Technology and Medical Decision Making

Michael W. Kattan, Ph.D.

Cleveland Clinic

kattanm@ccf.org

Professor of Medicine, Epidemiology and Biostatistics,
Cleveland Clinic Lerner College of Medicine of Case Western Reserve University

Chairman, Department of Quantitative Health Sciences,
Cleveland Clinic

Paul Cronan, PhD

M.D. Matthews Endowed Professor

Director of Graduate Programs

Information Systems Department

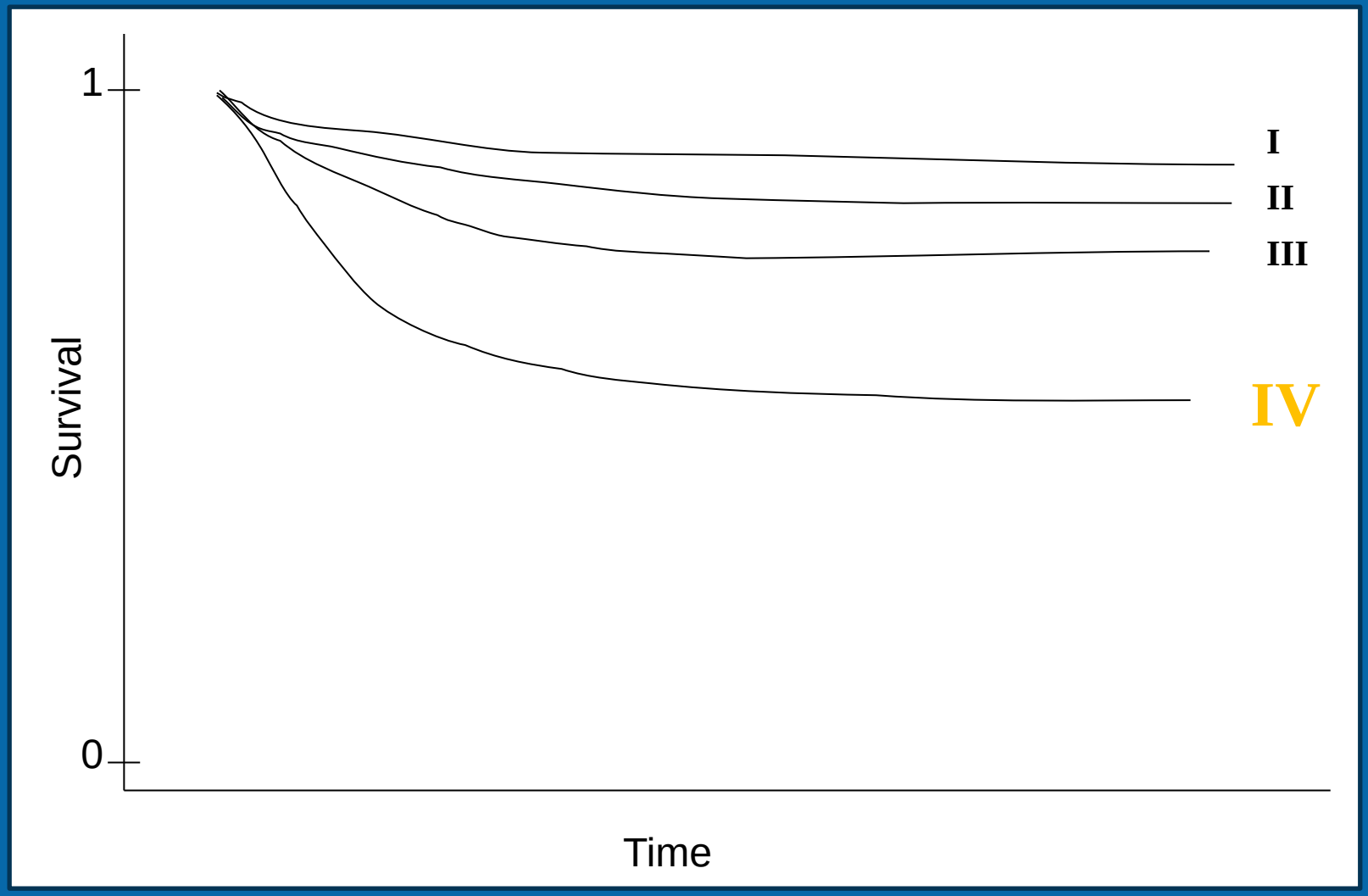
Walton College of Business

University of Arkansas

Cleveland Clinic

December 26, 1989

Hodgkin's Disease Prognosis



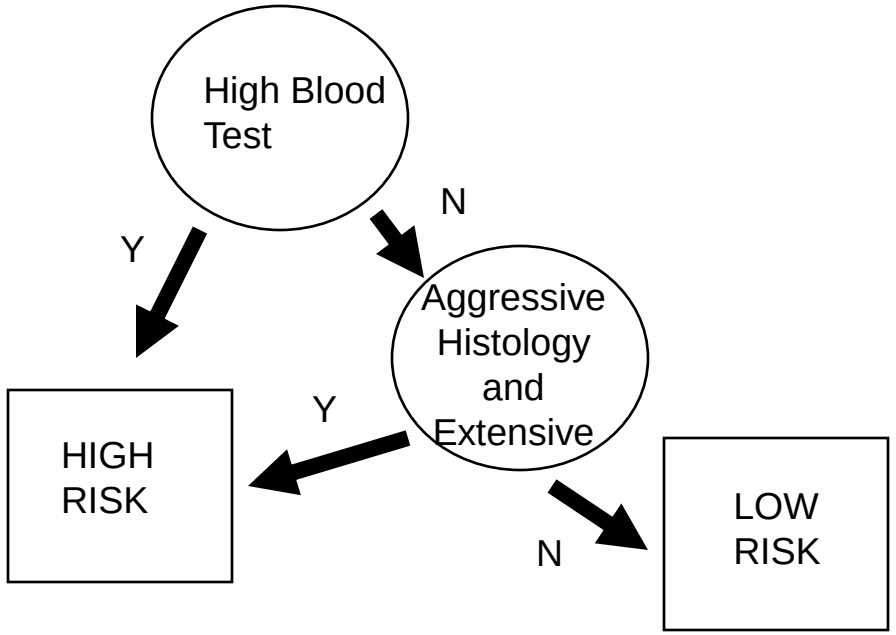
Many clinical counseling tools are not designed to predict accurately

Problems with my prediction:

- Didn't feel very tailored!
  - Not adjusted for age, comorbidities
  - Categories (e.g., extent of disease) were very broad
- Was this staging system really optimized for prediction?

How do we typically compute risk?

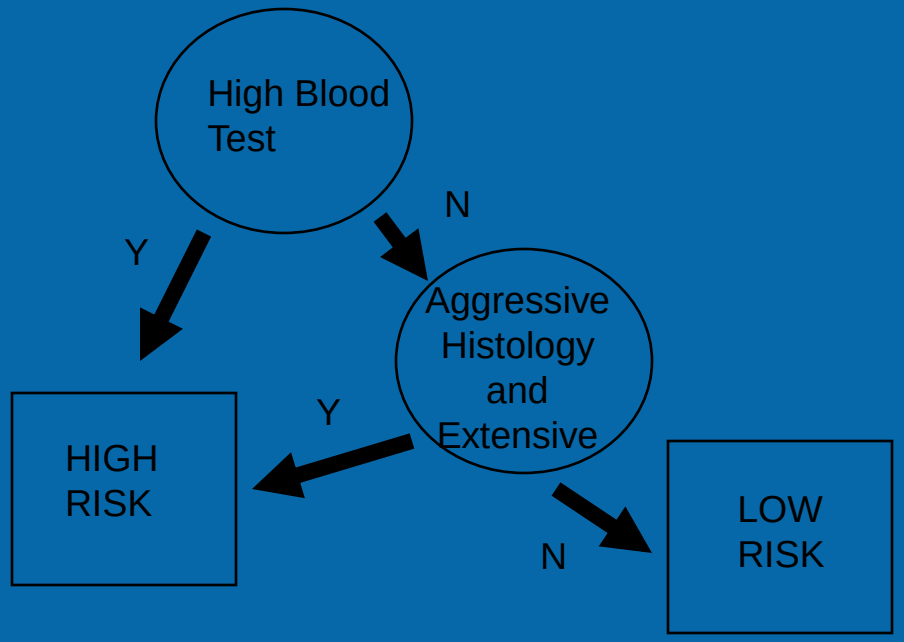
- **Based on features, we make a crude tree.**
- **Most cancer staging systems do this.**



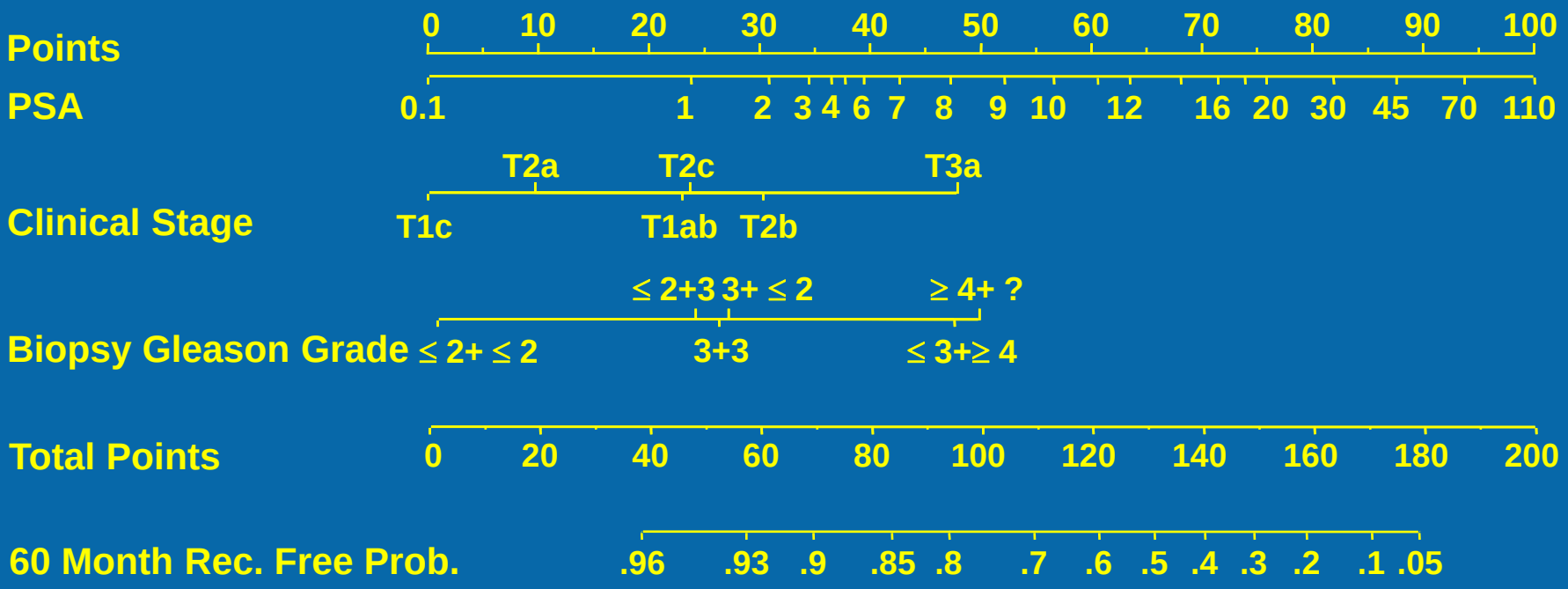
The problem with crude trees

- They are very easy to use.
- But they do not predict outcome optimally.
  - High risk groups are very heterogeneous.
    - A single risk factor may qualify a patient as high risk.

Other approaches, like a **Cox regression statistical model**, predict more accurately.



Preoperative Nomogram for Prostate Cancer Recurrence



Instructions for Physician: Locate the patient's PSA on the **PSA** axis. Draw a line straight upwards to the **Points** axis to determine how many points towards recurrence the patient receives for his PSA. Repeat this process for the **Clinical Stage** and **Biopsy Gleason Sum** axes, each time drawing straight upward to the **Points** axis. Sum the points achieved for each predictor and locate this sum on the **Total Points** axis. Draw a line straight down to find the patient's probability of remaining recurrence free for 60 months assuming he does not die of another cause first.

Instruction to Patient: "Mr. X, if we had 100 men exactly like you, we would expect between <predicted percentage from nomogram - 10%> and <predicted percentage + 10%> to remain free of their disease at 5 years following radical prostatectomy, and recurrence after 5 years is very rare."

Kattan MW et al: JNCI 1998; 90:766-771.

Cleveland Clinic

Some simple steps that will make a difference

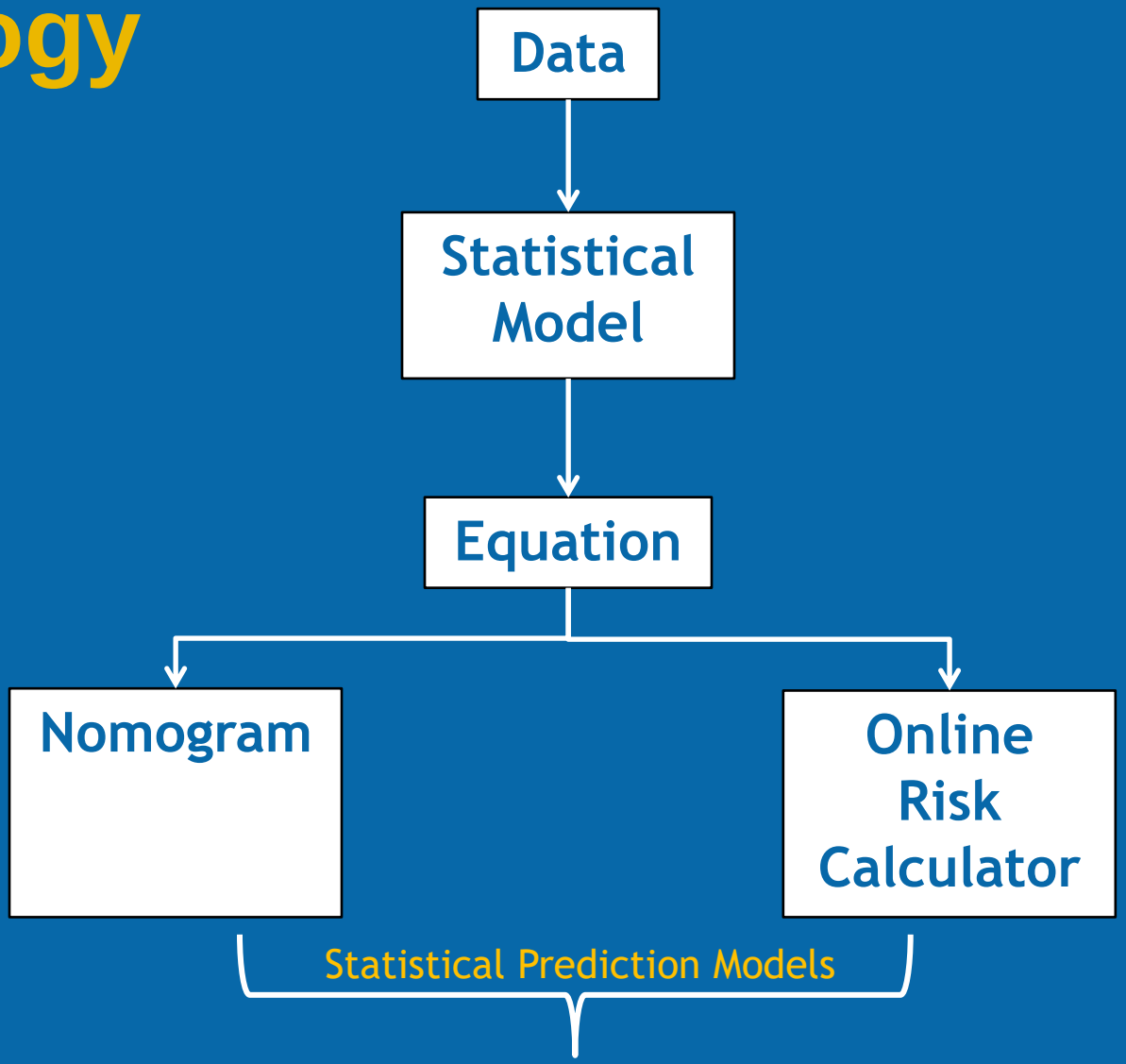
1. Build the most accurate model possible.

2. Take model to bedside
   - As a nomogram,
   - In stand-alone software (desktop, handheld, web)
   - Built into the electronic medical record

Doing this will predict patient outcome more accurately, resulting in

- **better patient counseling**
- **better treatment decision making**

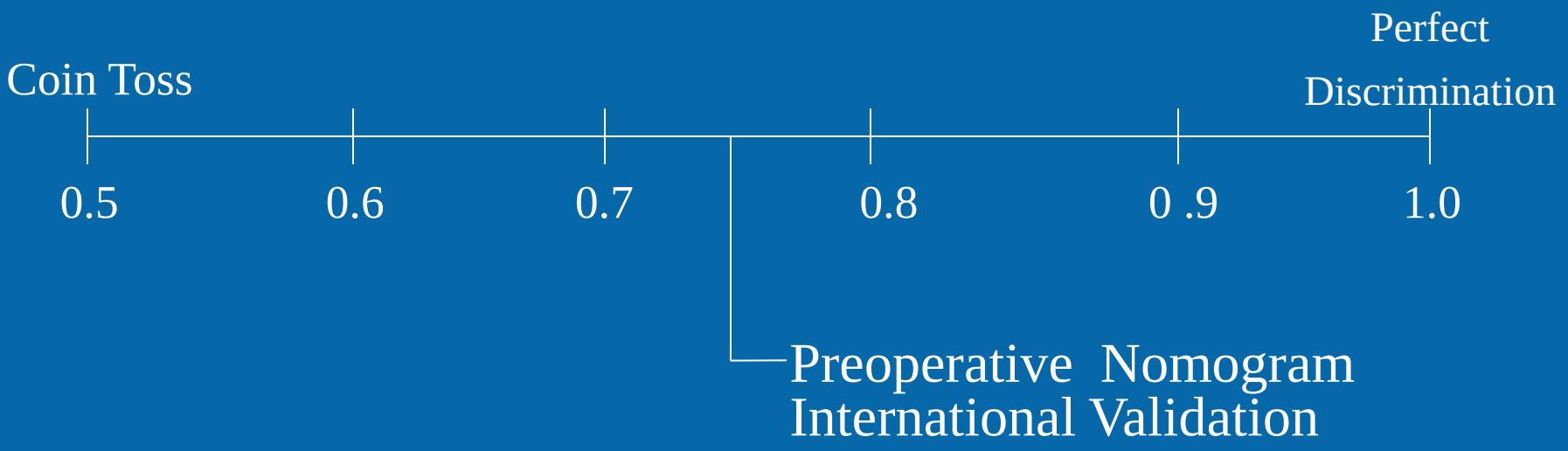
Terminology



Making a nomogram

- Usually a regression model (Cox or logistic)
  - Try machine learning techniques (neural nets, optimized trees like CART)
- Keep continuous variables continuous but relax linearity assumptions
- P-values for predictors don’t matter
- No variable selection or univariable screening
- Bottom line is its predictive accuracy

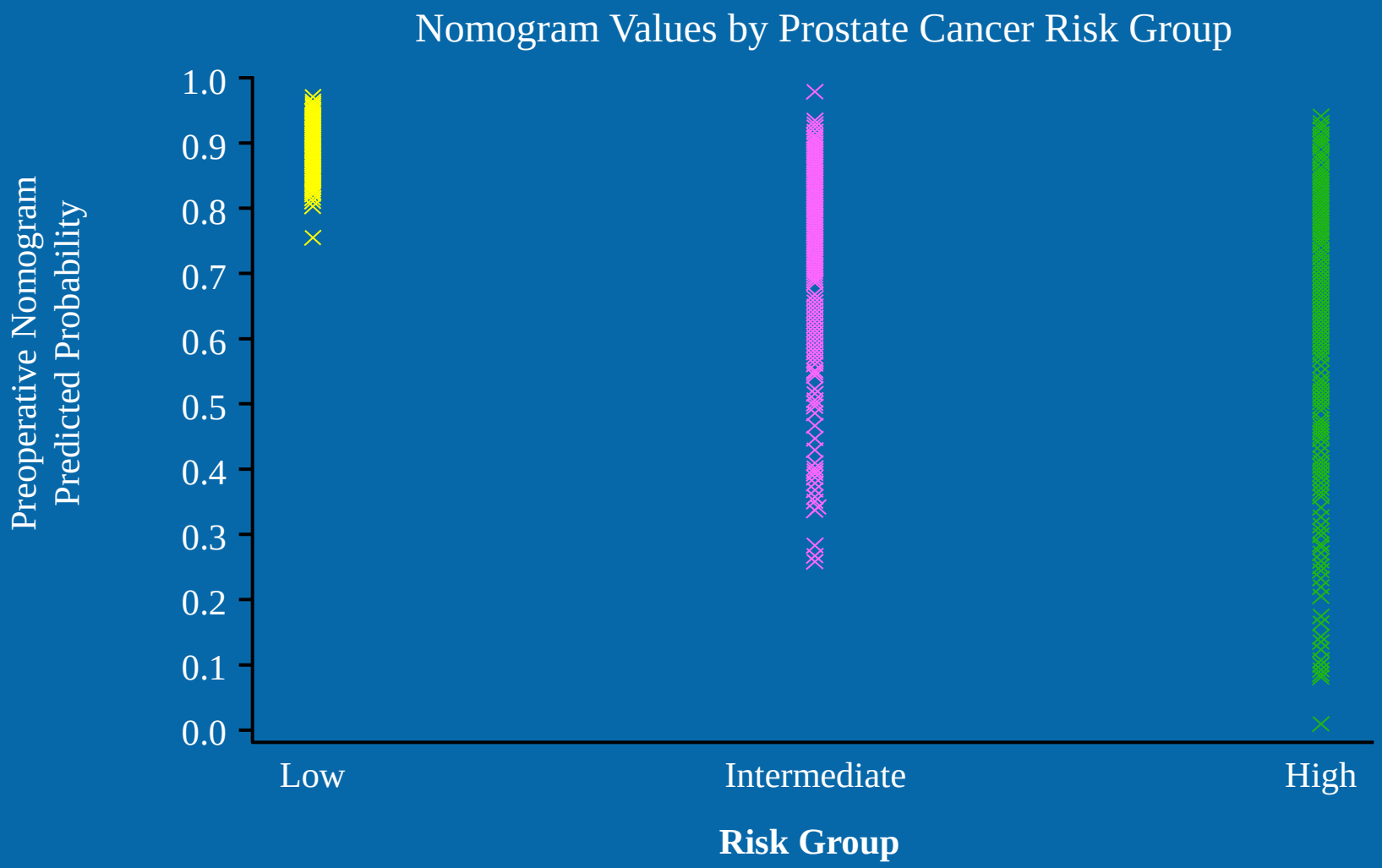
Nomogram Validation by Concordance Index (AUC)



1. Randomly select 2 patients
   a. One of whom fails (reaches the event of interest)
   b. The other must “survive” longer
2. Concordance index is the proportion of these pairs in which patient who fails first also had worse nomogram prediction.

SOURCE: Graefen et al., *JCO*, 2002.

CaPSURE Heterogeneity within Risk Groups



SOURCE: J Urol. 2005 Apr;173(4):1126-31

Cleveland Clinic

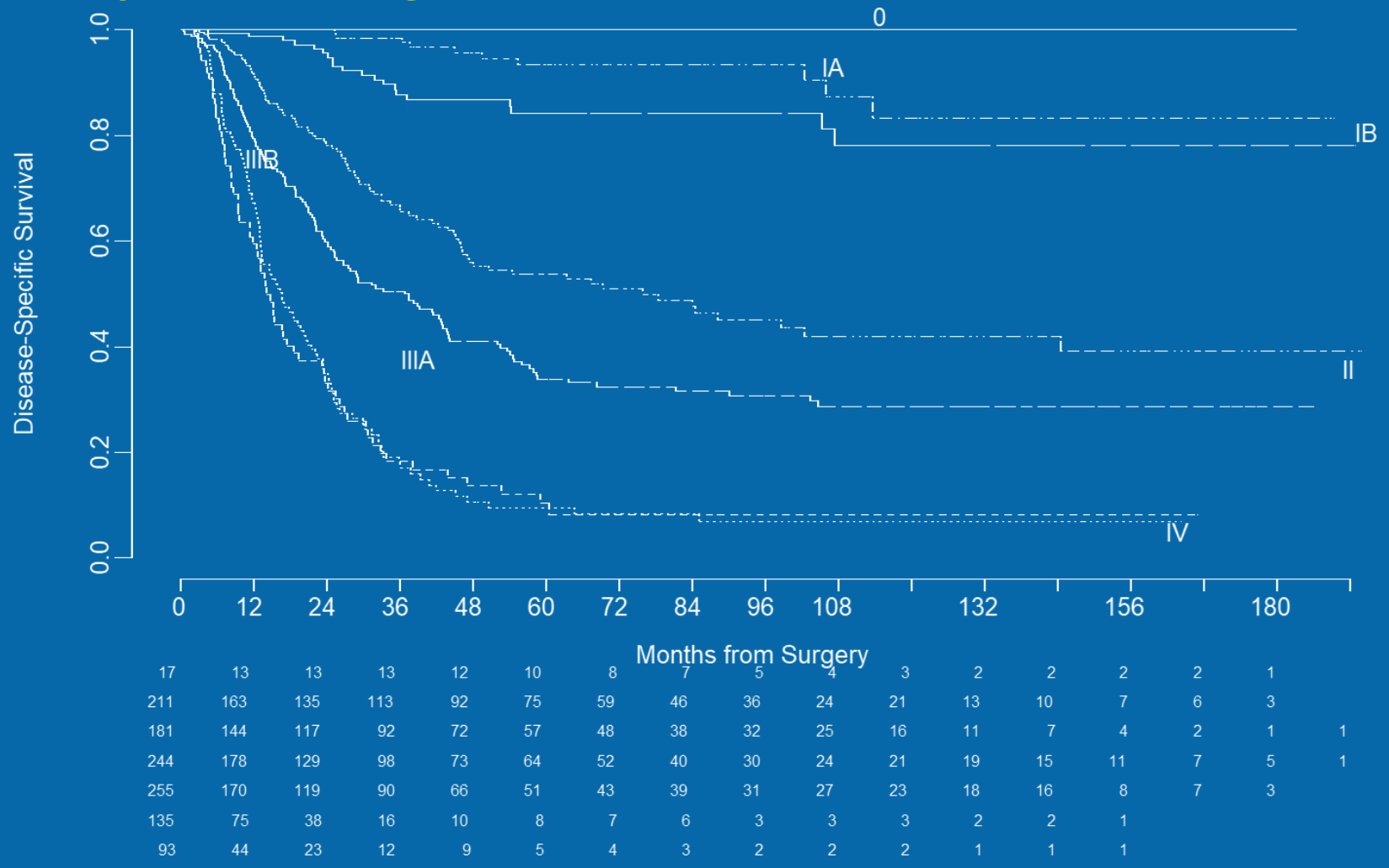
The consequence of risk stratification relative to a statistical model

- Mr. X, from the Cleveland Clinic:
  - PSA=6, clinical stage = T2c, biopsy Gleason sum=9, planned dose of 66.6 Gy without neoadjuvant hormones

- Radiation risk stratification: 81% @ 5 yr.
- Surgery nomogram: 68% @ 5yr.
- Radiation therapy nomogram: 24% @ 5yr.

SOURCE: Kattan MW, et al., *J Clin. Oncol.*, 2000.

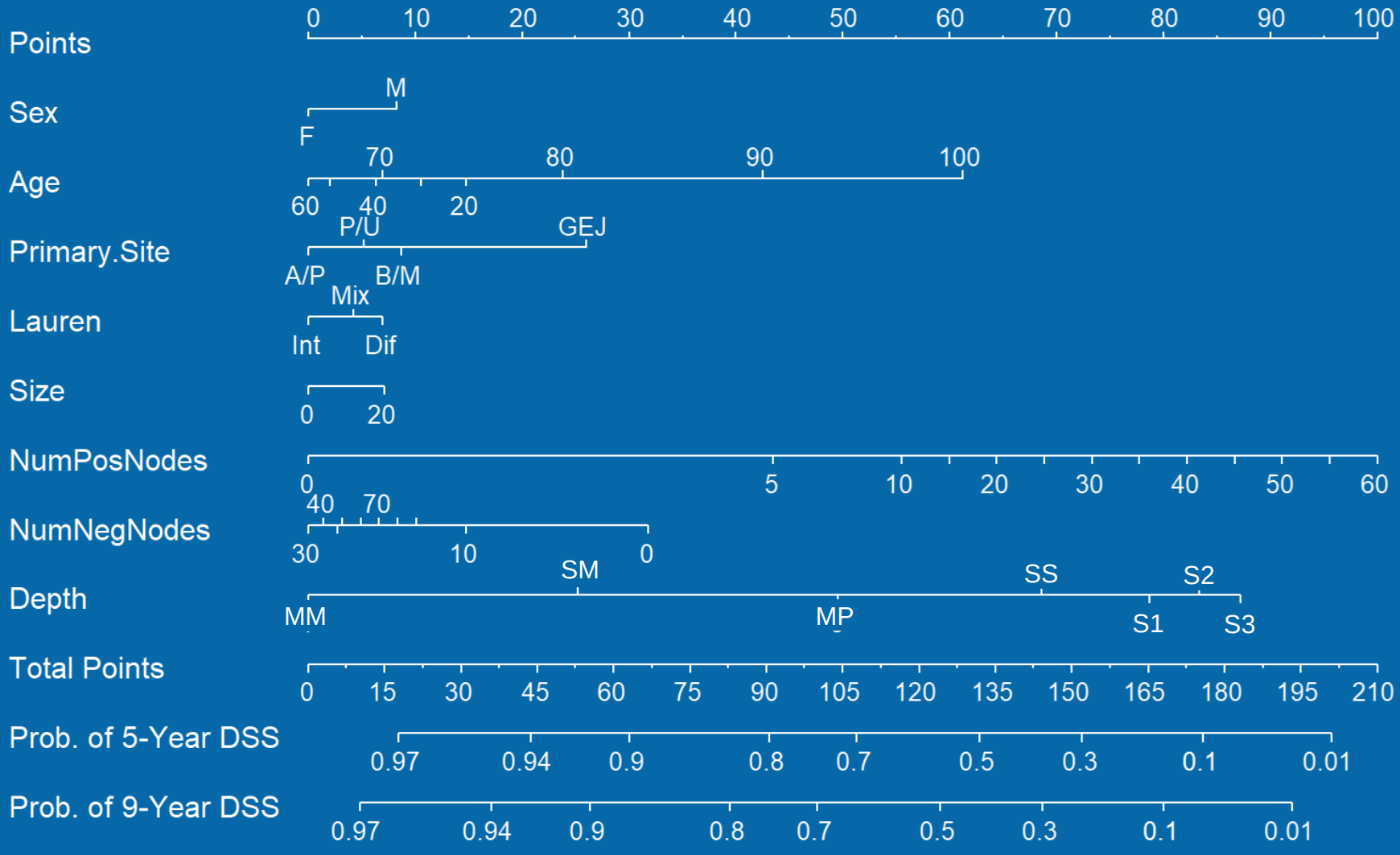
Gastric Cancer Disease-Specific Survival by AJCC Stage



SOURCE: Kattan et al., *JCO*, 2003

Cleveland Clinic

Gastric Cancer Disease-Specific Survival Nomogram



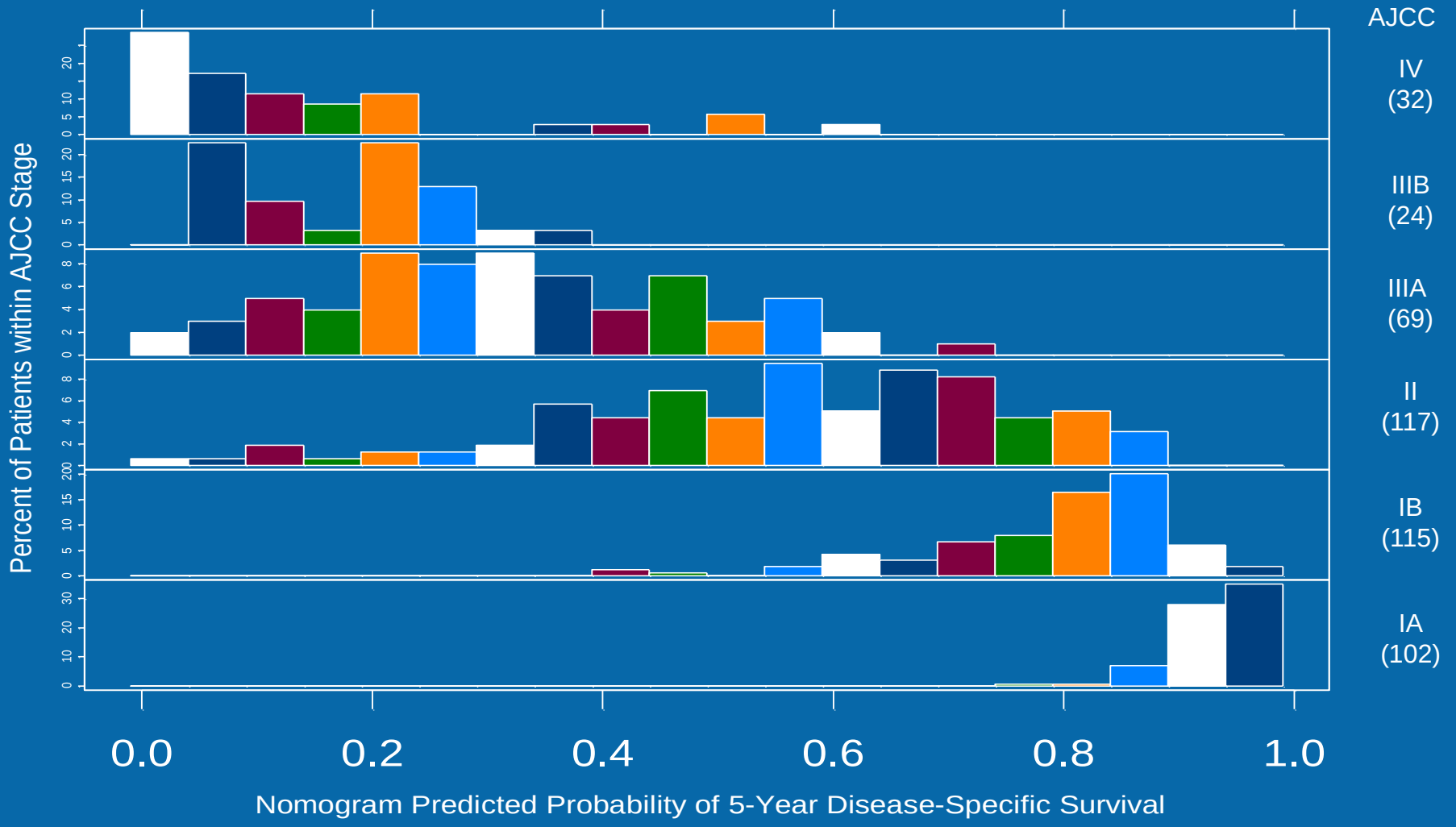
SOURCE: Kattan et al., *JCO*, 2003

Cleveland Clinic

How to tell if we are doing any better than existing models? Validation dataset

	Concordance Index	
Method	Original	Dutch Trial (n=459)
AJCC Stage	0.77	0.75
Nomogram	0.80	0.77
	(p<0.001)	(p<0.001)

Heterogeneity within stages



SOURCE: Kattan et al., *JCO*, 2003

Cleveland Clinic

Continuous Models vs. Staging/Grouping Systems

Model	Comparator	M C
Preop	L/I/H Risk Groups	0.67 vs. 0.64
Preop + IL6/TGFβ1	L/H Risk Groups	0.84 vs. 0.73
Pre XRT	L/I/H Risk Groups	0.76 vs. 0.69
Melanoma SLN+	AJCC Stage	0.69 vs. 0.66
Pancreatic Ca	AJCC Stage	0.64 vs. 0.56
Gastric Ca	AJCC Stage	0.77 vs. 0.75
Breast Ca	NPI Groups	0.69 vs. 0.64
Sarcoma	CART Groups	0.77 vs. 0.74

Why statistical prediction models?

It is an issue of alternatives

- **Predict based on knowledge and experience**
- Deny ability to predict at the individual patient level
- Report the overall average to all patients
- Assign the patient to a risk group, i.e. high, intermediate, or low
- Apply a statistical model

Urologists vs. Preoperative Nomogram

- 10 case descriptions from 1994 MSKCC patients presented to 17 urologists
  - In addition to PSA, biopsy Gleason grades, and clinical stage, urologists were provided with patient age, systematic biopsy details, previous biopsy results, and PSA history.
- Preoperative nomogram was provided.
- Urologists were asked to make their own predictions of 5 year progression-free probabilities with or without use of the preoperative nomogram.
- Concordance indices:
  - Nomogram = 0.67
  - Urologists = 0.55, $p<0.05$

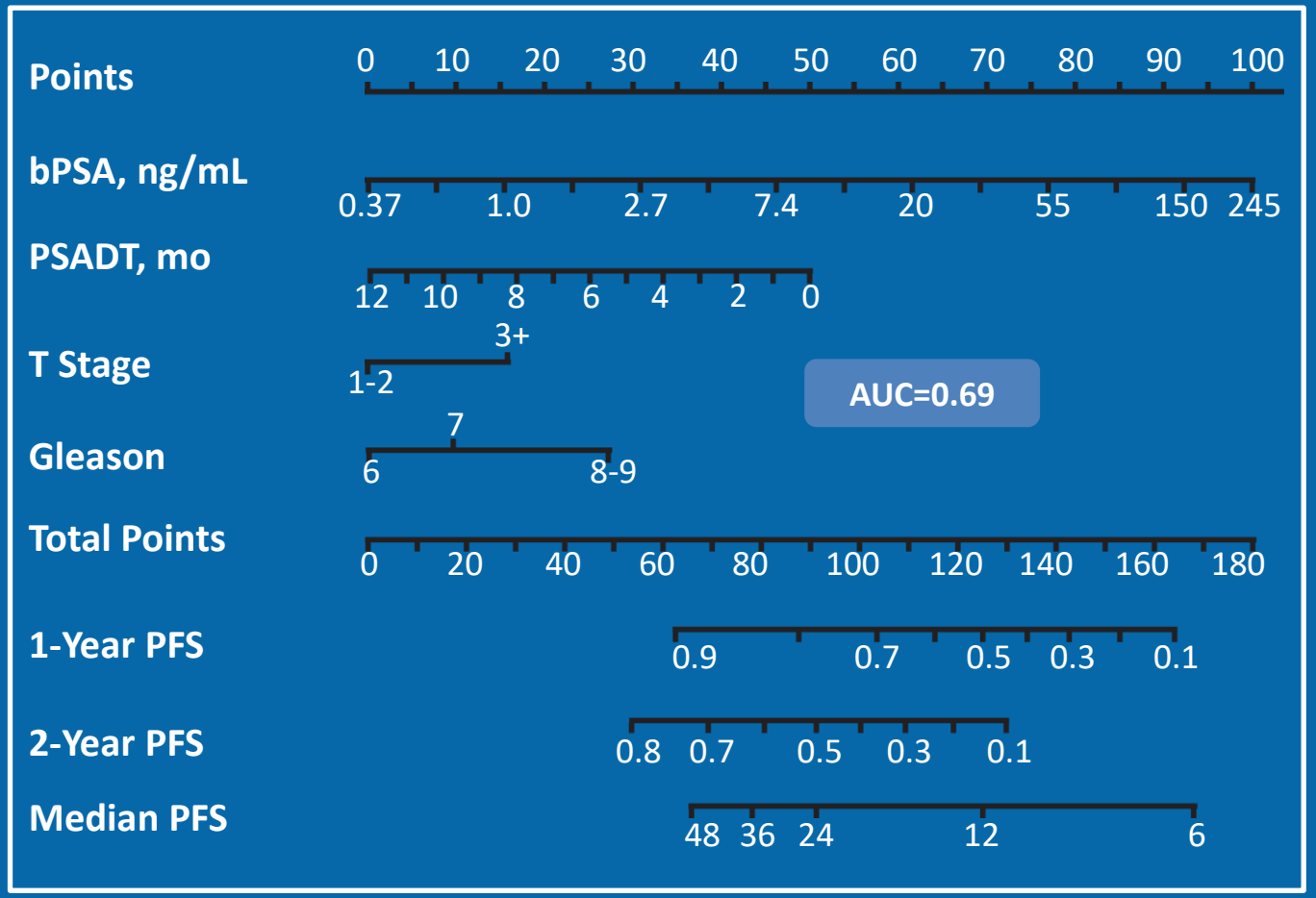
SOURCE: Ross P et al., *Semin Urol Oncol*, 2002.

All of these patients received radical prostatectomy, are now experiencing rising PSA, and have not started ADT.

Age	Race	Clinical Stage	Biopsy PSA	Biopsy Gleason Sum	Adjuvant Radiation	Months from Surgery to Today	Pathological Gleason Sum	Cap.inv	ECE	Margin	SVI	LN	PSA at BCR	PSA Doubling time (months)	If you had 100 patients just like this one, how many do you think would have a positive bone scan 1 year from today if left untreated? (Enter a number between 0 and 100)
67	W	T2A	2.7	7	N	16.12	9.00	P	P	N	P	P	2.5	3.62	
60	W	T2B	12.7	7	N	133.09	7.00	P	N	P	N	N	2959	11.65	
63	W	T1C	20.0	6	N	13.19	7.00	P	N	P	P	P	0.5	5.11	
72	W	T1C	13.2	7	N	9.64	7.00	P	N	P	N	N	0.6	3.04	
64	W	T2C	101.0	5	N	25.10	7.00	P	P	P	P	P	2	3.24	
57	W	T2B	11.1	4	N	9.18	7.00	P	P	P	P	N	6.4	1.51	
54	W	T2B	23.9	10	N	7.60	7.00	P	P	P	P	P	1.5	1.28	
65	W	T2A	13.5	6	N	103.16	7.00	P	P	P	N	N	8	8.52	
65	W	T1C	25.8	6	N	8.13	6.00	P	P	P	N	N	0.5	8.08	
61	W	T1C	13.5	6	N	34.90	7.00	P	P	P	N	N	0.7	11.58	
72	W	T1C	10.1	7	N	14.67	8.00	P	N	P	N	N	0.8	4.29	
67	W	T1C	26.8	6	N	10.43	6.00	P	P	P	N	N	1.4	3.92	
62	W	T2A	4.5	7	N	13.39	7.00	P	P	N	N	N	0.5	5.17	
69	W	T1C	4.7	7	N	11.32	8.00	P	P	N	P	N	3.4	1.76	
67	W	T1C	10.7	6	N	44.05	7.00	P	P	N	N	N	5.4	7.32	
65	W	T1C	7.4	6	N	37.50	7.00	P	N	N	N	N	6.9	5.79	
59	W	T1C	5.0	7	N	13.95	7.00	P	N	N	N	N	0.29	4.07	
57	W	T1C	13.3	7	N	3.82	7.00	P	N	N	N	N	0.5	2.21	
53	W	T1C	14.6	9	N	5.82	9.00	P	P	P	P	N	0.3	3.36	
62	W	T1C	14.6	8	N	20.53	7.00	P	N	P	N	N	1.3	6.10	
62	W	T1C	15.8	9	N	16.12	9.00	P	N	P	N	N	7.4	3.76	
63	W	T2A	7.1	7	N	21.81	7.00	P	P	N	N	N	1.2	5.61	
43	W	T1C	4.6	7	N	31.58	7.00	P	N	P	N	N	1	8.50	
57	W	T2A	4.4	7	N	9.05	7.00	P	N	N	N	N	0.3	4.84	
59	W	T1C	4.2	7	N	20.72	9.00	P	P	P	P	N	0.4	4.50	

Nomogram to Predict Bone Scan Positivity

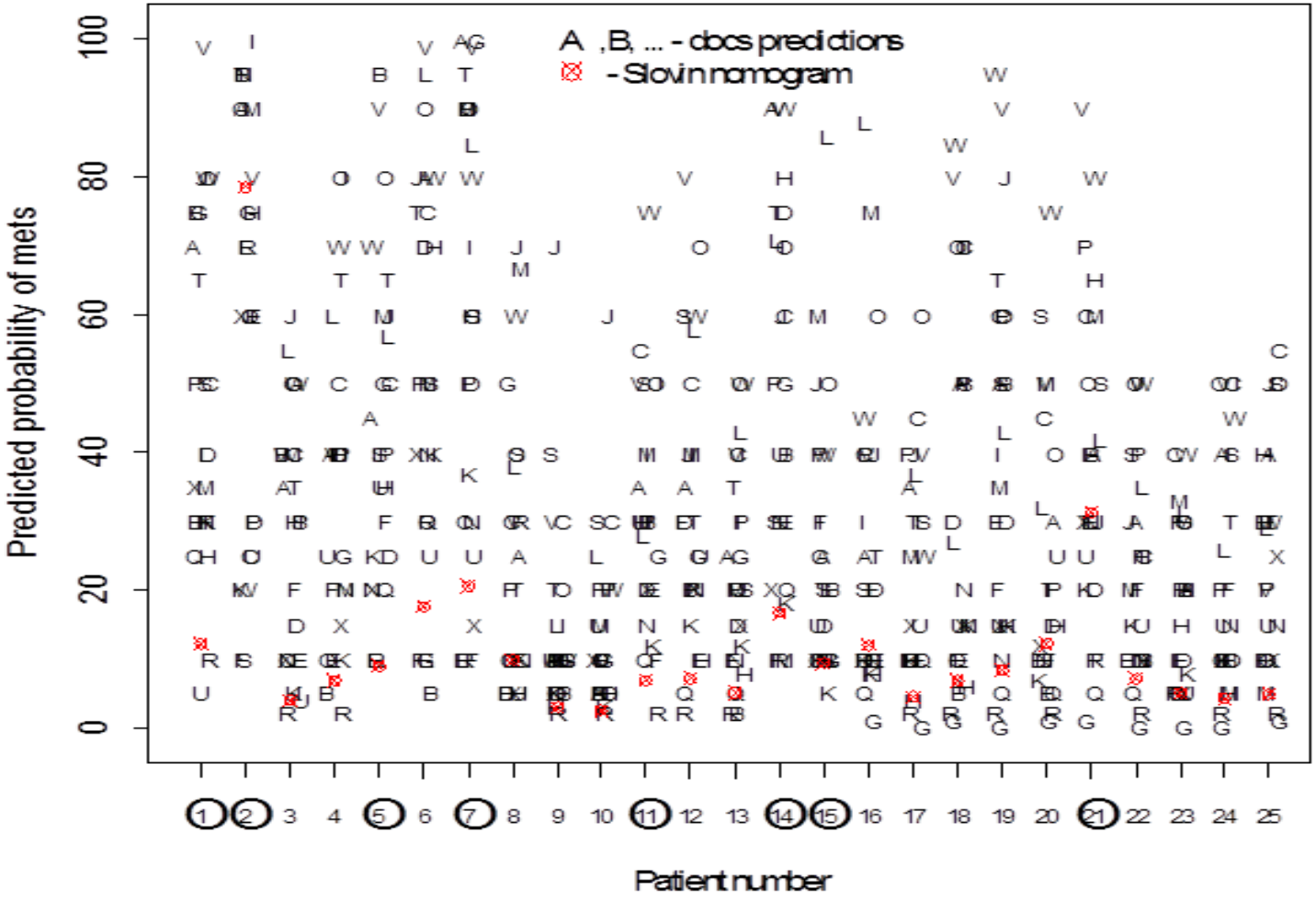
Nomogram Used to Predict Patient-Specific Probabilities of Metastasis-free Survival at 1 and 2 Years, and the Median Progression-free Survival Time



SOURCE: Slovin SF, et al. *Clin Can Res*. 2005;11:8669-8673.

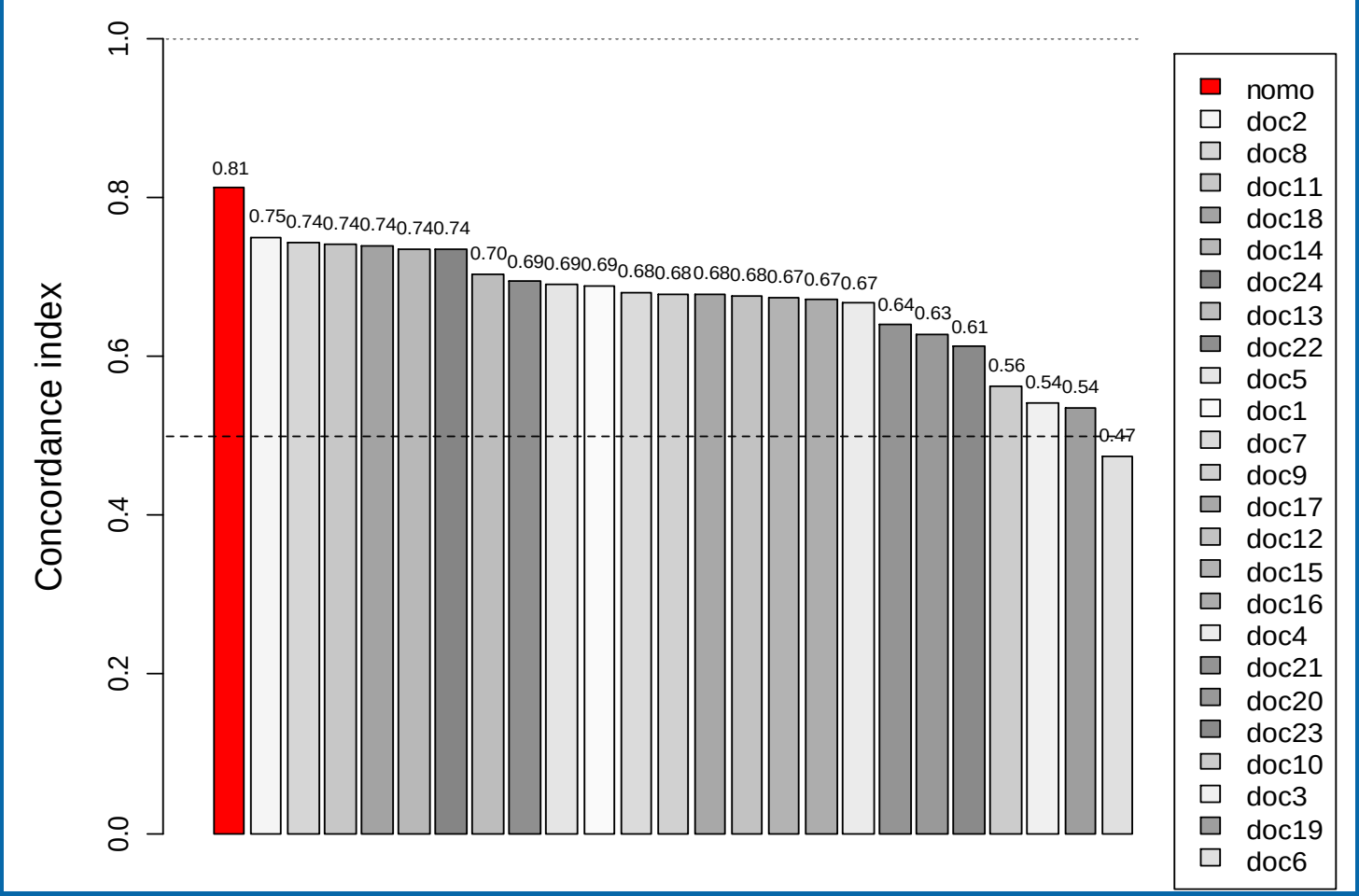
Cleveland Clinic

Clinical Gestalt is Highly Variable



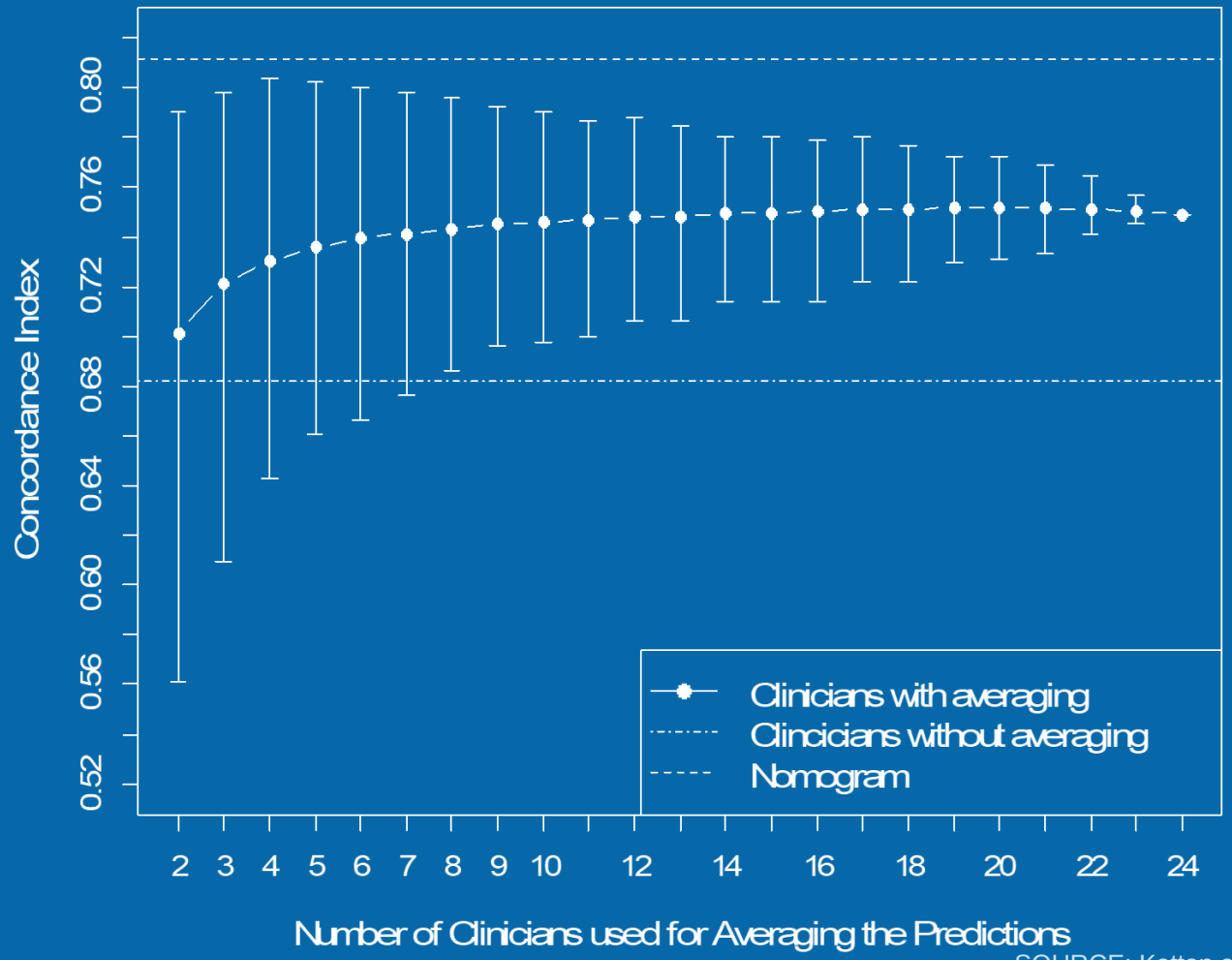
SOURCE: Kattan MW et al., Urology, 2013.

All clinicians lost to the nomogram



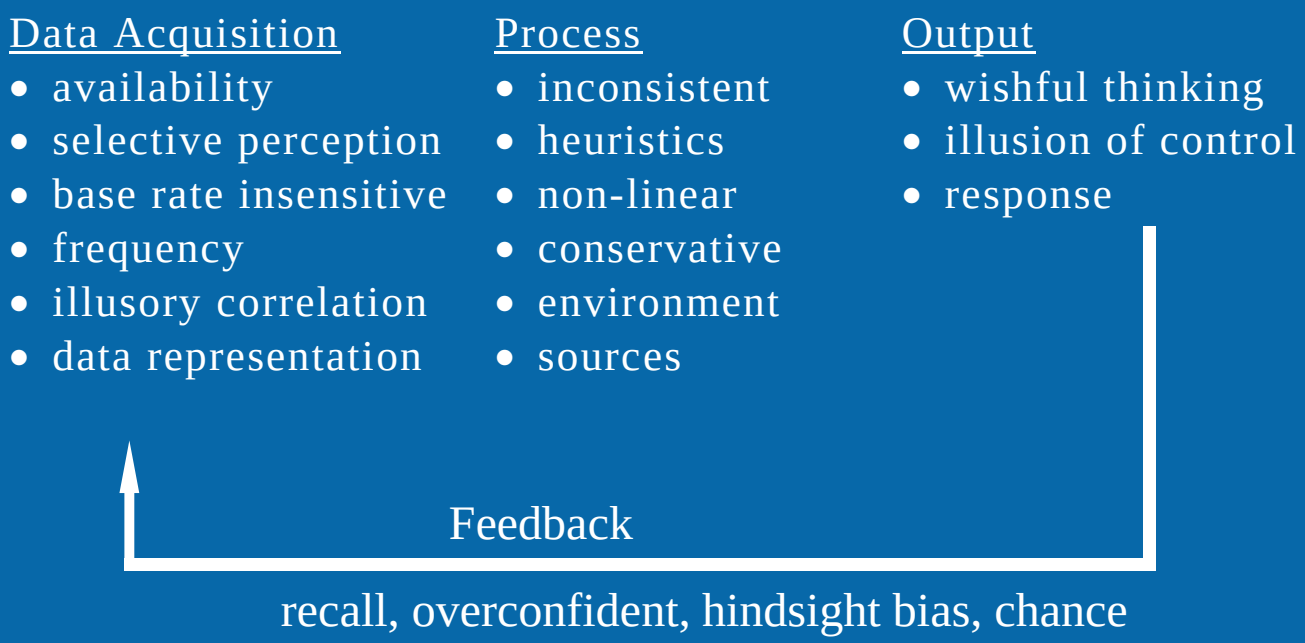
SOURCE: Kattan MW et al., Urology, 2013.

How about averaging the individual clinicians?



SOURCE: Kattan et al., *Med Dec Making*, 2015

Biases in Human Prediction



adapted from Hogarth, 1988

Comparative Effectiveness

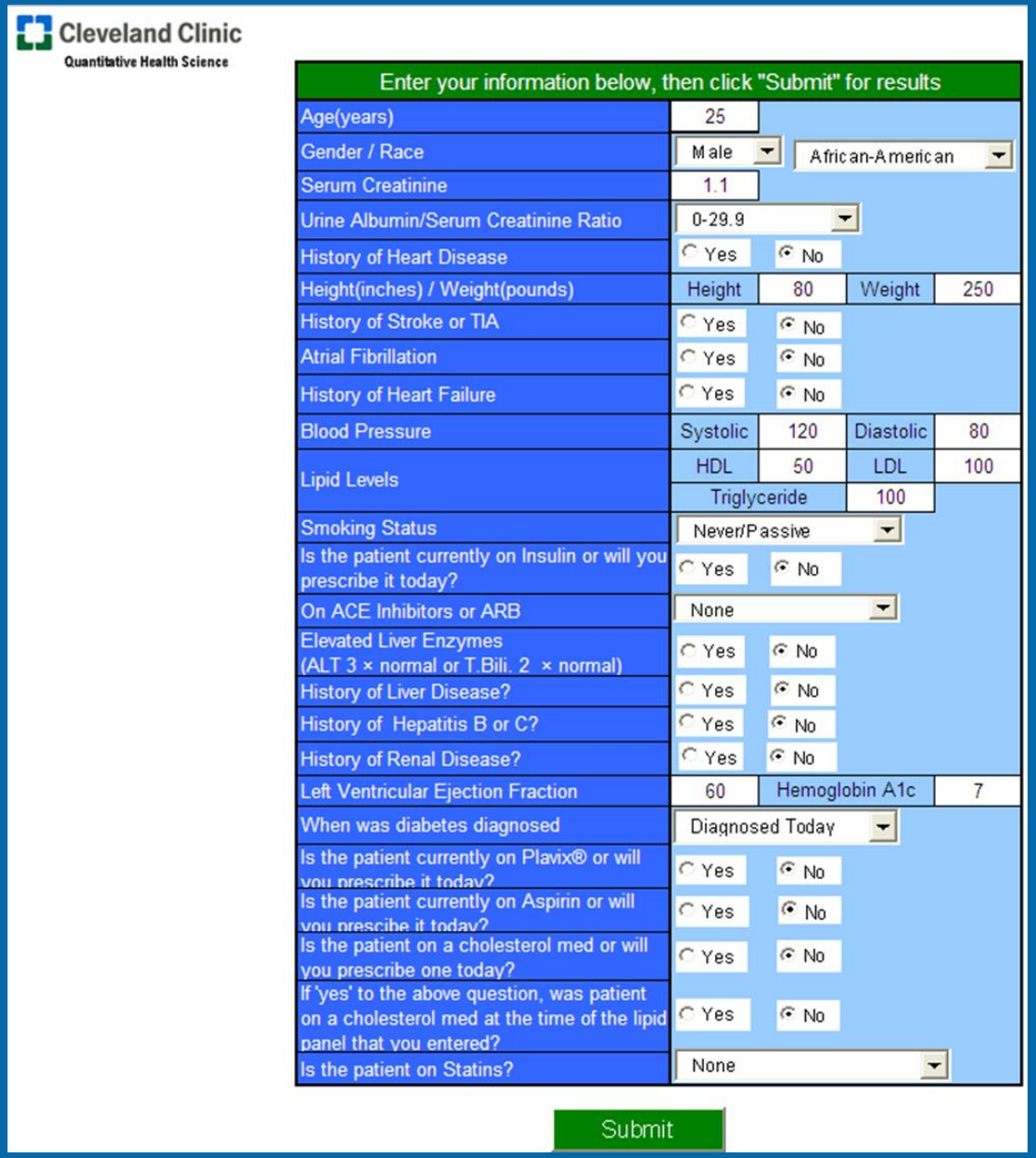
	Treatment options	
Benefits	T1	T2
B1		
B2		
Harms		
H1		
H2		

Probabilities go here

Must tailor the probabilities to the individual patient.

SOURCE: Kattan MW, *Med Decis Making*, 2009.

Risk Calculator



Cleveland Clinic

	DRUG CLASS			
OUTCOMES (6 year probabilities)	Big	Meg	SFU	TZD
Mortality	0.013	0.122	0.054	0.042
Stroke	0.016	0.021	0.018	0.016
Coronary Artery Disease	0.024	0.005	0.028	0.033
Liver Injury	0.073	0.114	0.105	0.089
Heart Failure	0.010	0.015	0.014	0.012
Renal Insufficiency	0.087	0.176	0.130	0.110
Diabetic Nephropathy	0.467	0.386	0.451	0.562

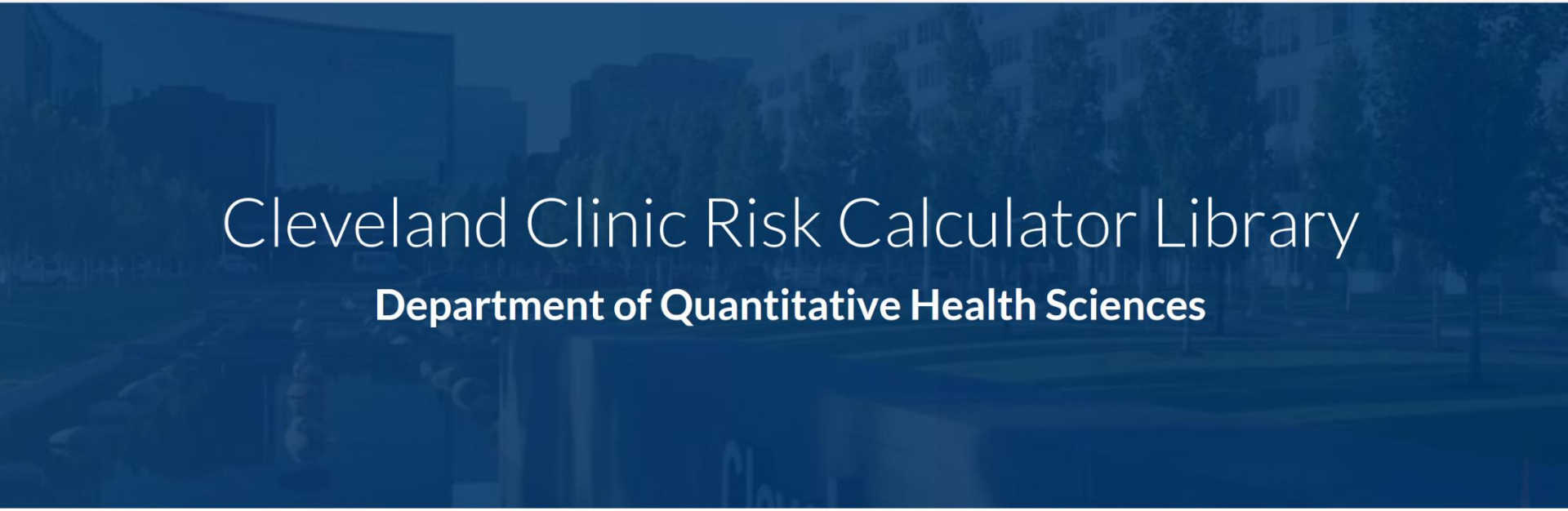
How did we make that risk calculator?

- Mined our own electronic health record at Cleveland Clinic (EPIC).
- Built statistical models relating baseline characteristics to each outcome
  - Tested them for accuracy
- Put all the prediction model equations in a single interface

The reason you need prediction models

- Is not because any model is perfect
- But a prediction model is better than any alternative

http://rcalc.ccf.org



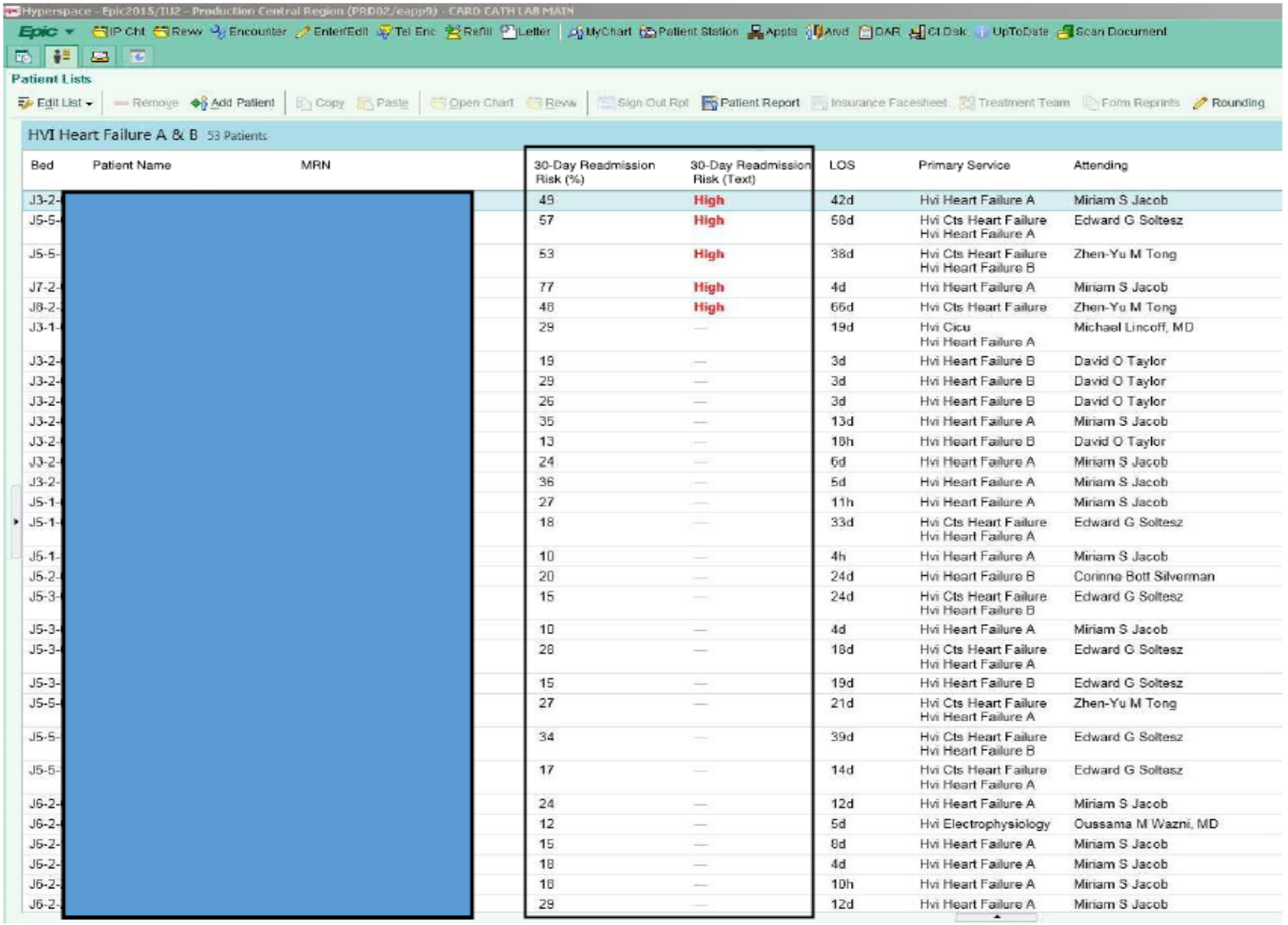
SELECT A CONDITION

BENIGN PROSTATIC HYPERPLASIA

For Patients Receiving Dutasteride Therapy

Cleveland Clinic

Integrating a Risk Calculator into the Electronic Health Record



Hyperspace - Epic2015/TU2 - Production Central Region (PRD02/eapp9) - CARD CATH LAB MAIN

Epic | IP Cht | Rewv | Encounter | Enter/Edit | Tel Enc | Refill | Letter | MyChart | Patient Station | Appts | Arvd | OAR | Ct Disk | UpToDate | Scan Document

Patient Lists

Edit List | Remove | Add Patient | Copy | Paste | Open Chart | Revw | Sign Out Rpt | Patient Report | Insurance Facesheet | Treatment Team | Form Reprints | Rounding

HVI Heart Failure A & B 53 Patients

Bed	Patient Name	MRN	30-Day Readmission Risk (%)	30-Day Readmission Risk (Text)	LOS	Primary Service	Attending
J3-2-	[redacted]	[redacted]	49	High	42d	Hvi Heart Failure A	Miriam S Jacob
J5-5-	[redacted]	[redacted]	57	High	58d	Hvi Cts Heart Failure Hvi Heart Failure A	Edward G Soltesz
J5-5-	[redacted]	[redacted]	53	High	38d	Hvi Cts Heart Failure Hvi Heart Failure B	Zhen-Yu M Tong
J7-2-	[redacted]	[redacted]	77	High	4d	Hvi Heart Failure A	Miriam S Jacob
J8-2-	[redacted]	[redacted]	48	High	66d	Hvi Cts Heart Failure	Zhen-Yu M Tong
J3-1-	[redacted]	[redacted]	29	—	19d	Hvi Cicu Hvi Heart Failure A	Michael Lincoff, MD
J3-2-	[redacted]	[redacted]	19	—	3d	Hvi Heart Failure B	David O Taylor
J3-2-	[redacted]	[redacted]	29	—	3d	Hvi Heart Failure B	David O Taylor
J3-2-	[redacted]	[redacted]	26	—	3d	Hvi Heart Failure B	David O Taylor
J3-2-	[redacted]	[redacted]	35	—	13d	Hvi Heart Failure A	Miriam S Jacob
J3-2-	[redacted]	[redacted]	13	—	18h	Hvi Heart Failure B	David O Taylor
J3-2-	[redacted]	[redacted]	24	—	6d	Hvi Heart Failure A	Miriam S Jacob
J3-2-	[redacted]	[redacted]	36	—	5d	Hvi Heart Failure A	Miriam S Jacob
J5-1-	[redacted]	[redacted]	27	—	11h	Hvi Heart Failure A	Miriam S Jacob
J5-1-	[redacted]	[redacted]	18	—	33d	Hvi Cts Heart Failure Hvi Heart Failure A	Edward G Soltesz
J5-1-	[redacted]	[redacted]	10	—	4h	Hvi Heart Failure A	Miriam S Jacob
J5-2-	[redacted]	[redacted]	20	—	24d	Hvi Heart Failure B	Corinne Bott Silverman
J5-3-	[redacted]	[redacted]	15	—	24d	Hvi Cts Heart Failure Hvi Heart Failure B	Edward G Soltesz
J5-3-	[redacted]	[redacted]	10	—	4d	Hvi Heart Failure A	Miriam S Jacob
J5-3-	[redacted]	[redacted]	28	—	18d	Hvi Cts Heart Failure Hvi Heart Failure A	Edward G Soltesz
J5-3-	[redacted]	[redacted]	15	—	19d	Hvi Heart Failure B	Edward G Soltesz
J5-5-	[redacted]	[redacted]	27	—	21d	Hvi Cts Heart Failure Hvi Heart Failure A	Zhen-Yu M Tong
J5-5-	[redacted]	[redacted]	34	—	39d	Hvi Cts Heart Failure Hvi Heart Failure B	Edward G Soltesz
J5-5-	[redacted]	[redacted]	17	—	14d	Hvi Cts Heart Failure Hvi Heart Failure A	Edward G Soltesz
J6-2-	[redacted]	[redacted]	24	—	12d	Hvi Heart Failure A	Miriam S Jacob
J6-2-	[redacted]	[redacted]	12	—	5d	Hvi Electrophysiology	Oussama M Wazni, MD
J6-2-	[redacted]	[redacted]	15	—	8d	Hvi Heart Failure A	Miriam S Jacob
J6-2-	[redacted]	[redacted]	18	—	4d	Hvi Heart Failure A	Miriam S Jacob
J6-2-	[redacted]	[redacted]	18	—	10h	Hvi Heart Failure A	Miriam S Jacob
J6-2-	[redacted]	[redacted]	29	—	12d	Hvi Heart Failure A	Miriam S Jacob

Conclusions

More accurate predictions can be helpful for a lot of things!

The most accurate predictions presently available should be used, and these are likely from statistical prediction models.

Personalized predictions are key to effective informed consent and are the backbone of medical decision making.

Clinician judgment, risk groups, risk factor counting, and overall treatment effects from RCTs, are all less helpful.