به نام خدا

Clinical Prediction Models

Based on the “Ewout W. Steyerberg’s” book

1st session

Introduction

Davood Khalili, MD, MPH, PhD

Department of Biostatistics and Epidemiology

Research Institute for Endocrine Sciences

Shahid Beheshti University of Medical Sciences

Diagnosis and Prognosis are pivotal in medicine

- **Screening**
  - For high risk people
  - For patients
- **Diagnosis**
- **Therapy**

Prediction Models & Decision-Making

- Traditionally, the probabilities of diagnostic and prognostic outcomes are implicitly assessed.
- Medicine has been much more subjective than in the current era of “evidence-based medicine”.
- Moving towards sharing decision-making with patients by communication about risks and benefits.

Prediction Models & Decision-Making

- Clinical prediction models provide the evidence-based input for shared decision-making, by providing estimates of the individual probabilities of risks and benefits

- Clinical prediction models combine a number of characteristics (e.g., related to the patient, the disease, or treatment) to predict a diagnostic or prognostic outcome.

- Publications with clinical prediction models have increased steeply over recent years

“Statistical Modelling”

To Explain or to Predict?

- **Causal explanation (Testing theories)**
- **Prediction**
- Description

Statistical Science
2010, Vol. 25, No. 3, 289–310
DOI: 10.1214/10-STS330

To Explain or to Predict?

Galit Shmueli

Explaining and Predicting Are Different

Four major aspects are:

Causation–Association: In explanatory modeling f represents an underlying causal function, and X is assumed to cause Y. In predictive modeling f captures the association between $\mathcal{X}$ and $\mathcal{Y}$.

Statistical Science
2010, Vol. 25, No. 3, 289–310
DOI: 10.1214/10-STS330

To Explain or to Predict?

Galit Shmueli

Explaining and Predicting Are Different

Theory–Data: In explanatory modeling, f is carefully constructed based on $\mathcal{F}$ in a fashion that supports interpreting the estimated relationship between X and Y and testing the causal hypotheses. In predictive modeling, f is often constructed from the data. Direct interpretability in terms of the relationship between X and Y is not required, although sometimes transparency of f is desirable.

Statistical Science
2010, Vol. 25, No. 3, 289–310
DOI: 10.1214/10-STS330

To Explain or to Predict?

Galit Shmueli

Explaining and Predicting Are Different

Retrospective–Prospective: Predictive modeling is forward-looking, in that f is constructed for predicting new observations. In contrast, explanatory modeling is retrospective, in that f is used to test an already existing set of hypotheses.

Statistical Science
2010, Vol. 25, No. 3, 289–310
DOI: 10.1214/10-STS330

To Explain or to Predict?

Galit Shmueli

Explaining and Predicting Are Different

Bias–Variance:

In explanatory modeling the focus is on minimizing bias to obtain the most accurate representation of the underlying theory. In contrast, predictive modeling seeks to minimize the combination of bias and estimation variance, occasionally sacrificing theoretical accuracy for improved empirical precision.

“wrong” model can sometimes predict better than the correct one.

Statistical Science
2010, Vol. 25, No. 3, 289–310
DOI: 10.1214/10-STS330

To Explain or to Predict?

Galit Shmueli

The process of statistical modeling through the explain/predict lens, differ from goal definition to model use and reporting.

1 Study Design and Data Collection

2 Data Preparation

3 Exploratory Data Analysis

4 Choice of Variables

5 Choice of Methods

6 Validation, Model Evaluation and Model Selection

7 Model Use and Reporting

Applications of Prediction Models

Public health

- Targeting of preventive interventions

Framingham for CVD & BCDDP, BRCAPRO for breast cancer

Clinical practice

- Diagnostic work-up (e.g. test ordering in renal artery stenosis)
- Starting treatment (threshold for treatment)

Therapeutic decision-making

- Intensity of treatment
- Surgical decision making
- Delaying treatment

Risk factors in four prediction models for breast cancer

Risk factor	Gailmodel	Clausmodel	Myriad tables	BRCAPRO model
Woman's personal information				
Age	+	+	+	+
Race/ethnicity	+			
Ashkenazi Jewish			+	+
Breast biopsy	+			
Atypical hyperplasia	+			
Hormonal factors				
Age at menarche	+			
Age at first live birth	+			
Age at menopause	+			
Family history				
1st degree relatives with breast cancer	+	+	Age <50/≥50	Age for all affected
2nd degree relatives with breast cancer		+	Age <50/≥50	Age for all affected
1st or 2nd degree with ovarian cancer			+	Age for all affected
Bilateral breast cancer				+
Male breast cancer				+
Outcome predicted	Incident breast cancer			BRCA 1/2 mutation

BRCAPRO was shown to perform at least as good as experienced genetic counselors

Applications of Prediction Models

Public health

- Targeting of preventive interventions

Framingham for CVD & BCDDP, BRCAPRO for breast cancer

Clinical practice

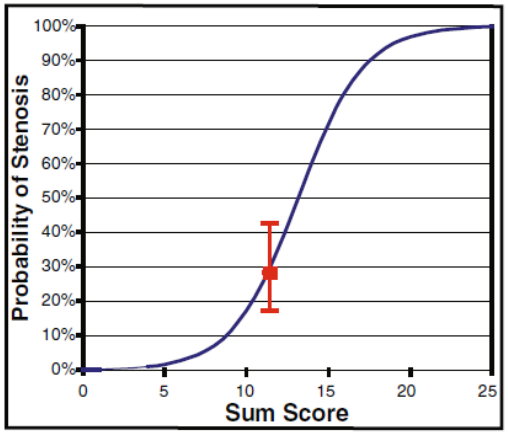
- Diagnostic work-up (e.g. test ordering in renal artery stenosis)
- Starting treatment (threshold for treatment)

Therapeutic decision-making

- Intensity of treatment
- Surgical decision making
- Delaying treatment

Prediction rule for renal artery stenosis as implemented in an Excel spreadsheet

	A	B	C	D	E	F	G	H
1	**Prediction rule for renal artery stenosis**							
2								
3	**Predictor**			Value	Score			
4	Smoking	former or current =1		1	-			
5	Current age	years		45	4.4			
6	Gender	male = 1		1	0			
7	Atherosclerotic vascular disease*	yes = 1		0	0			
8	Onset of hypertension within 2 years	yes = 1		1	1			
9	Body mass index >= 25 kg/m2	yes = 1		0	2			
10	Presence of abdominal bruit	yes = 1		0	0			
11	Serum creatinine concentration	µmol/L		112	4.1			
12	Serum cholesterol level > 6.5 mmol/L**	yes = 1		0	0			
17	*Sumscore*				*11*			
18				Formula	Score chart			
19	*Predicted probability of renal artery stenosis*			28%	25%			
20	*Confidence interval*		17%	-	43%	See figure for graphical illustration		
21	* femoral or carotid bruit, angina pectoris, claudication, myocardial infarction, CVA, or vascular surgery							
22	** or cholesterol lowering therapy							



The diagnostic accuracy of the regression model was similar to that of renal scintigraphy, which had a sensitivity of 72% and a specificity of 90%.

Applications of Prediction Models

Public health

- Targeting of preventive interventions

Framingham for CVD & BCDDP, BRCAPRO for breast cancer

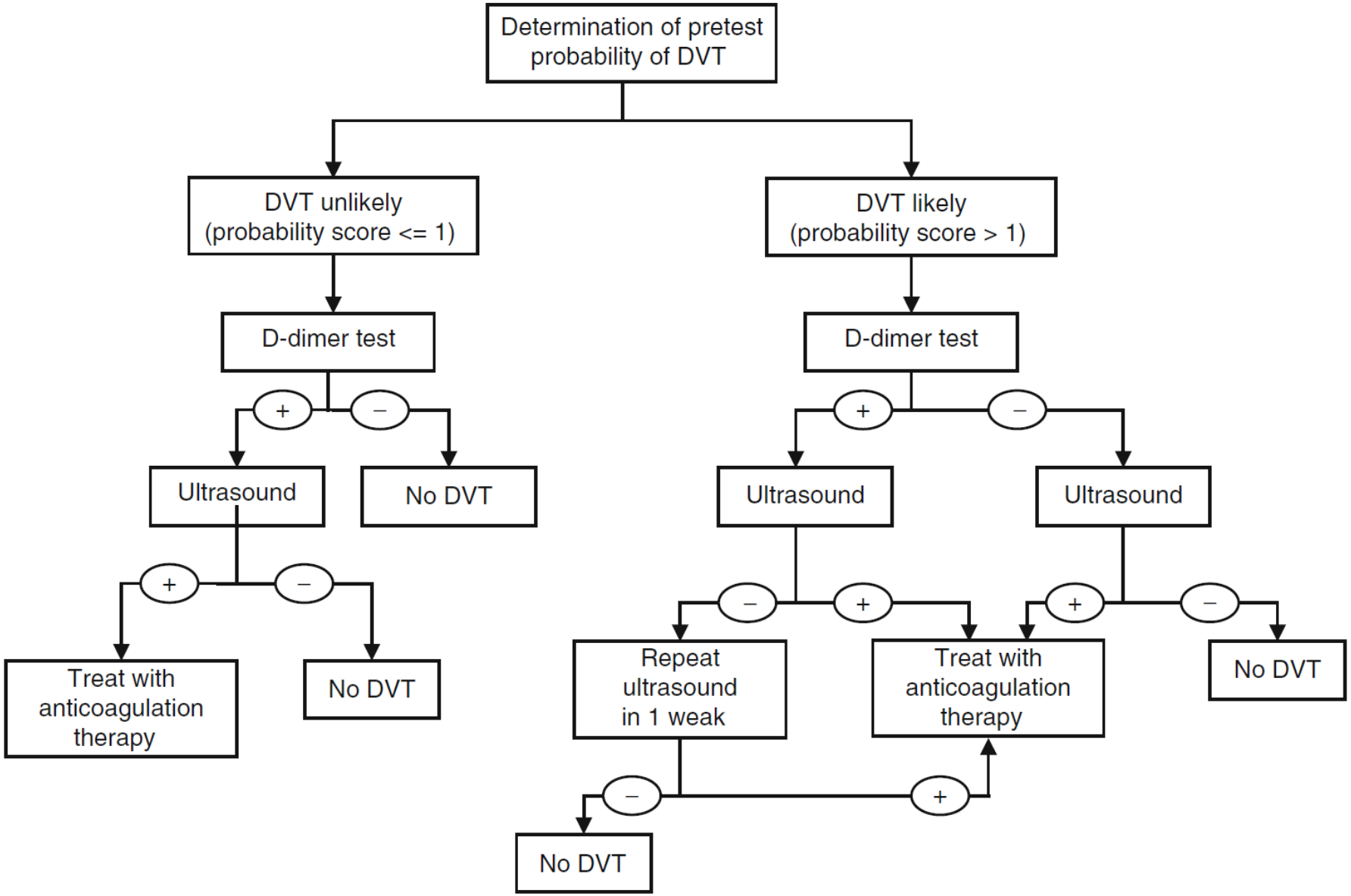
Clinical practice

- Diagnostic work-up (e.g. test ordering in renal artery stenosis)
- Starting treatment (threshold for treatment)

Therapeutic decision-making

- Intensity of treatment
- Surgical decision making
- Delaying treatment

Probability of Deep Venous Thrombosis



Applications of Prediction Models

Public health

- Targeting of preventive interventions

Framingham for CVD & BCDDP, BRCAPRO for breast cancer

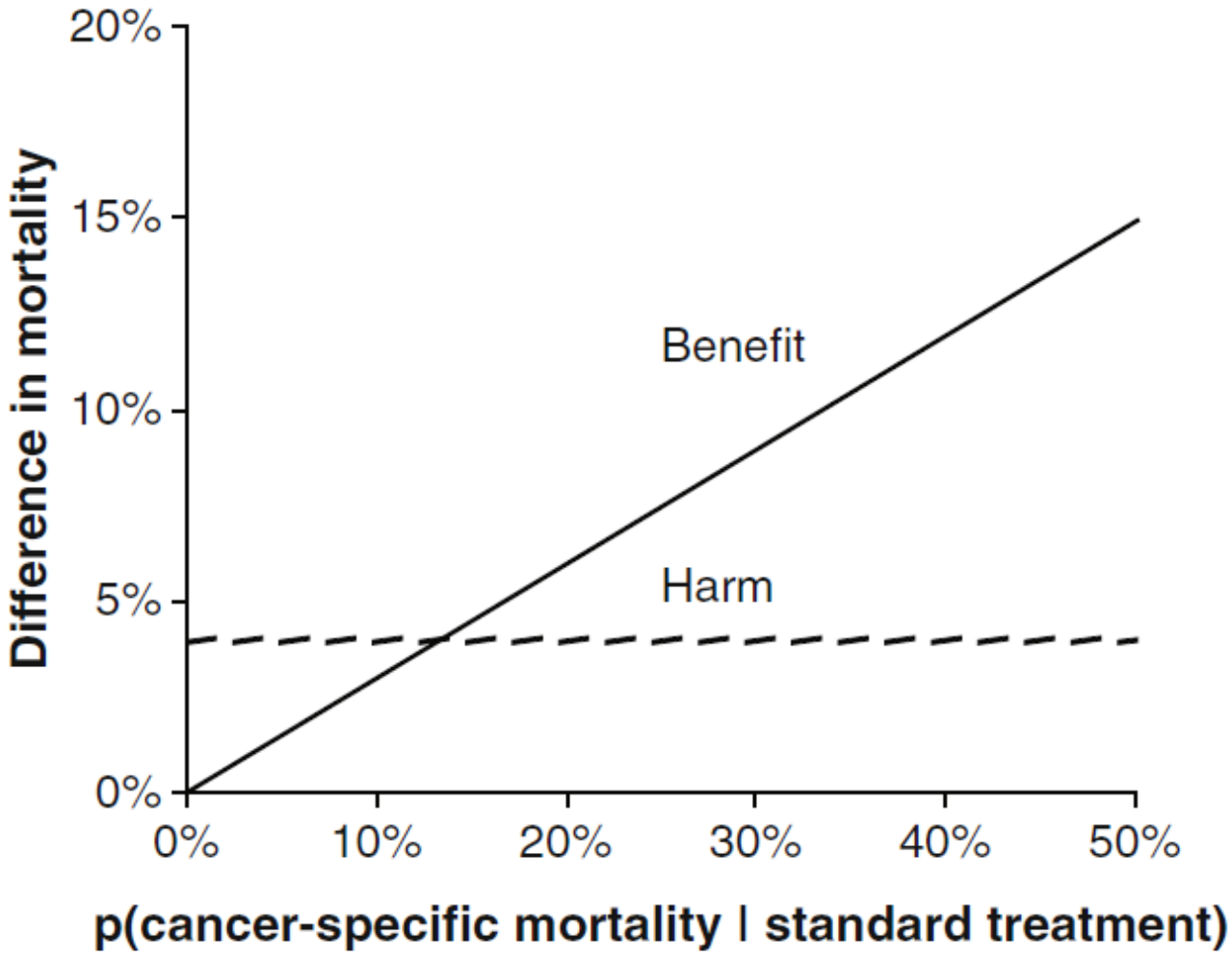
Clinical practice

- Diagnostic work-up (e.g. test ordering in)
- Starting treatment (threshold for treatment)

Therapeutic decision-making

- Intensity of treatment (e.g. chemotherapy or statin therapy)
- Surgical decision making
- Delaying treatment

Graphical illustration of weighing benefit and harm of treatment.



Applications of Prediction Models

Public health

- Targeting of preventive interventions

Framingham for CVD & BCDDP, BRCAPRO for breast cancer

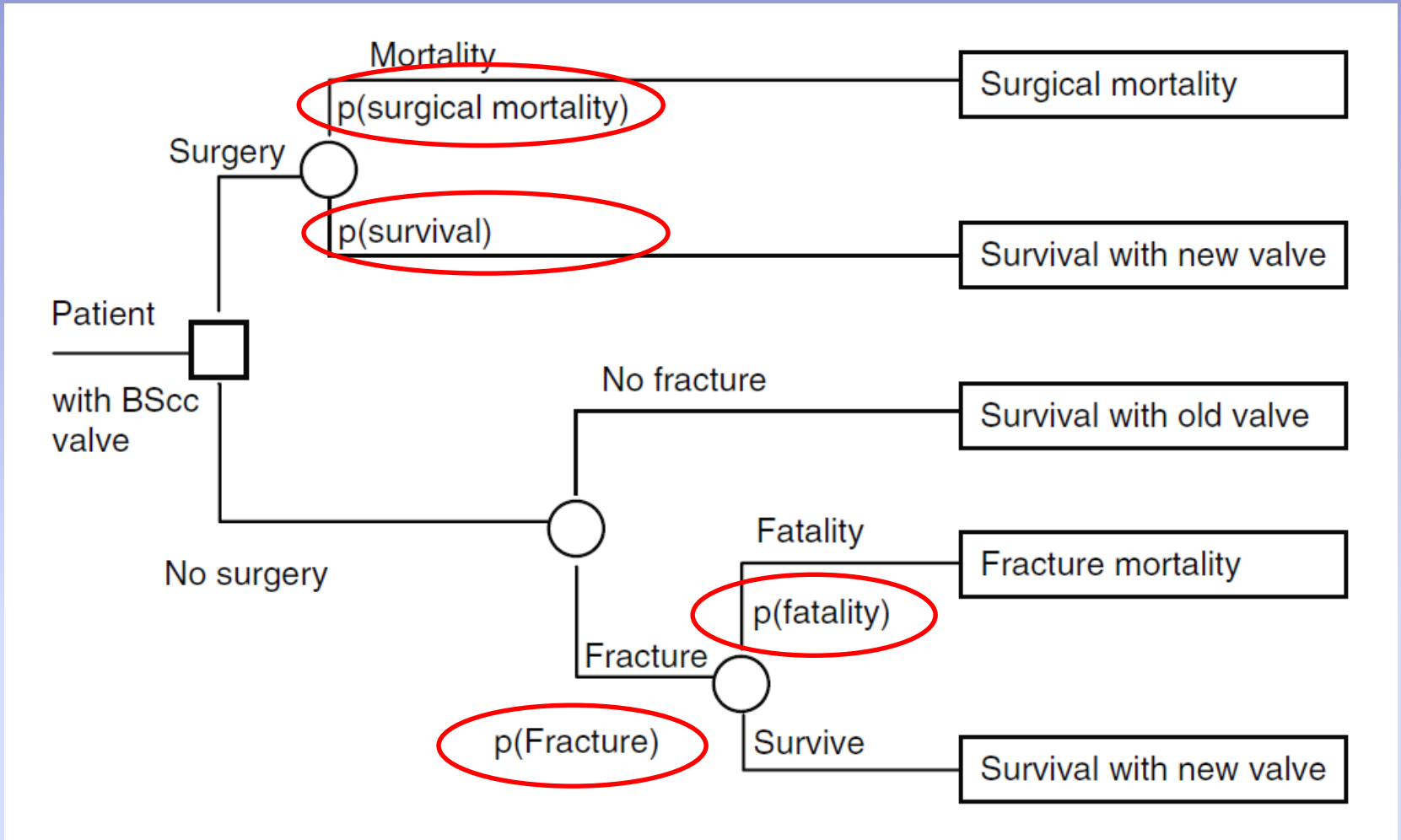
Clinical practice

- Diagnostic work-up (e.g. test ordering in)
- Starting treatment (threshold for treatment)

Therapeutic decision-making

- Intensity of treatment
- Surgical decision making
- Delaying treatment

Surgical decision-making on short-term vs. long-term risk in replacement of a risky BScc heart valve.



Patient characteristics used in the decision analysis of replacement of risky heart valves

Characteristic	Surgical risk	Survival	Fracture	Fatality fracture
Patient related				
Age (years)	+	+	+	+
Sex (male/female)		+		
Time since implantation (years)		+		
Valve related				
Position (aortic/mitral)	+	+	+	+
Opening angle (60°/70°),			+	
Size (<29 mm or >=29 mm)			+	
Production characteristics			+	
Type of prediction model	Logistic regression	Poisson regression	Poisson regression	Logistic regression

Applications of Prediction Models

Public health

- Targeting of preventive interventions

Framingham for CVD & BCDDP, BRCAPRO for breast cancer

Clinical practice

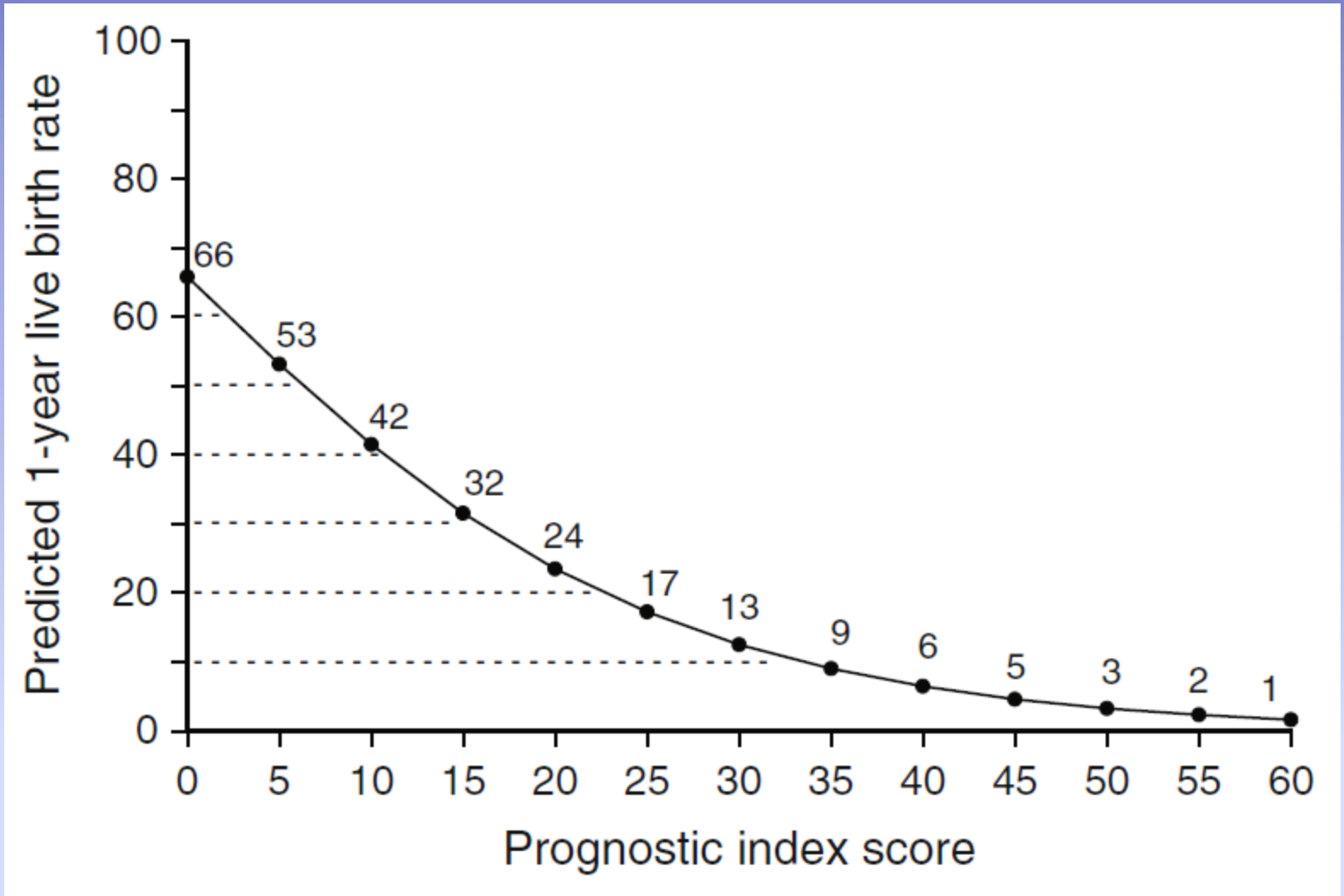
- Diagnostic work-up (e.g. test ordering in)
- Starting treatment (threshold for treatment)

Therapeutic decision-making

- Intensity of treatment
- Surgical decision making
- <u>Delaying treatment</u>

Score chart to estimate the chance of spontaneous pregnancy within 1 year

							Subfertility Score
Woman's age (years)	21–25	26–31	32–35	36–37	38–39	40–41	
Score	**0**	**3**	**7**	**10**	**13**	**15**	……..
Duration of subfertility (yrs)	1	2	3–4	5–6	7–8		
Score	**0**	**3**	**7**	**12**	**18**		……..
Type of subfertility	Secondary			Primary			
Score	**0**			**8**			
Motility (%)	≥60	40–59	20–39	0–19			
Score	**0**	**2**	**4**	**6**			……..
Referral status	Secondary care			Tertiary care			
Score	**0**			**4**			……..
				Prognostic index score (Sum)			……..


Predicted 1-year live birth rate
100
80
60
40
20
0
66
53
42
32
24
17
13
9
6
5
3
2
1
0
5
10
15
20
25
30
35
40
45
50
55
60
Prognostic index score

Why Study This Book?

The development and applications of prediction models are often suboptimal in medical publications.

Issues include:

(a) Better predictive modelling is sometimes easily possible;
e.g. a large data set
with high quality data is available, but all continuous predictors are dichotomized, which is known to have several disadvantages.

(b) Small samples are used:

- Studies are underpowered, with unreliable answers to difficult questions such as “Which are the most important predictors in this prediction problem?”

- The problem of small sample size is aggravated by doing a complete case analysis.

- Predictors are omitted that should reasonably have been included based on subject matter knowledge.

- Stepwise selection methods are abundant, which are especially risky in small data sets.

- No attempts are made towards validation, or validation is done inefficiently.

(c) Claims are exaggerated:

- Often we see statements such as ‘the predictors were identified’; in many instances such findings may not be reproducible and may largely represent noise.

- Models are not internally valid, with overoptimistic expectations of model performance in new patients.

- One modern method with a fancy name is claimed as being superior to a more traditional regression approach, while no convincing evidence exists.

(d) Poor generalizability:

- If models are not internally valid, we cannot expect them to generalize.

- Models are developed for each local situation, discarding earlier findings on effects of predictors and earlier models; a framework for continuous improvement and updating of prediction models is required.

The Book Structure

It has been found that people learn by example, by checklists, and by own discovery!

American Journal of Epidemiology Vol. 170, No. 4

Book Review

Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating

By Ewout W. Steyerberg

- **The book is aimed at <u>epidemiologists and applied biostatisticians</u> who want to develop or apply a prediction model.**

- **Only a basic knowledge of regression analysis is required, including linear regression, logistic regression, and Cox regression.**

- **The book contains few formulas and little theoretical justification, so a strong background in mathematics is not required.**

- **Also provided is R code for implementing many of the techniques discussed in the book, which will be very useful for readers to reinforce the ideas being presented.**

- **There is a discussion of issues in designing studies for prediction research, including selecting study subjects and choosing predictors and outcome variables.**

- **The book includes a 7-step checklist to be considered when developing a valid prediction model.**

- **There are enlightening discussions about missing values, selection of main effects and interactions, estimation of model parameters with shrinkage methods and evaluation of performance and clinical usefulness.**

Structure of the Book

Part I:

Prediction Models in Medicine

Part II:

Developing Valid Prediction Models

Part III:

Generalizability of Prediction Models

Part IV:

Applications


از توجه شما
سپاسگزارم