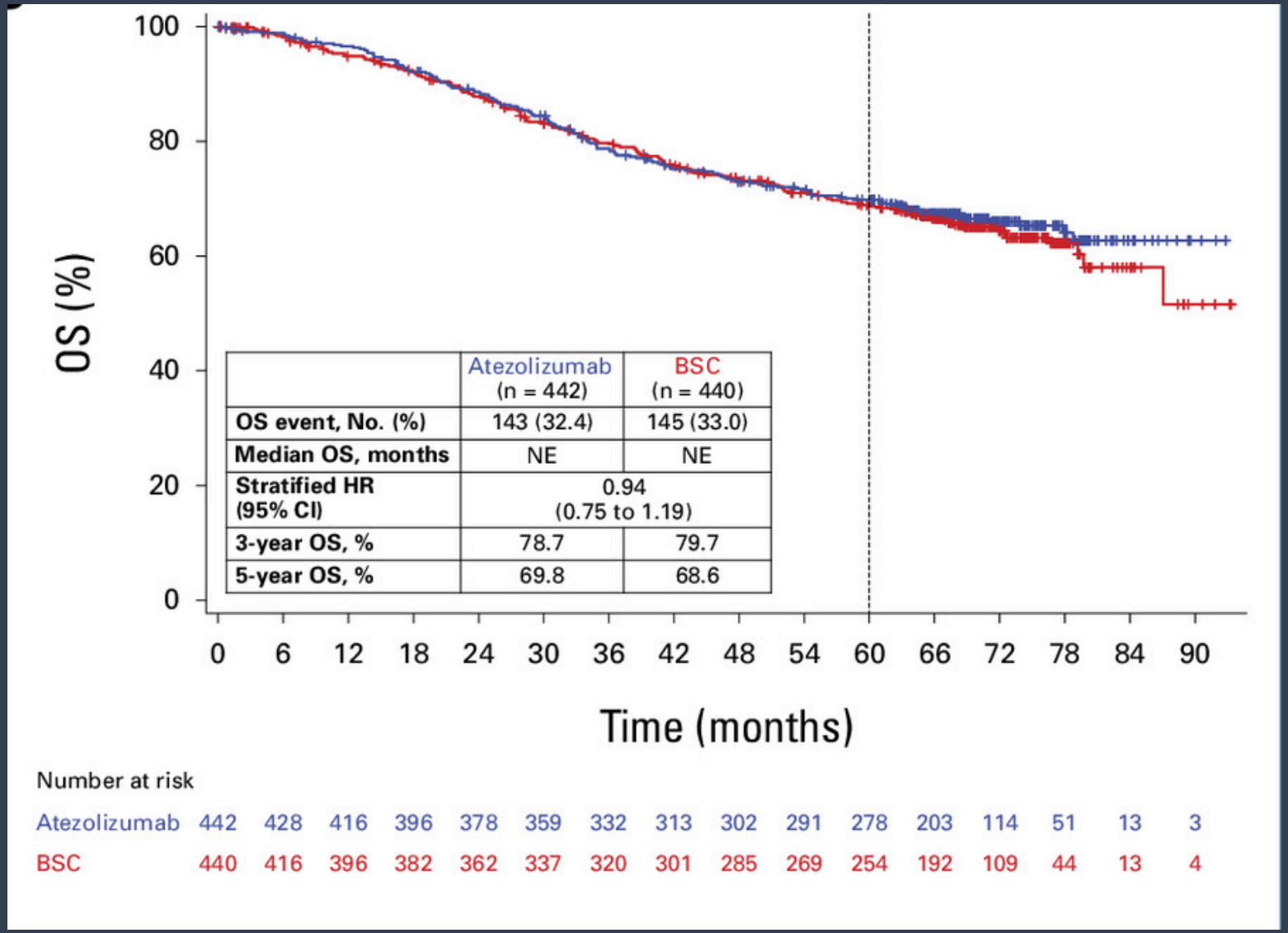
Spigel et al Lancet Resp Med 2025

IMpower010



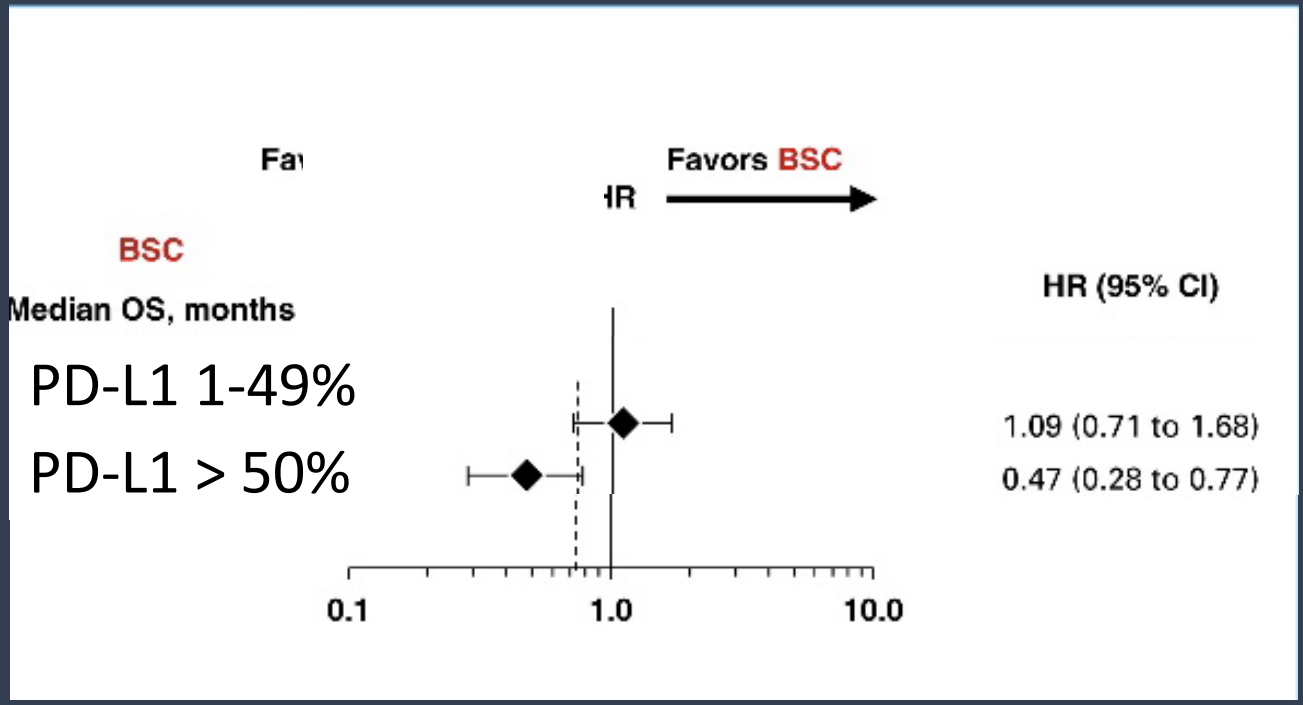
IMpower010 Lancet 2021

Overall Survival



Felip JCO 2025

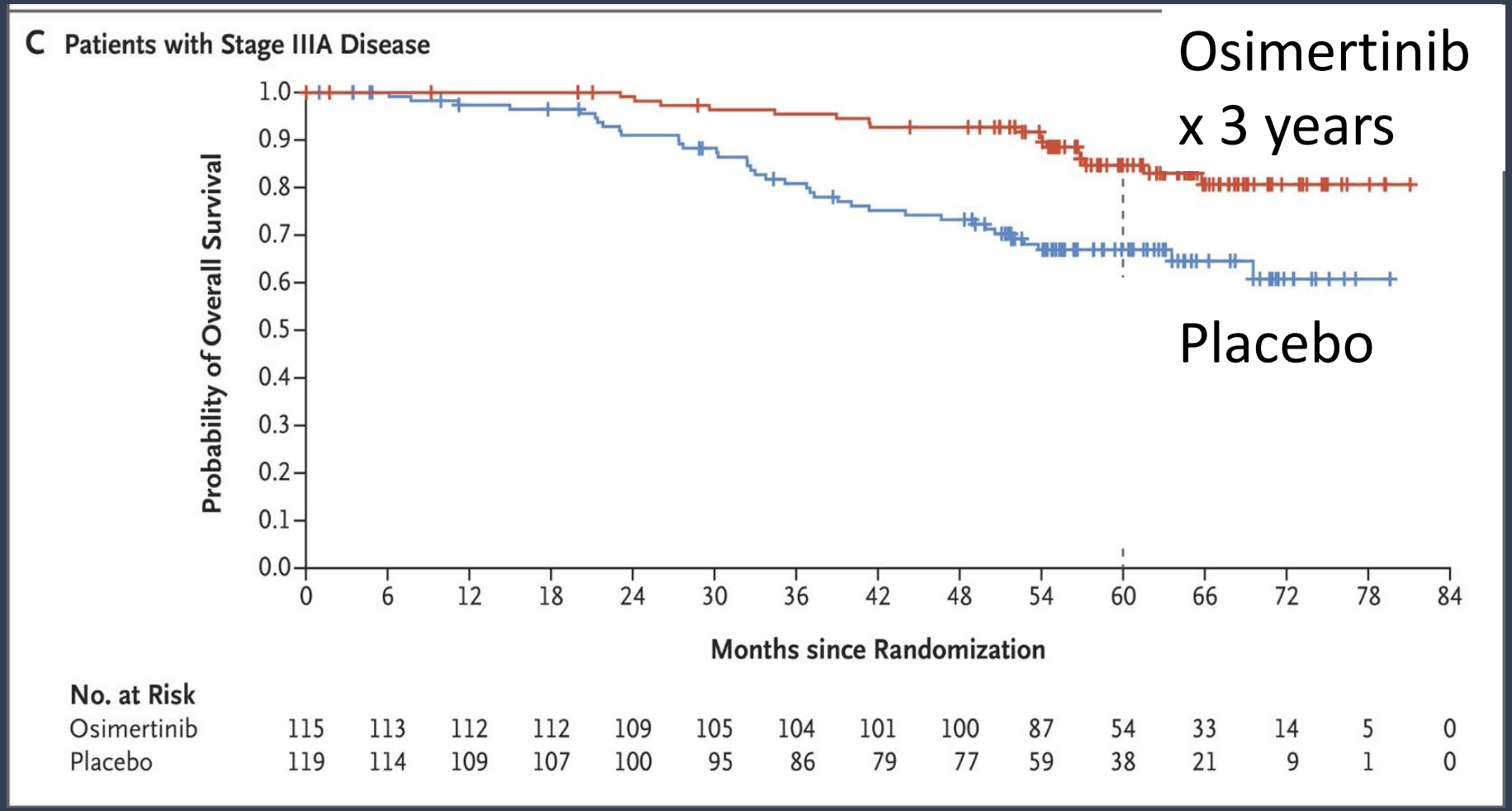
Stage II - III group of NSCLC



Felip JCO 2025

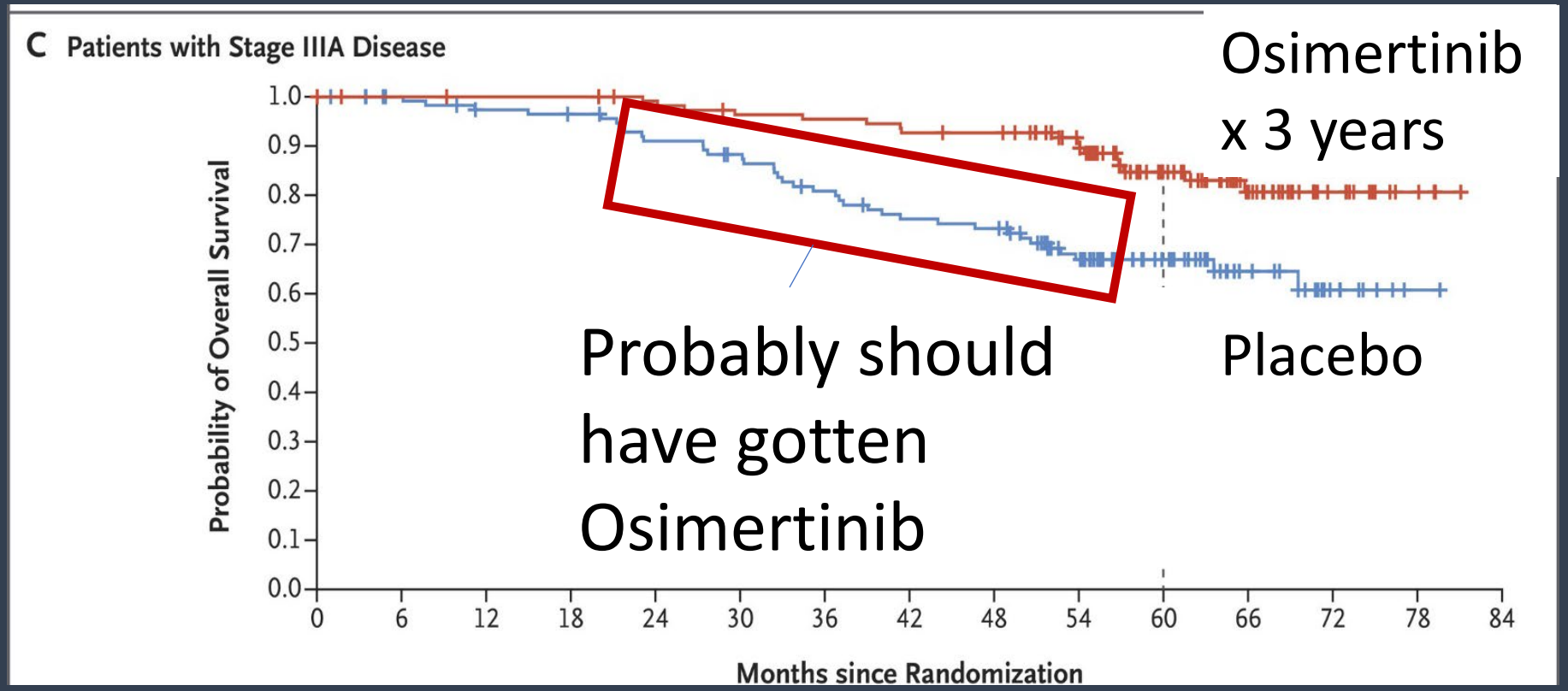
EGFR Mutated NSCLC

Stage III EGFR Mutated Lung Cancer



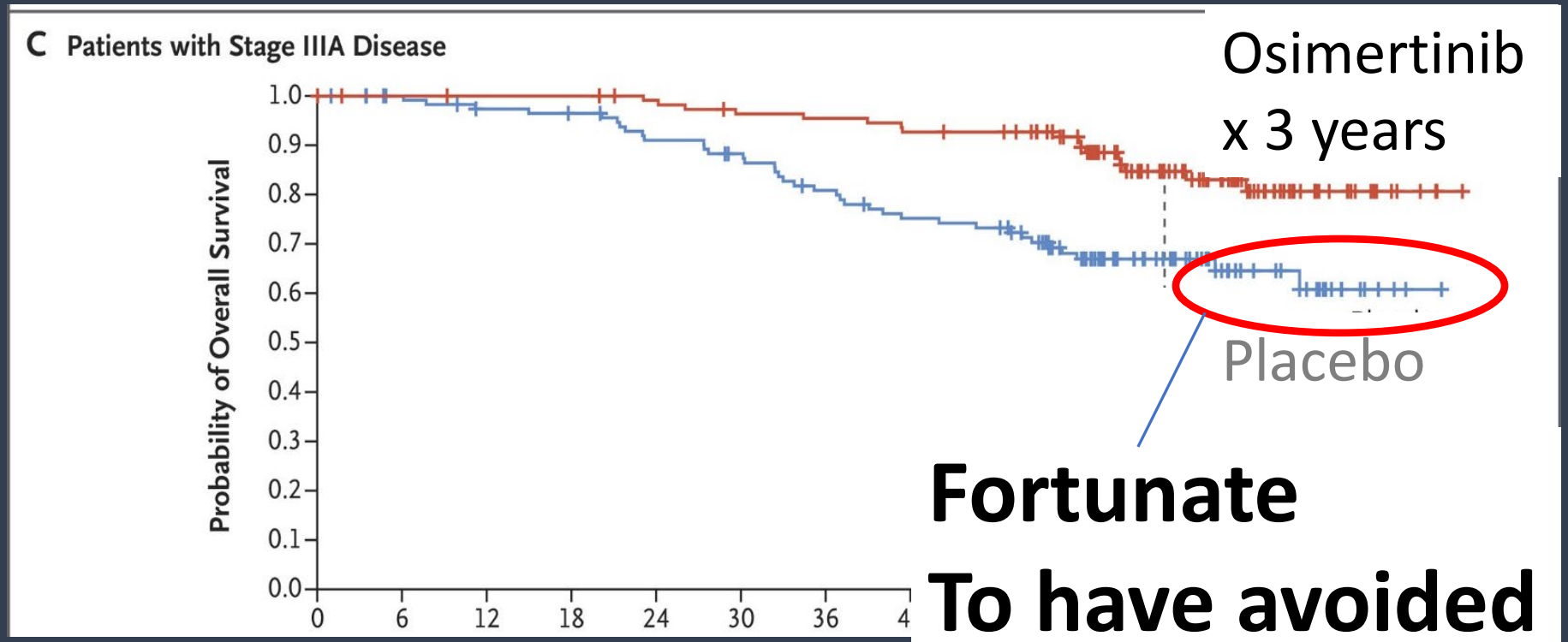
Tsuboi NEJM 2023

Stage III EGFR Mutated Lung Cancer



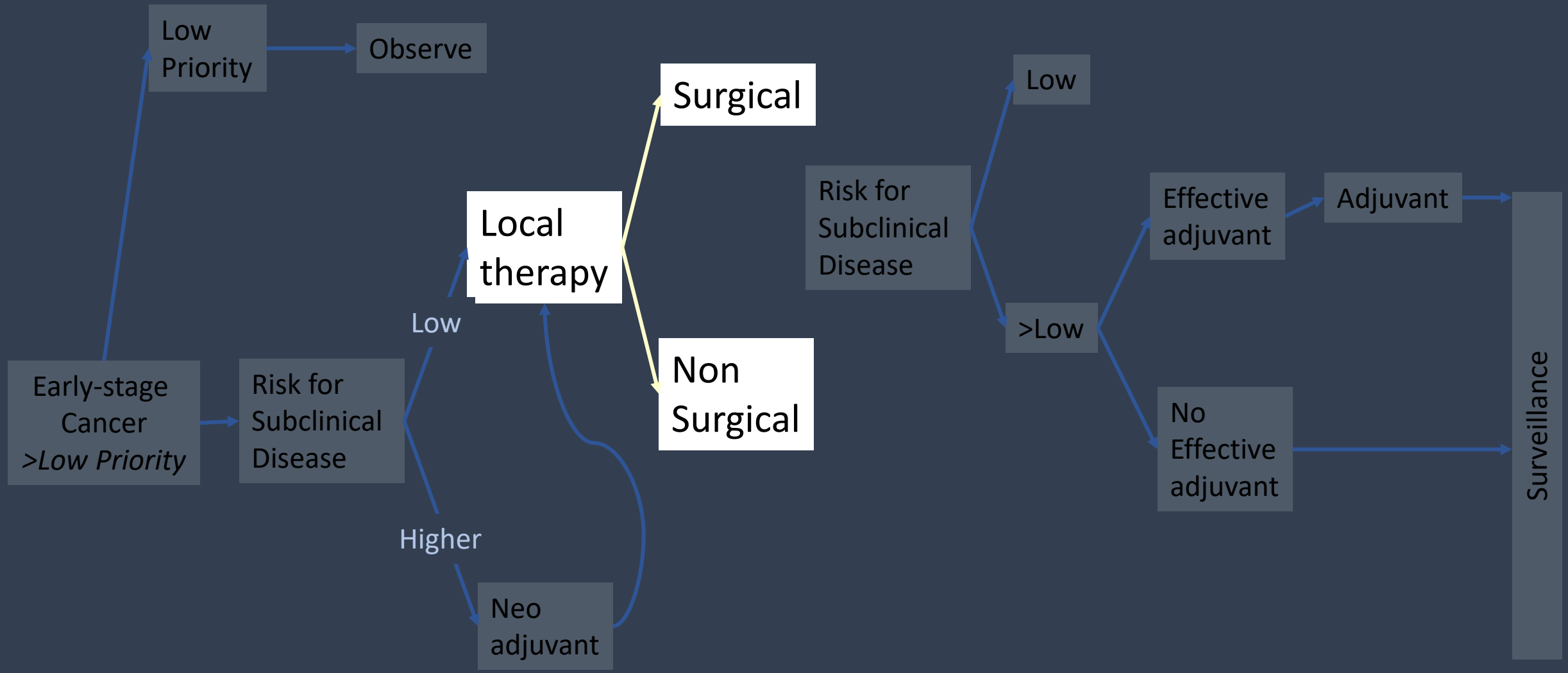
Tsuboi NEJM 2023

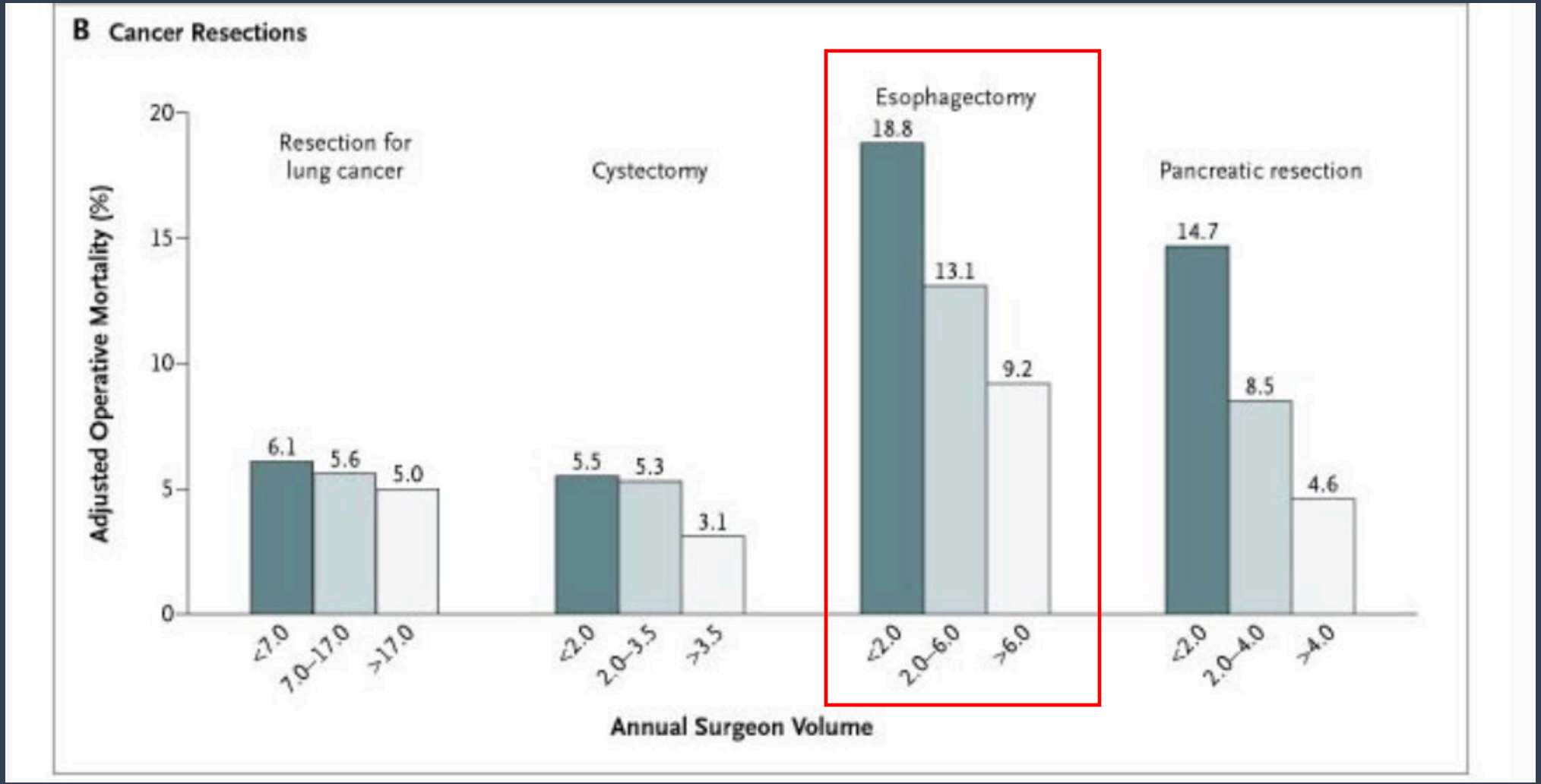
Stage III EGFR Mutated Lung Cancer



Fortunate
To have avoided
Osimertinib 3y?

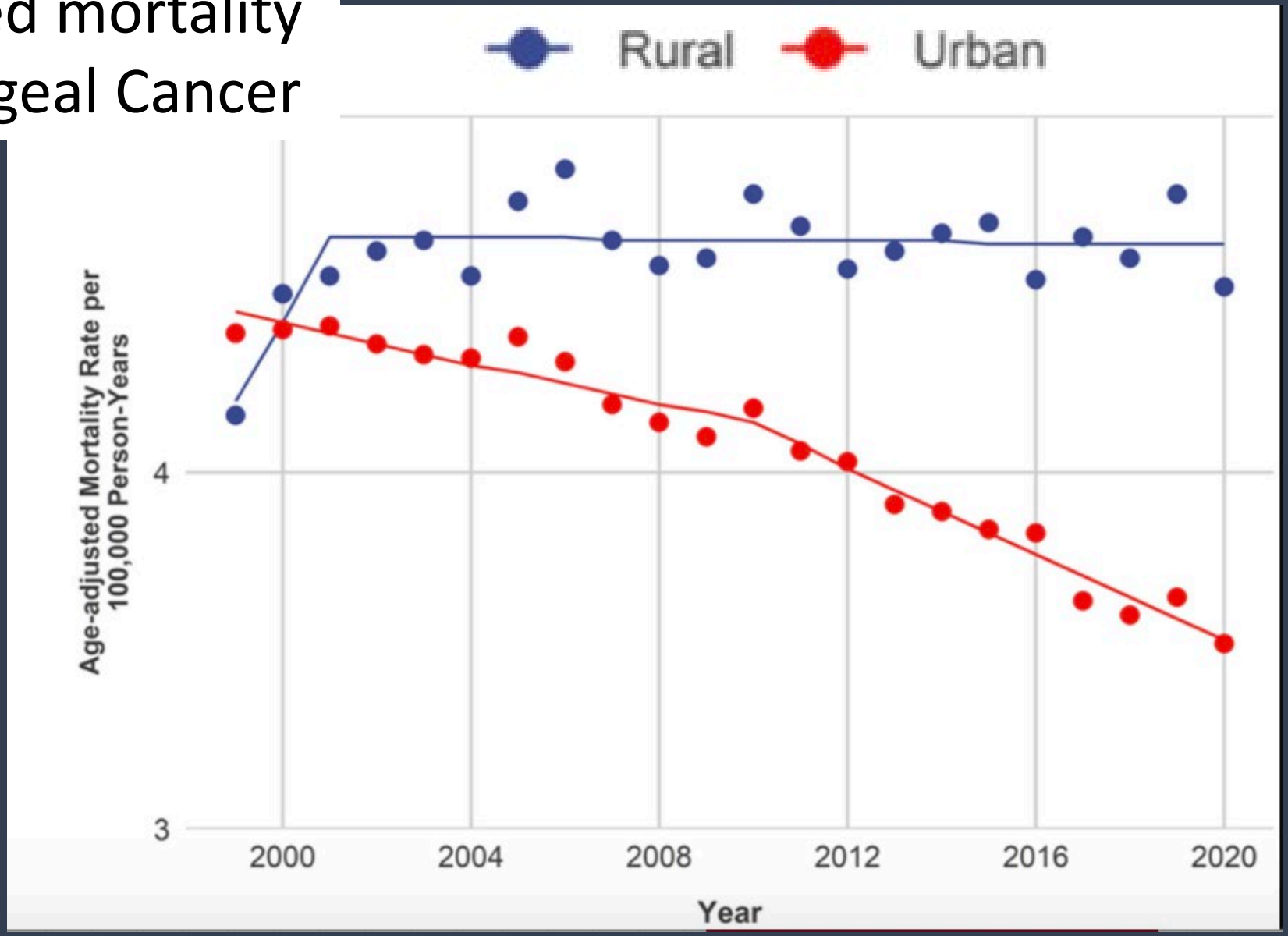
Tsuboi NEJM 2023


Low Priority
Observe
Early-stage Cancer
>Low Priority
Risk for Subclinical Disease
Low
Higher
Local therapy
Surgical
Non Surgical
Neo adjuvant
Risk for Subclinical Disease
Low
>Low
Effective adjuvant
Adjuvant
No Effective adjuvant
Surveillance



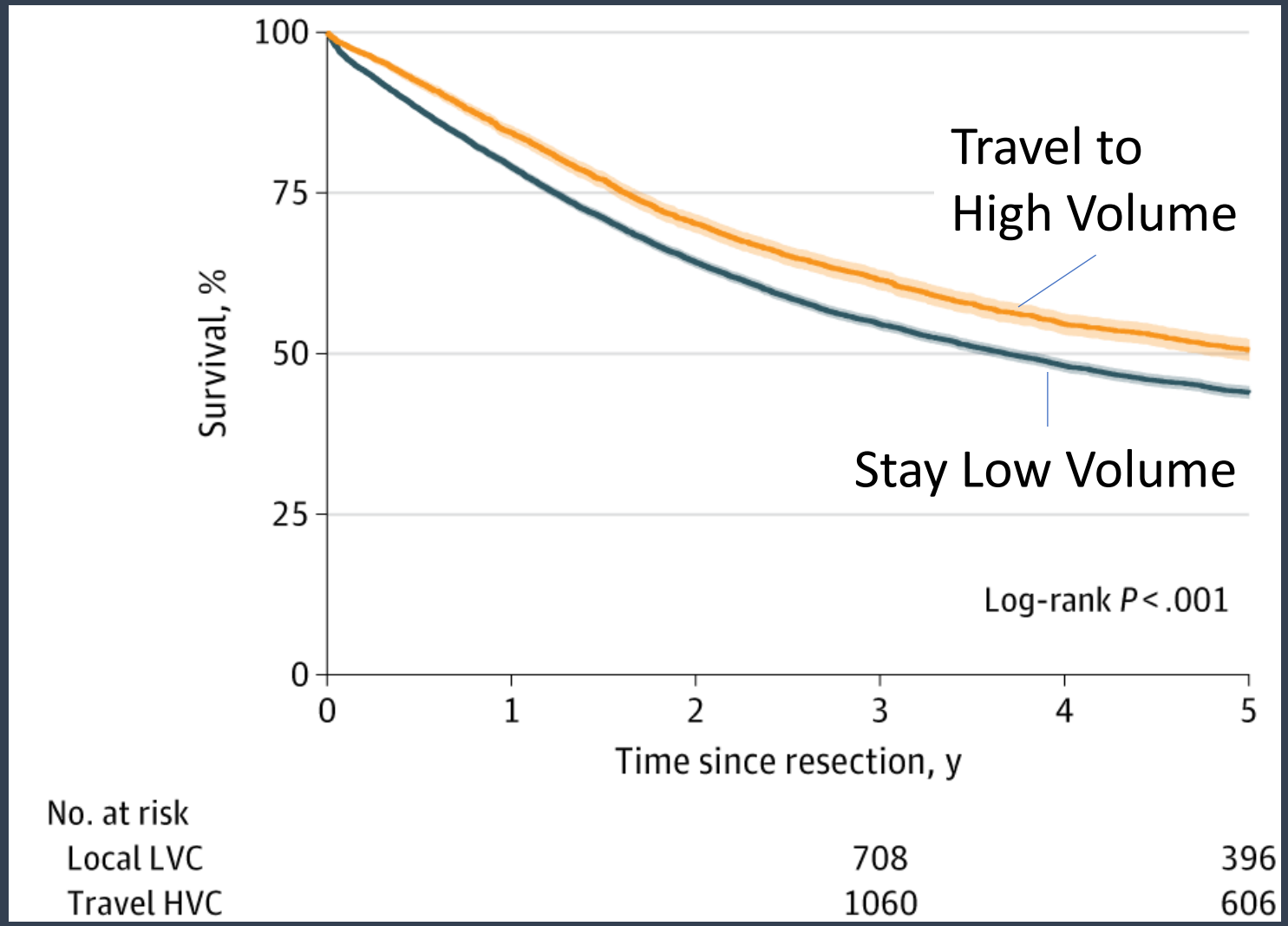
Birkmeyer NEJM 2003

Age adjusted mortality For Esophageal Cancer



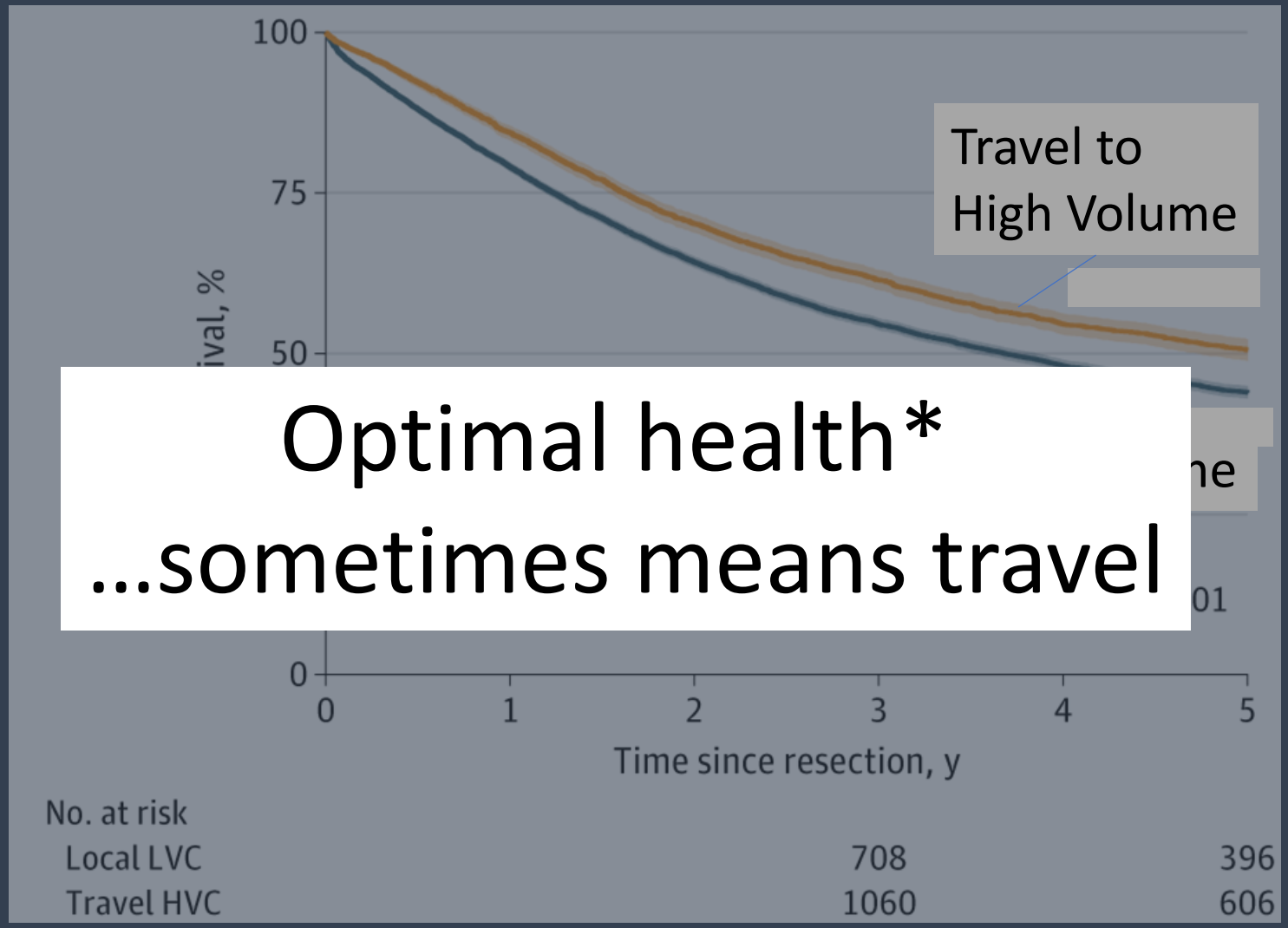
Elnaggar et al Clinics and Research in Hepatology and Gastroenterology 2025

5-year Survival Surgically-managed Esophageal Cancer



Sakowitz JAMA Net Open 2024

5-year Survival Surgically-managed Esophageal Cancer



Optimal health*
…sometimes means travel

Sakowitz JAMA Net Open 2024

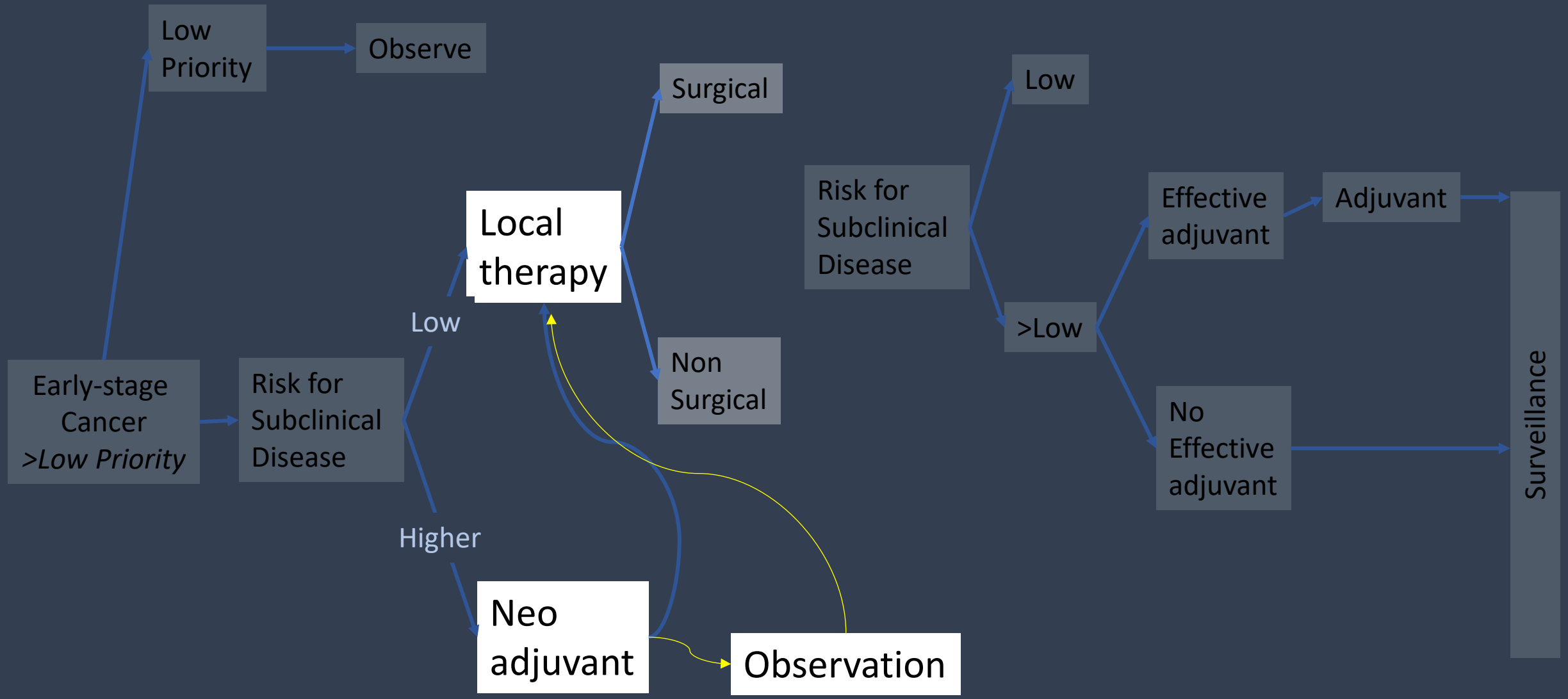
For a significant portion of the US cancer population

Suboptimal care…is their best option

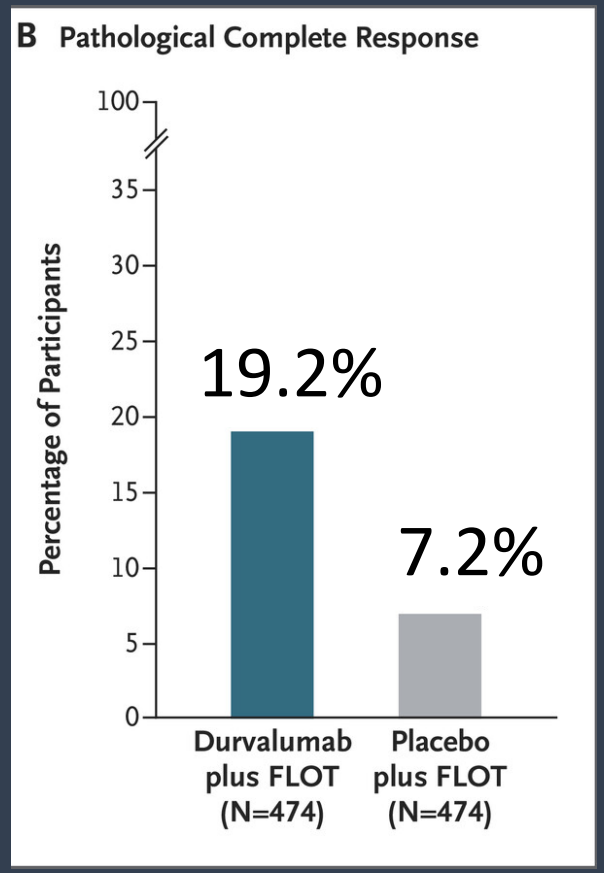
For a significant portion of the US cancer population

Suboptimal care…is their best option

Cannot take if off the table

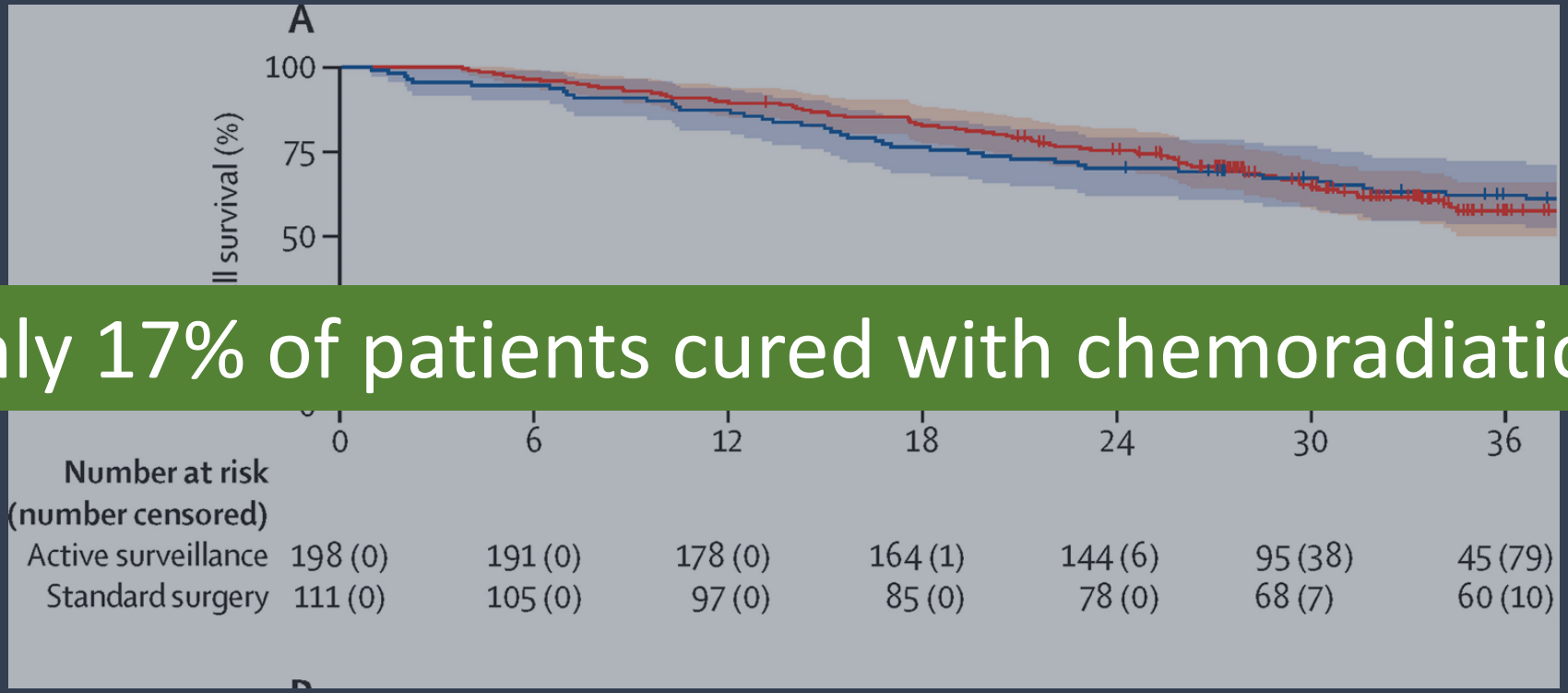

Low Priority
Observe
Early-stage Cancer
>Low Priority
Risk for Subclinical Disease
Low
Higher
Local therapy
Surgical
Non Surgical
Neo adjuvant
Observation
Risk for Subclinical Disease
Low
>Low
Effective adjuvant
Adjuvant
No Effective adjuvant
Surveillance

Pathologic Complete Response to Preoperative Treatment Esoph CA



Jernigan NEJM 2025

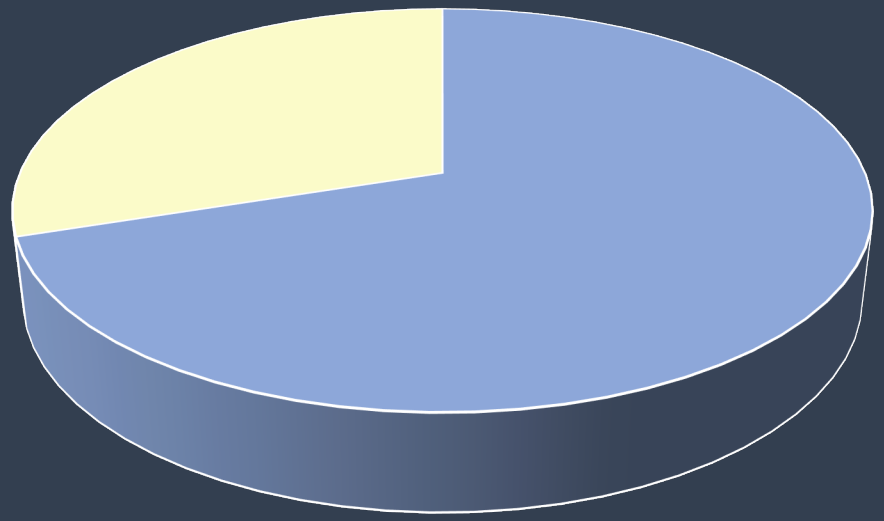
SANO trial Esoph CA complete clinical response active surveillance vs esophagectomy



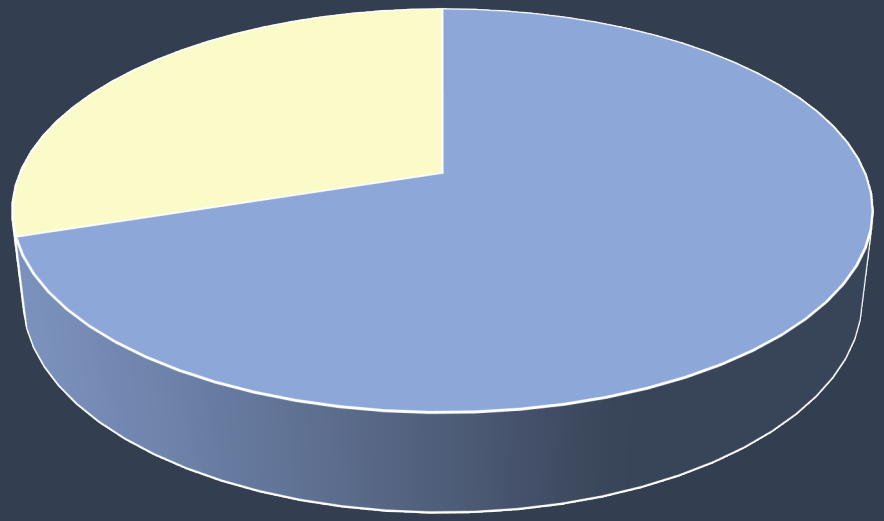
Only 17% of patients cured with chemoradiation

van der Wilk Lancet 2025

Cured After Surgery



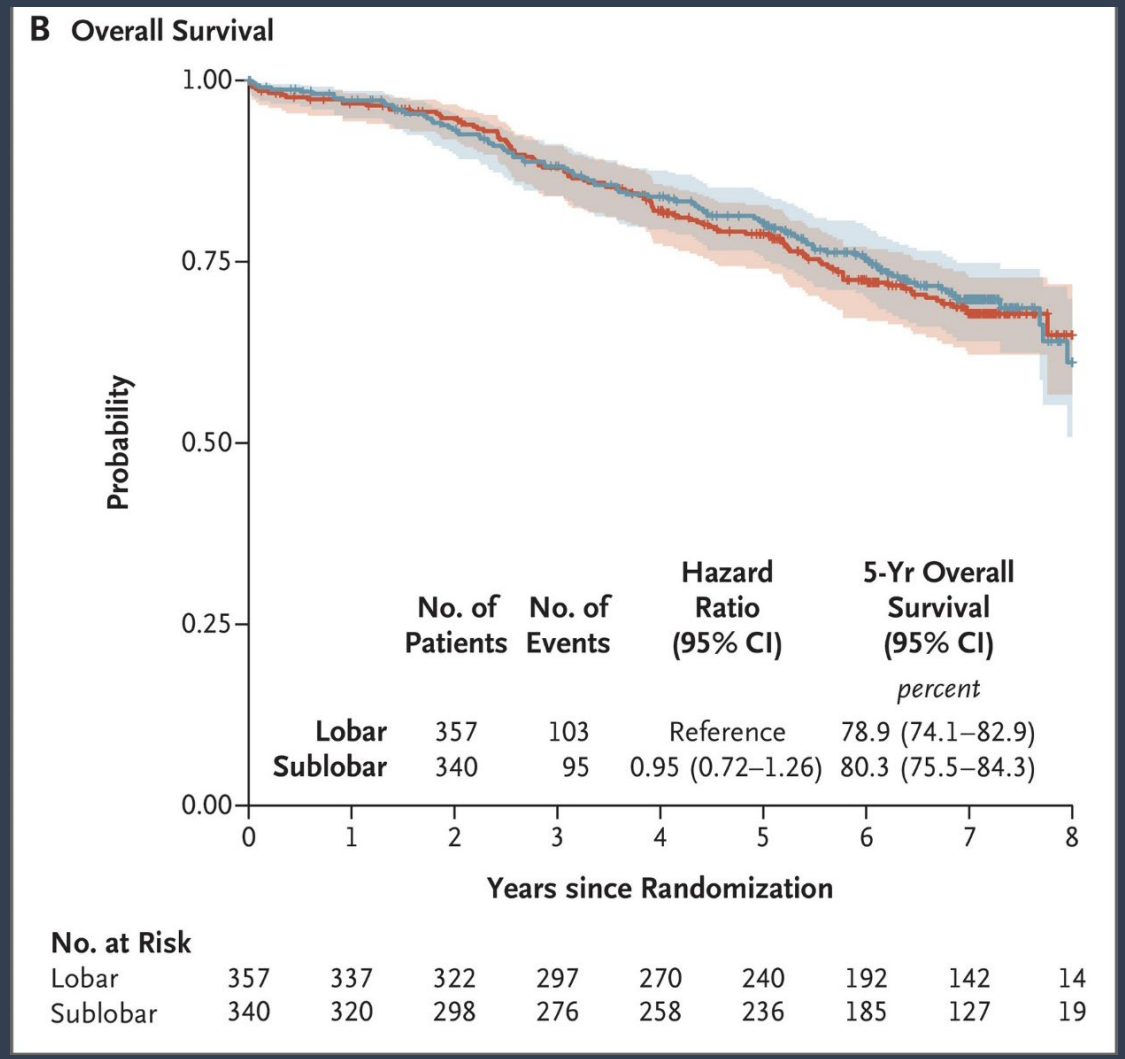
Cured
After Surgery



Curable
Only if surgery

?

Overall Survival Early Lung Ca

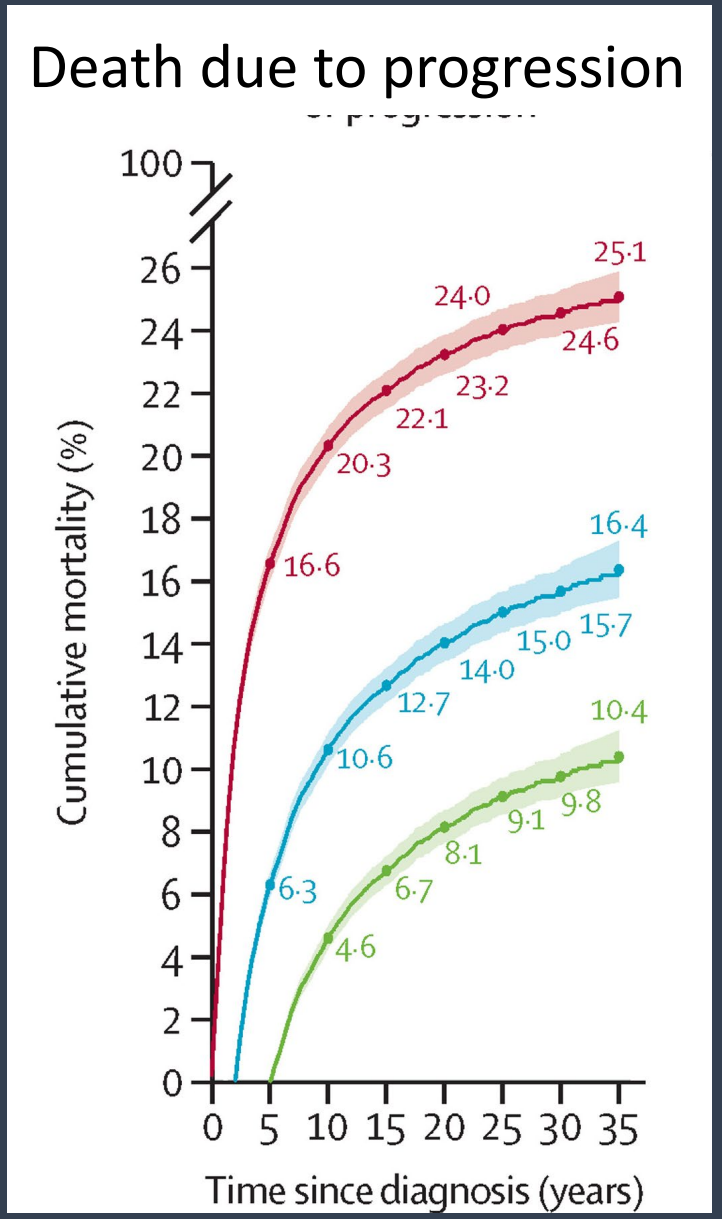


14.5% death from index cancer

13.3% Died of noncancer Related causes

Altorki NEJM 2023

Death due to progression



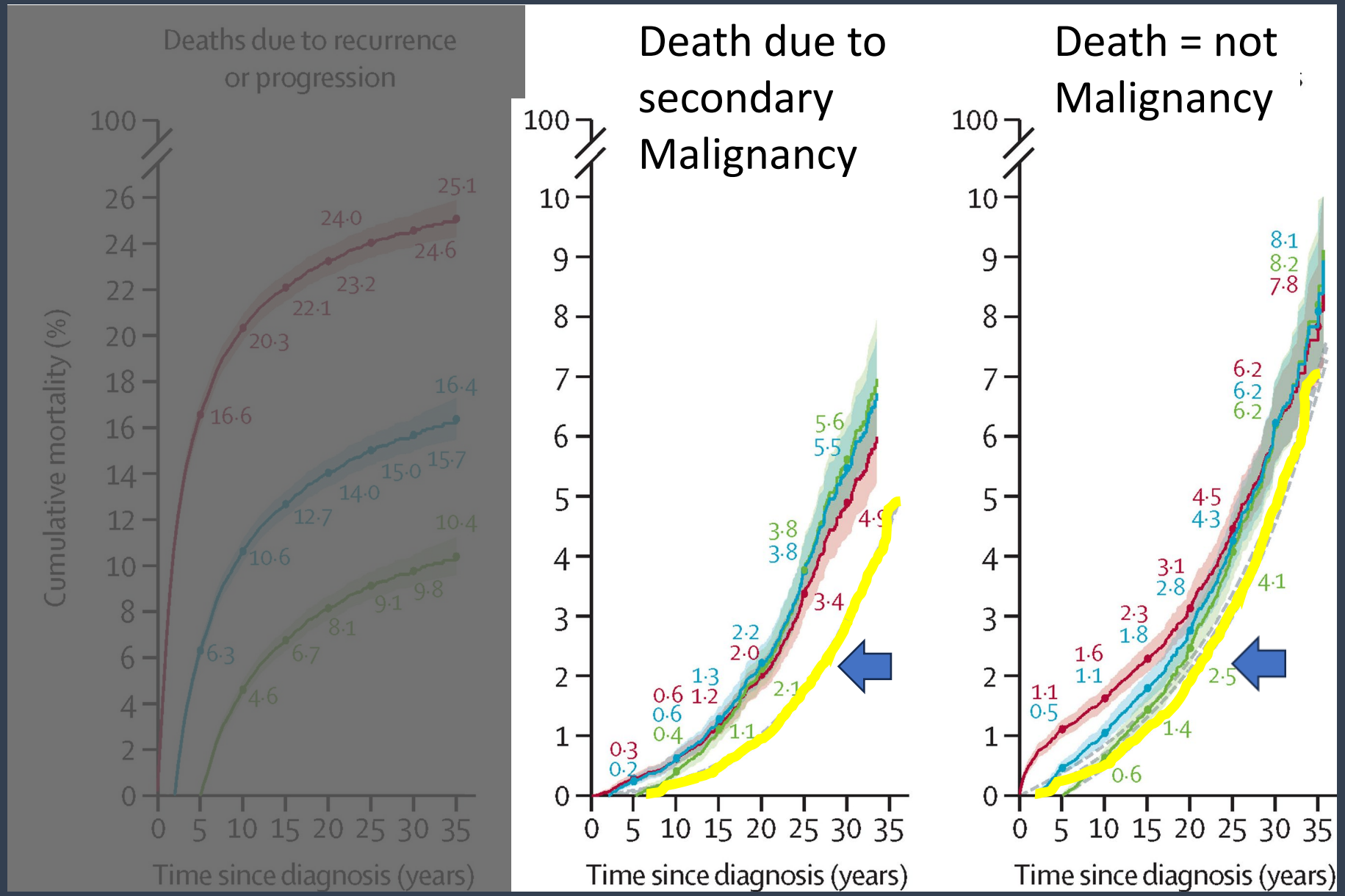
Survivors of AYA cancers in Alberta Canada

Red = all survivors

Blue = 2 year survivors

Green = 5 year survivors

Hughes Lancet public health 2025



Hughes Lancet public health 2025

JAMA Oncology

Research Letter

Smoking Status of the US Cancer Population and a New Perspective From the National Cancer Database

Giorgio Caturegli, MD[1]; Xuan Zhu, MPH[2]; Bryan Palis, MA[2] ;
Timothy W. Mullett, MD, MBA[3]; Benjamin J. Resio, MD[1]; Daniel J. Boffa, MD, MBA[1]

Still Smoking at Cancer Diagnosis

14.7% (234,548) Overall in US Cancer population

34.8% (63,461) Lung Cancer

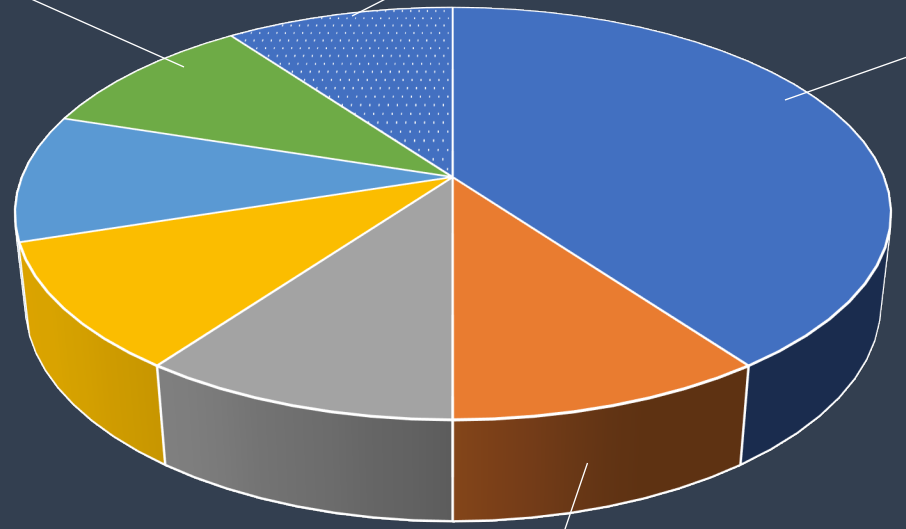
Survivorship Opportunities

- Obesity
- Smoking
- Cancer Screening
- Cholesterol
- Exercise
- Mental health

Eligible for targeted Tx – did not need it

Will die sooner
of something else

Curable ONLY with surgery

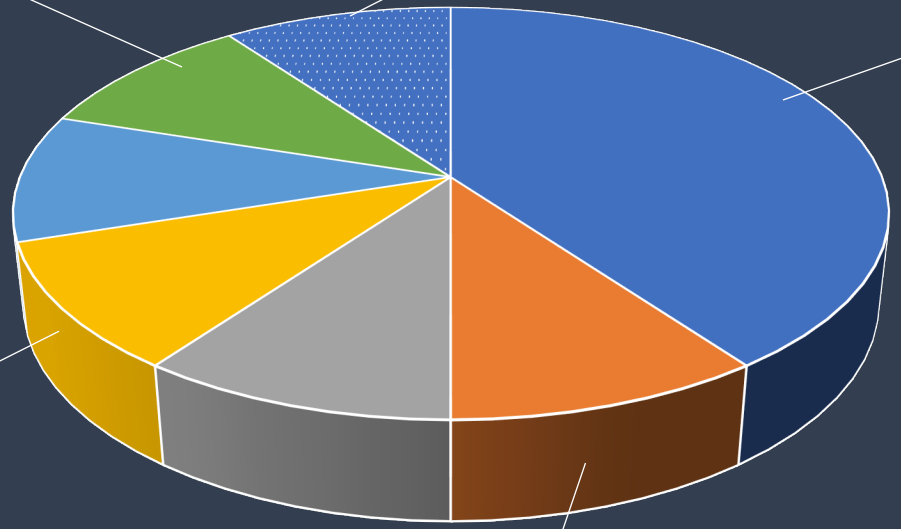


Currently Incurable

Eligible for targeted Tx – did not need it

Will die sooner
of something else

Curable ONLY with surgery



Future Death
Preventable
Cause

Currently Incurable

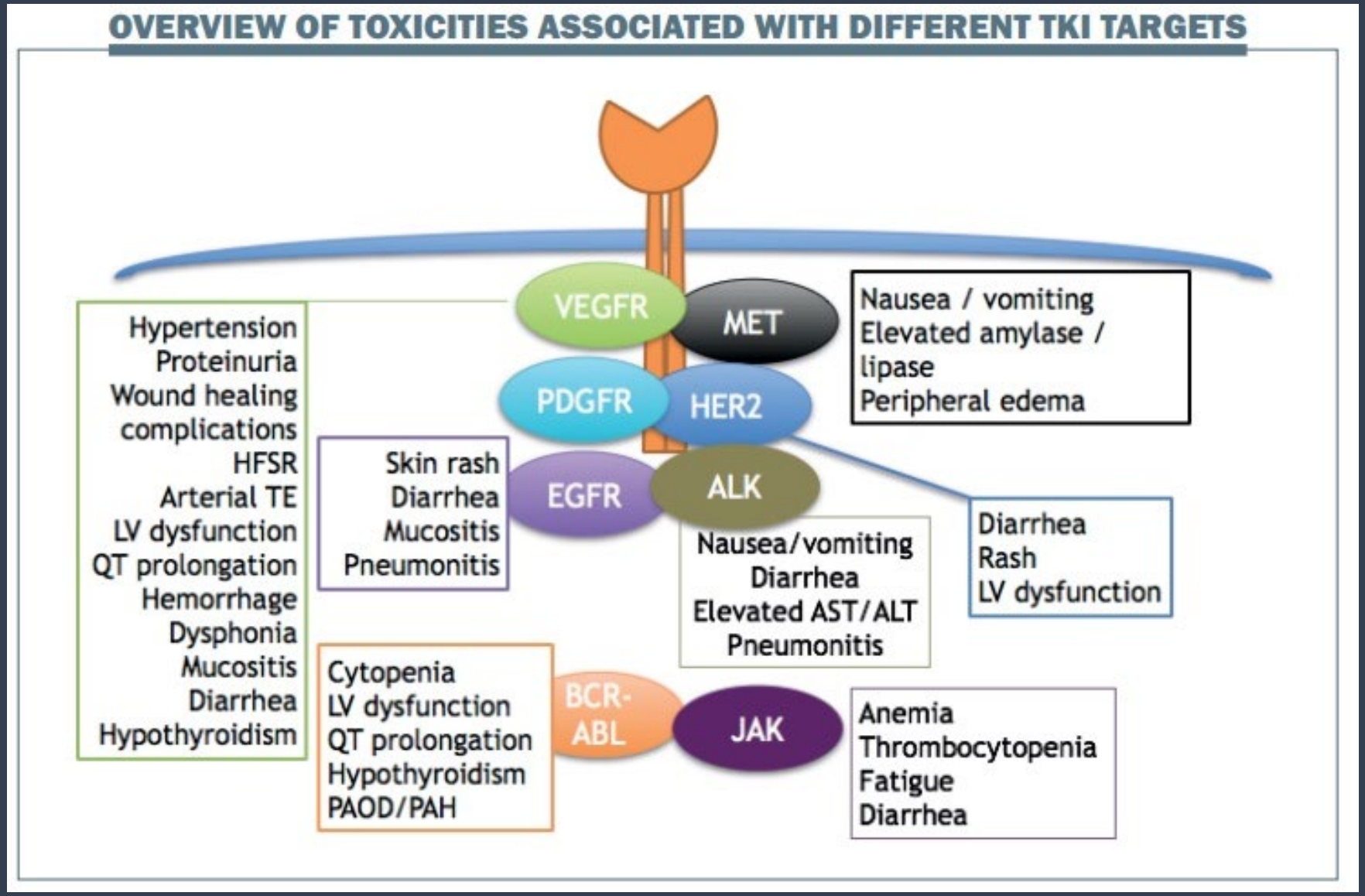
Better Reaction to Bad News

Treating subclinical disease more effectively

20,000 prescription drug products FDA approved

Toxicities

Drugs are doing many different things



Cancer World Archive 2015

nature medicine

Article

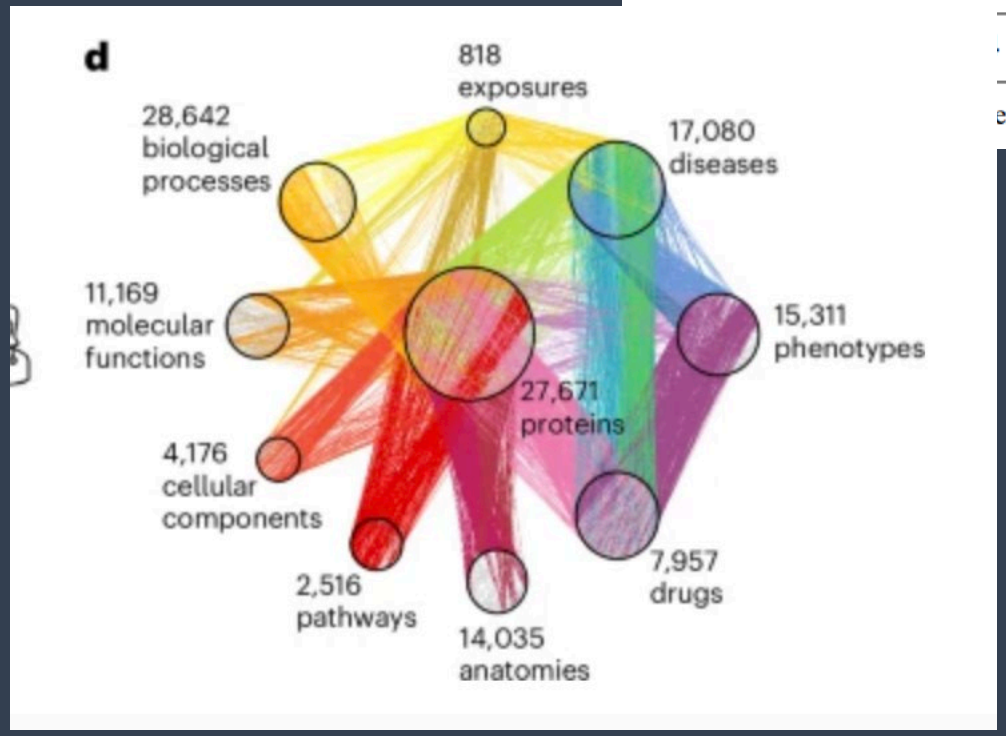
https://doi.org/10.1038/s41591-024-03233-x

A foundation model for clinician-centered drug repurposing

Received: 23 September 2023

er 2024

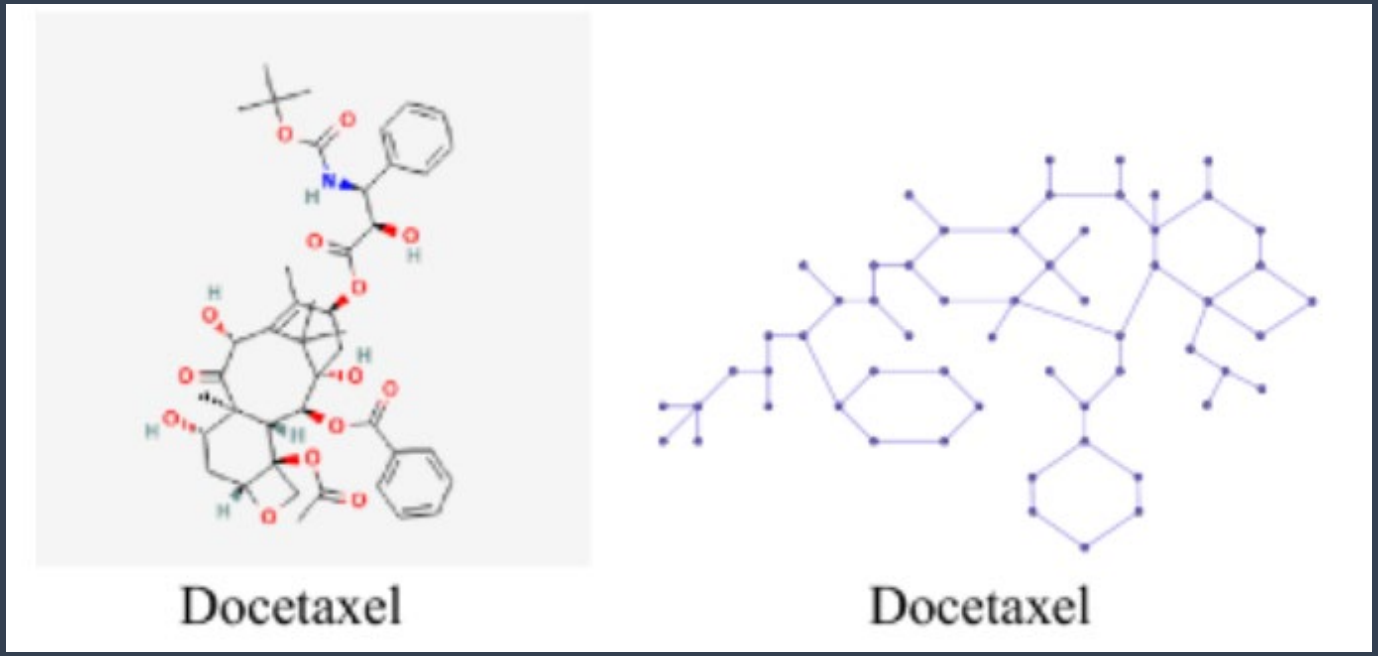
Kexin Huang[1,9,10], Payal Chandak[2,10], Qianwen Wang[1], Shreyas Havaldar[3], Akhil Vaid[3,4], Jure Leskovec[5], Girish N. Nadkarni[4], Benjamin S. Glicksberg[3,4], Nils Gehlenborg[1] & Marinka Zitnik[1,6,7,8]



TxGNN

New Drugs

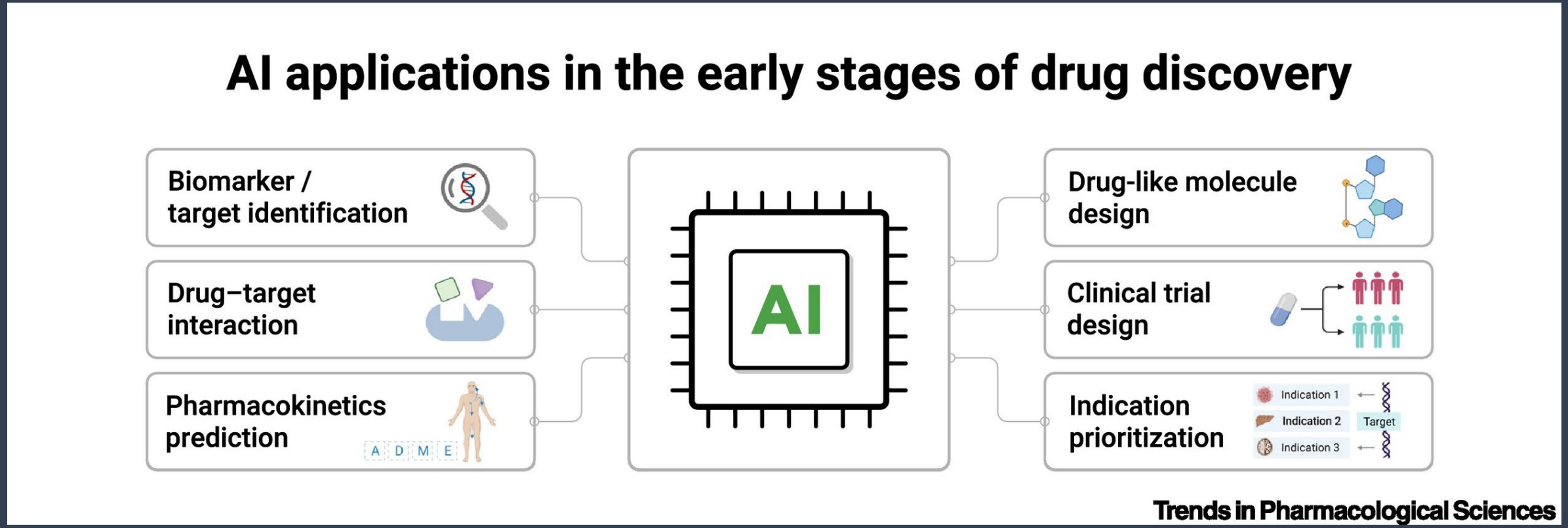
- **QS - PR** Quantitative Structural – Property Relationship
- **QS - AR** Quantitative Structural – Activity Relationship



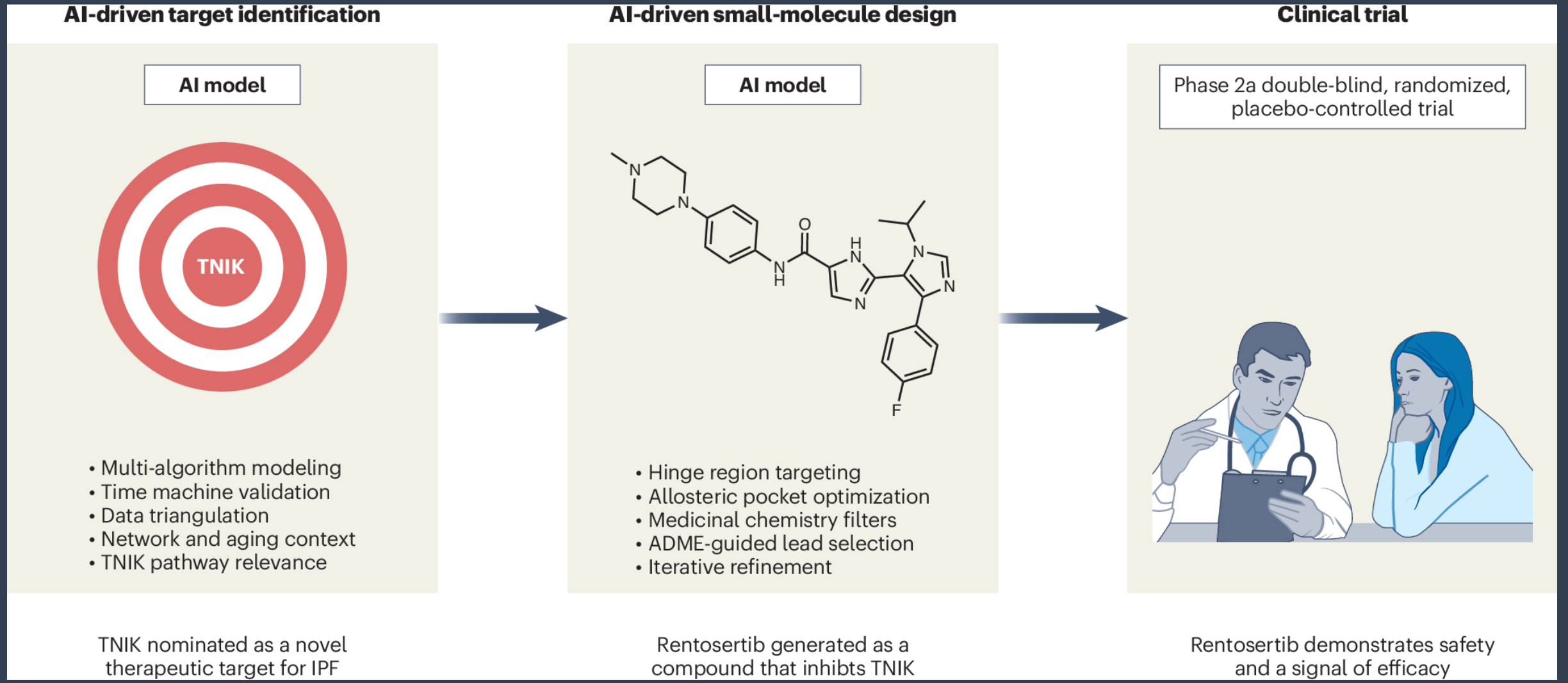
Simplified Molecular Input Line Entry System SMILES

CC1=C2[C@H](C(=O)[C@@]3([C@H](C[C@@H]4[C@]([C@H]3[C@@H]([C@@](C2(C)C)(C[C@@H]1OC(=O)[C@@H]([C@H](C5=CC=CC=C5)NC(=O)OC(C)(C)C)O)O)OC(=O)C6=CC=CC=C6)(CO4)OC(=O)C)O)C)O

Index name	Index symbol	Formula
First entire neighborhood index	$NM_1^{\varepsilon}(\Gamma)$	$\sum_{x \in V(\Gamma) \cup E(\Gamma)} \delta^2(x)$
		$= \sum_{x \in V(\Gamma)} \delta^2(x) + \sum_{x \in E(\Gamma)} \delta^2(x)$
Modified entire neighborhood index	$MNM_1^{\varepsilon}(\Gamma)$	$\sum_{\substack{x \text{ is either adjacent} \\ \text{or incident to } y}} [\delta(x) + \delta(y)]$
		$= \sum_{uv \in E(\Gamma)} [\delta(u) + \delta(v)] + \sum_{ef \in E(L(\Gamma))} [\delta(e) + \delta(f)] + \sum_{\substack{v \text{ incident} \\ \text{to } e \text{ in } \Gamma}} [\delta(v) + \delta(e)]$
Second entire neighborhood index	$NM_2^{\varepsilon}(\Gamma)$	$\sum_{\substack{x \text{ is either adjacent} \\ \text{or incident to } y}} \delta(x)\delta(y)$
		$= \sum_{uv \in E(\Gamma)} [\delta(u)\delta(v)] + \sum_{ef \in E(L(\Gamma))} [\delta(e)\delta(f)] + \sum_{\substack{v \text{ incident} \\ \text{to } e \text{ in } \Gamma}} [\delta(v)\delta(e)]$
Entire neighborhood forgotten index	$NF^{\varepsilon}(\Gamma)$	$\sum_{x \in V(\Gamma) \cup E(\Gamma)} \delta^3(x)$
		$= \sum_{x \in V(\Gamma)} \delta^3(x) + \sum_{x \in E(\Gamma)} \delta^3(x)$
Modified entire neighborhood forgotten index	$MNF^{\varepsilon}(\Gamma)$	$\sum_{\substack{x \text{ is either adjacent} \\ \text{or incident to } y}} [\delta^2(x) + \delta^2(y)]$
		$= \sum_{uv \in E(\Gamma)} [\delta^2(u) + \delta^2(v)] + \sum_{ef \in E(L(\Gamma))} [\delta^2(e) + \delta^2(f)] + \sum_{v \text{ incident}}$



Pun Trends in Pharmacological Sciences2023



Zitnik Nature Medicine 2025

Zu et al Nature Medicine 2025

ARTIFICIAL INTELLIGENCE

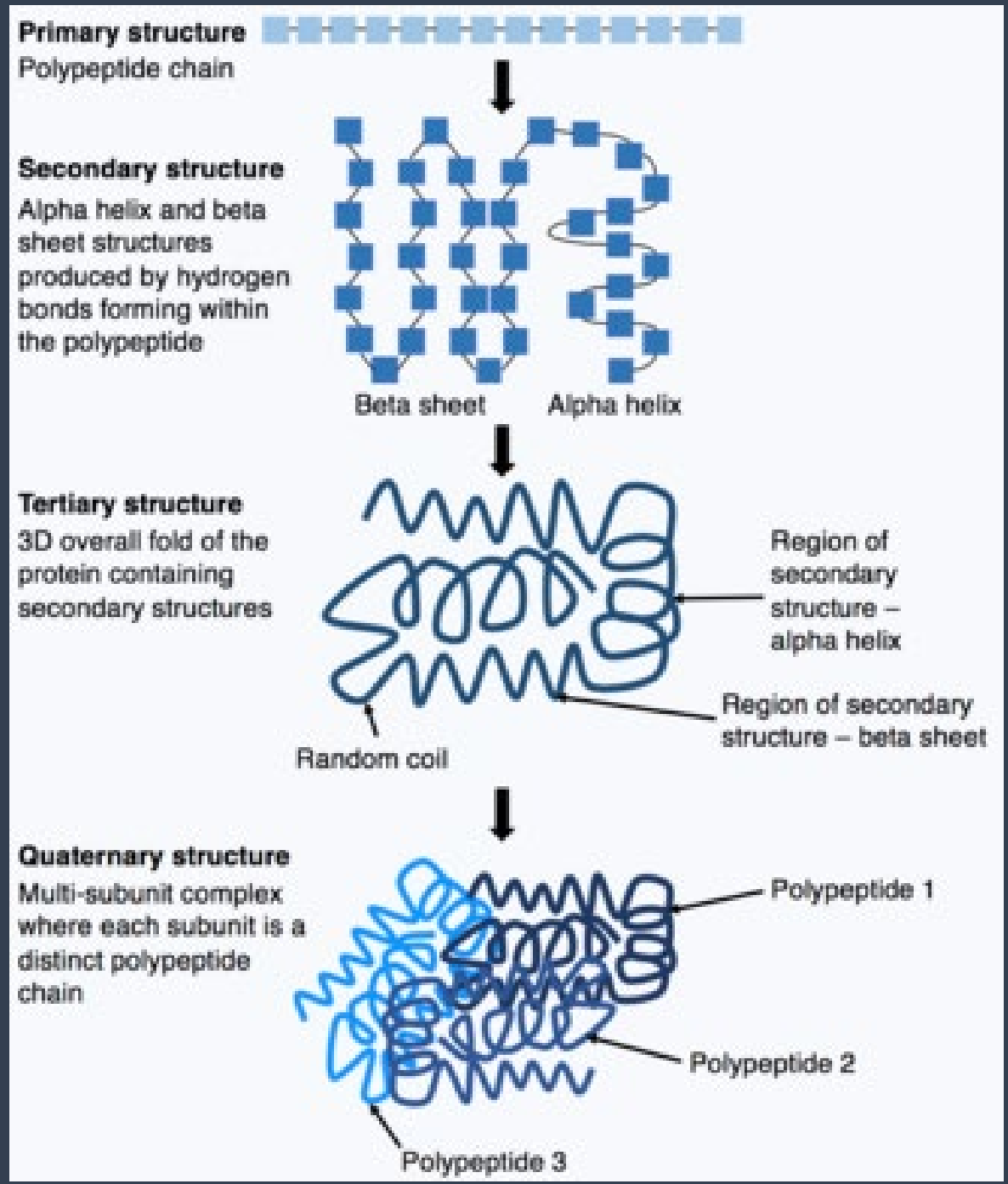
DeepMind's protein-folding AI has solved a 50-year-old grand challenge of biology

AlphaFold can predict the shape of proteins to within the width of an atom. The breakthrough will help scientists design drugs and understand disease.

By Will Douglas Heaven

November 30, 2020

AlphaFold


Primary structure
Polypeptide chain
Secondary structure
Alpha helix and beta sheet structures produced by hydrogen bonds forming within the polypeptide
Beta sheet
Alpha helix
Tertiary structure
3D overall fold of the protein containing secondary structures
Region of secondary structure – alpha helix
Region of secondary structure – beta sheet
Random coil
Quaternary structure
Multi-subunit complex where each subunit is a distinct polypeptide chain
Polypeptide 1
Polypeptide 2
Polypeptide 3

RFdiffusion – design antibodies

Watson Nature 2023

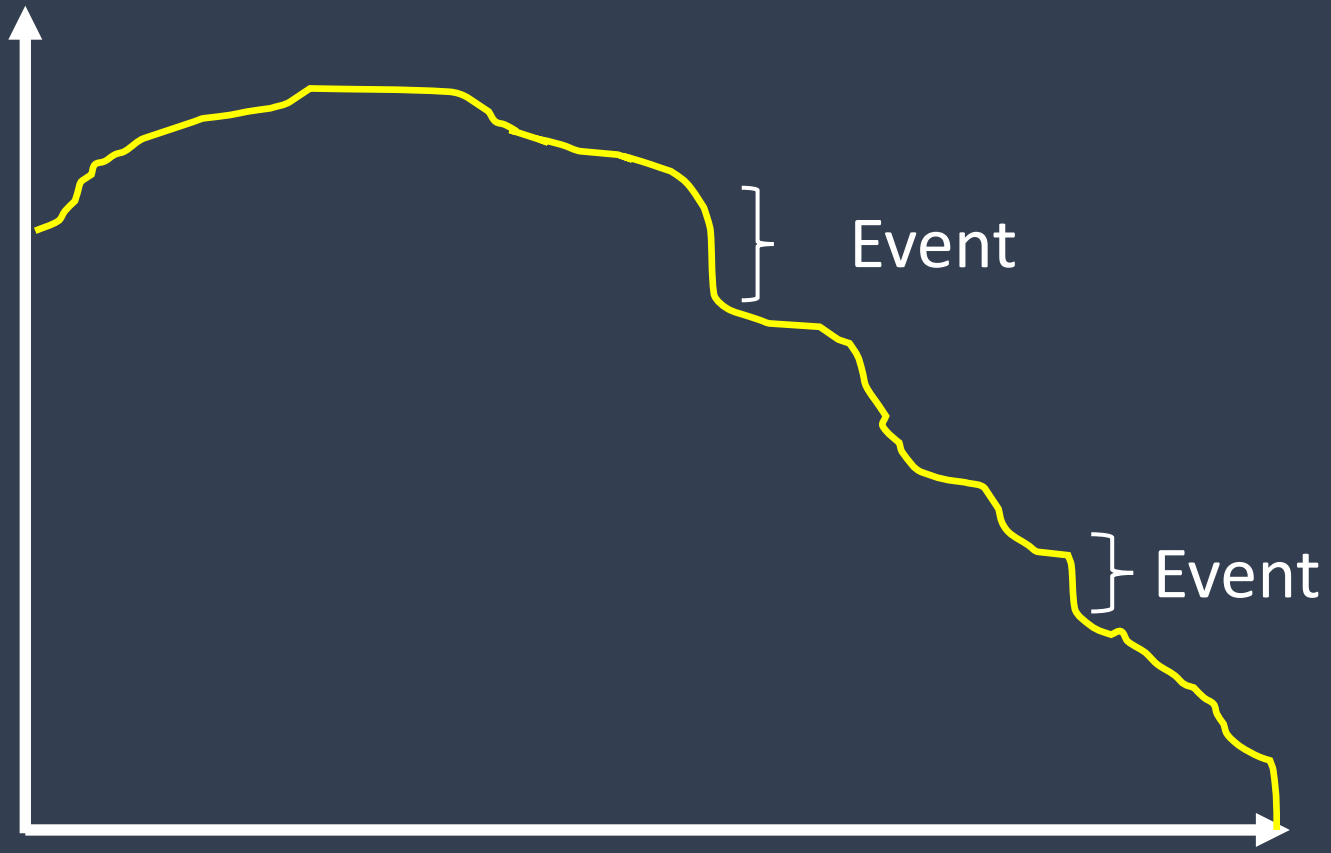
Moravec's Paradox

Tasks which are easy for humans are hard for Artificial Intelligence (AI), and vice versa

% Optimal Health

Years

% Optimal Health

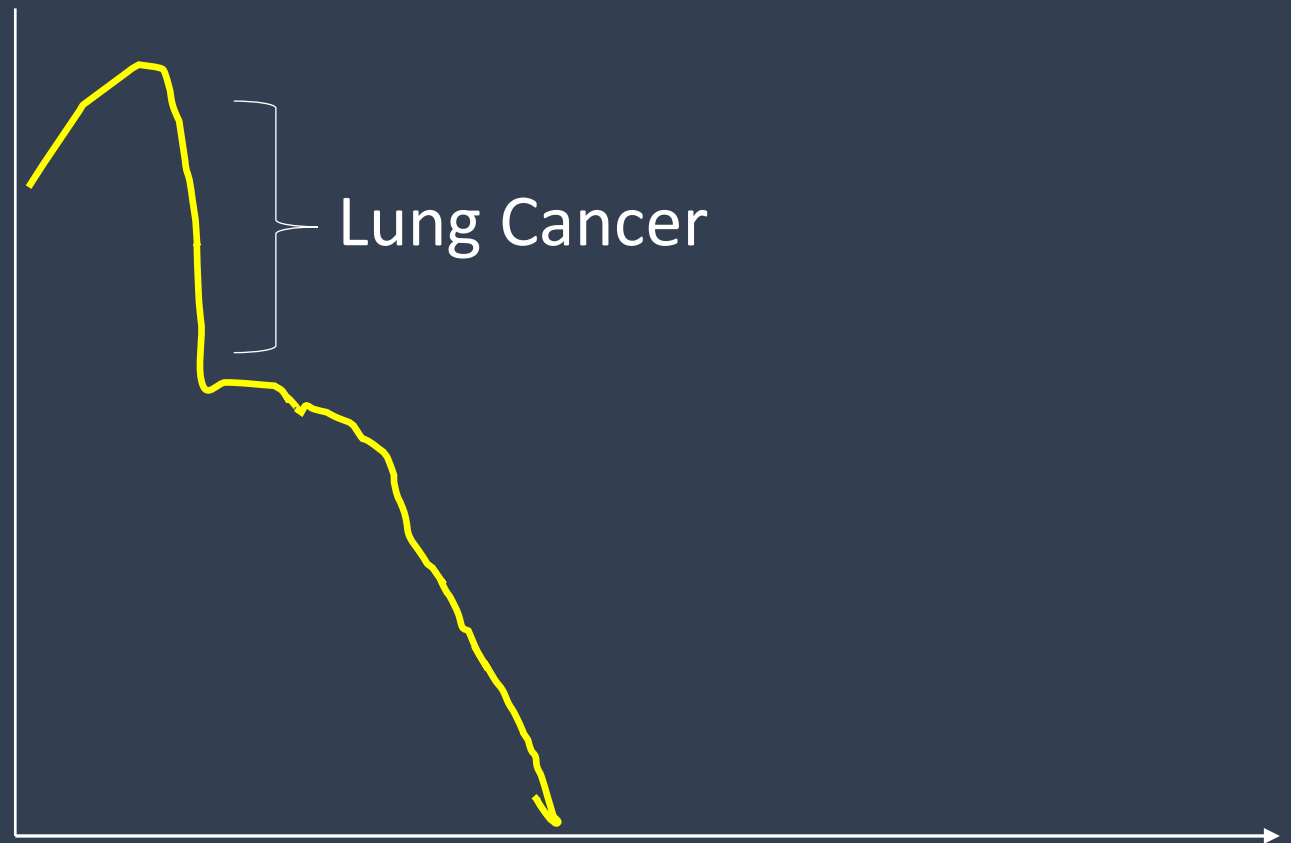


Years

% Optimal Health

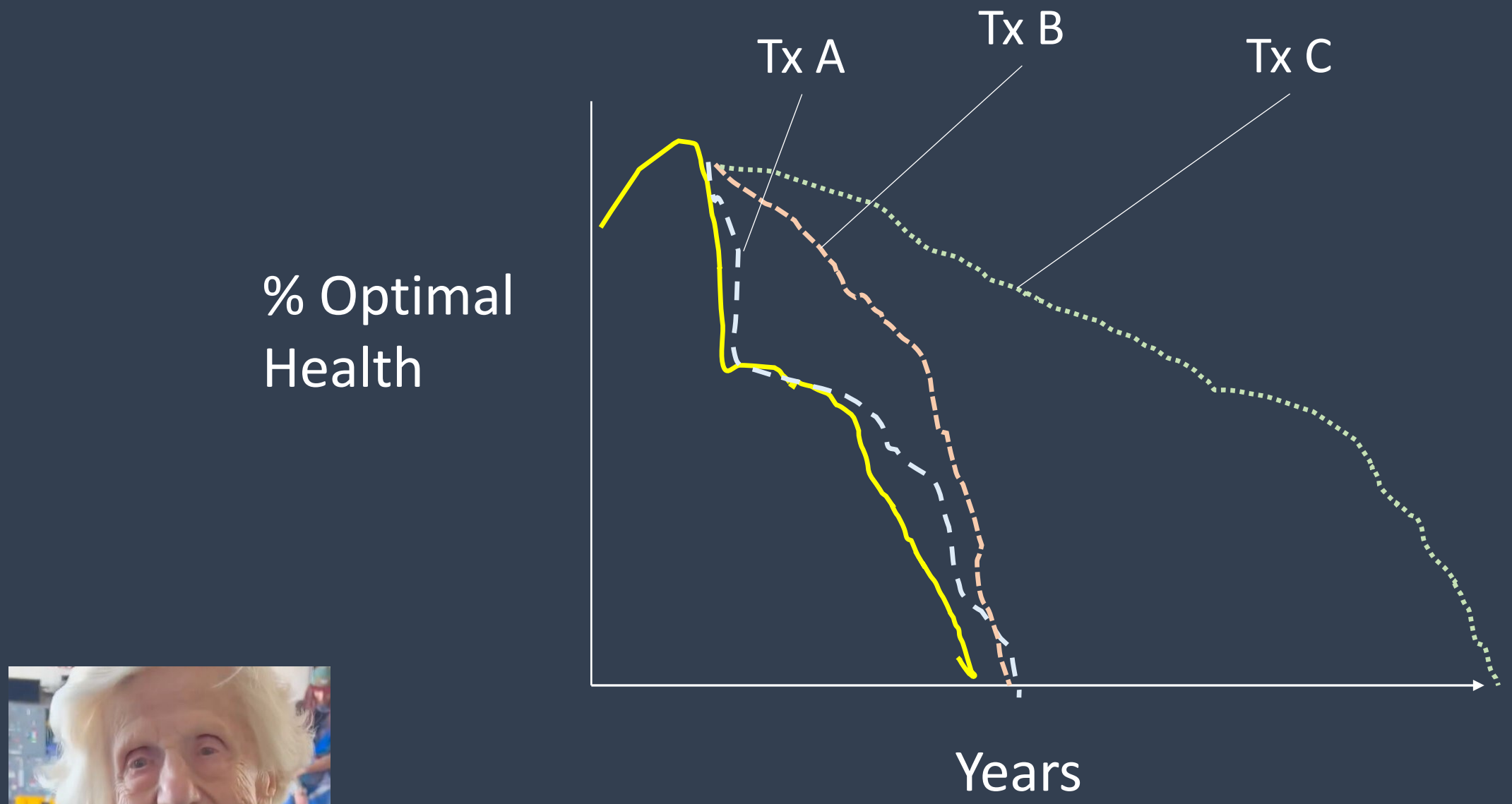
Years

% Optimal Health



Tx A
% Optimal Health
Years

Tx A
Tx B
% Optimal Health
Years


Tx A
Tx B
Tx C
% Optimal Health
Years

Thank you