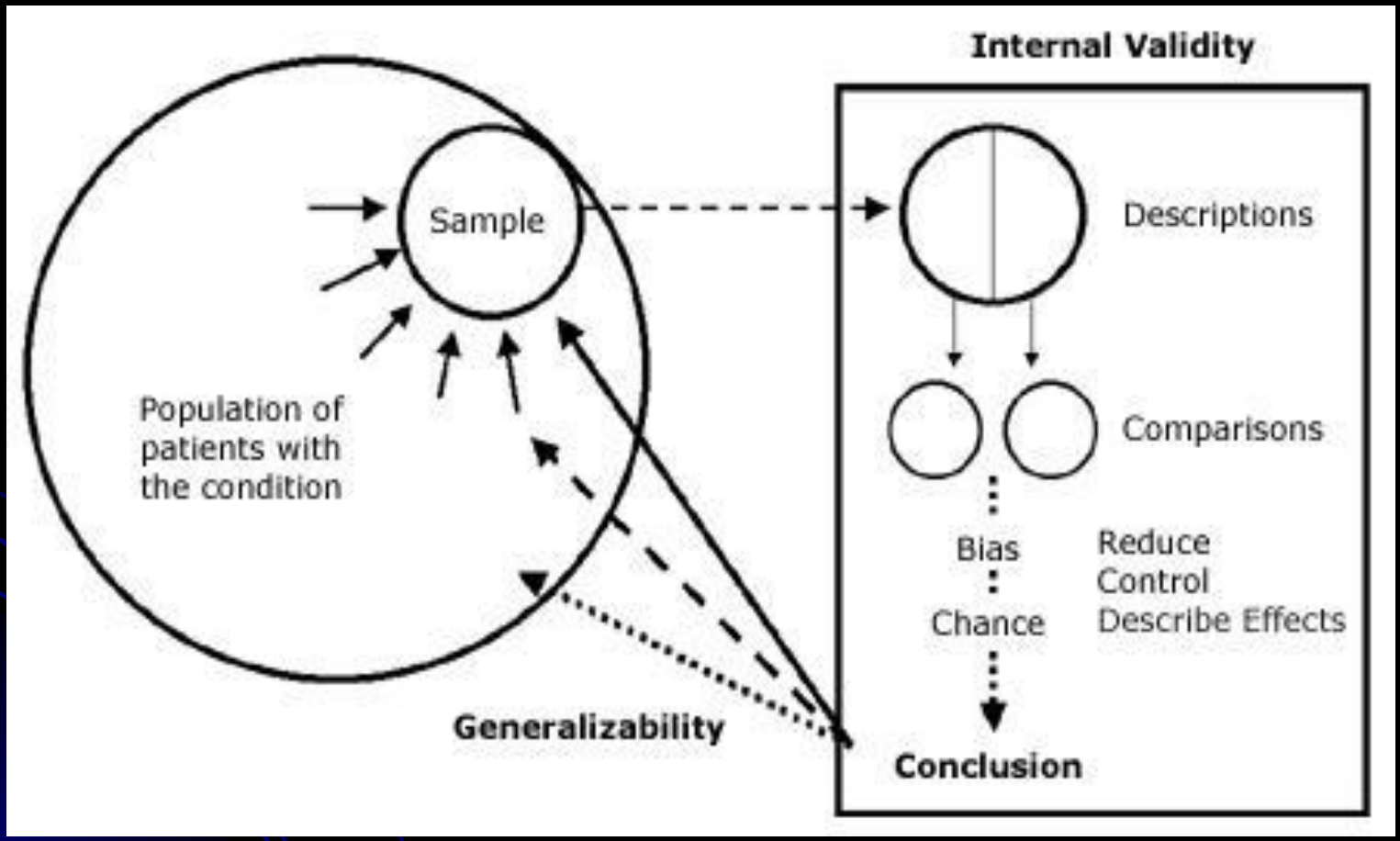
如何批判性阅读论文?

黄久佐

医学博士，主治医师

北京协和医院整形外科

Step3: Critical appraisal of evidence



Internal and external validity of research evidence

Appraising therapeutic studies

1. Are the results valid?

2. What are the results?

3. Will they help me look after my patients?

Users' Guides for an Article About Therapy

Are the results valid?

- Did intervention and control groups start with the same prognosis?
  - Were patients randomized?
  - Was randomization concealed?
  - Were patients in the study groups similar with respect to known prognostic factors?
- Was prognostic balance maintained as the study progressed?
  - To what extent was the study blinded?
- Were the groups prognostically balanced at the study's completion?
  - Was follow-up complete?
  - Were patients analyzed in the groups to which they were randomized?
  - Was the trial stopped early?

What are the results?

- How large was the treatment effect?
- How precise was the estimate of the treatment effect?

How can I apply the results to patient care?

- Were the study patients similar to my patient?
- Were all patient-important outcomes considered?
- Are the likely treatment benefits worth the potential harm and costs?

Users' Guides for an Article About Therapy

Are the results valid?

- Did intervention and control groups start with the same prognosis?
  - Were patients randomized?
  - Was randomization concealed?
  - Were patients in the study groups similar with respect to known prognostic factors?
  - Was prognostic balance maintained as the study progressed?
  - To what extent was the study blinded?
- Were the groups prognostically balanced at the study's completion?
  - Was follow-up complete?
  - Were patients analyzed in the groups to which they were randomized?
  - Was the trial stopped early?

To what extent was the study blinded?

Five Groups That Should, if Possible, Be Blind to Treatment Assignment	
Patients	To avoid placebo effects
Clinicians	To prevent differential administration of therapies that affect the outcome of interest (cointervention)
Data collectors	To prevent bias in data collection
Adjudicators of outcome	To prevent bias in decisions about whether or not a patient has had an outcome of interest
Data analysts	To avoid bias in decisions regarding data analysis

Users' Guides for an Article About Therapy

Are the results valid?

- Did intervention and control groups start with the same prognosis?
  - Were patients randomized?
  - Was randomization concealed?
  - Were patients in the study groups similar with respect to known prognostic factors?
- Was prognostic balance maintained as the study progressed?
  - To what extent was the study blinded?
- Were the groups prognostically balanced at the study's completion?
  - Was follow-up complete?
  - Were patients analyzed in the groups to which they were randomized?
  - Was the trial stopped early?

When Does Loss to Follow-up Seriously Threaten Validity?

	Trial A		Trial B	
	Treatment	Control	Treatment	Control
Number of patients randomized	1000	1000	1000	1000
Number (%) lost to follow-up	30 (3)	30 (3)	30 (3)	30 (3)
Number (%) of deaths	200 (20)	400 (40)	30 (3)	60 (6)
RRR not counting patients lost to follow-up	0.2/0.4 = 0.50		0.03/0.06 = 0.50	
RRR—worst-case scenario[a]	0.17/0.4 = 0.43		0.00/0.06 = 0	

Abbreviation: RRR, relative risk reduction.

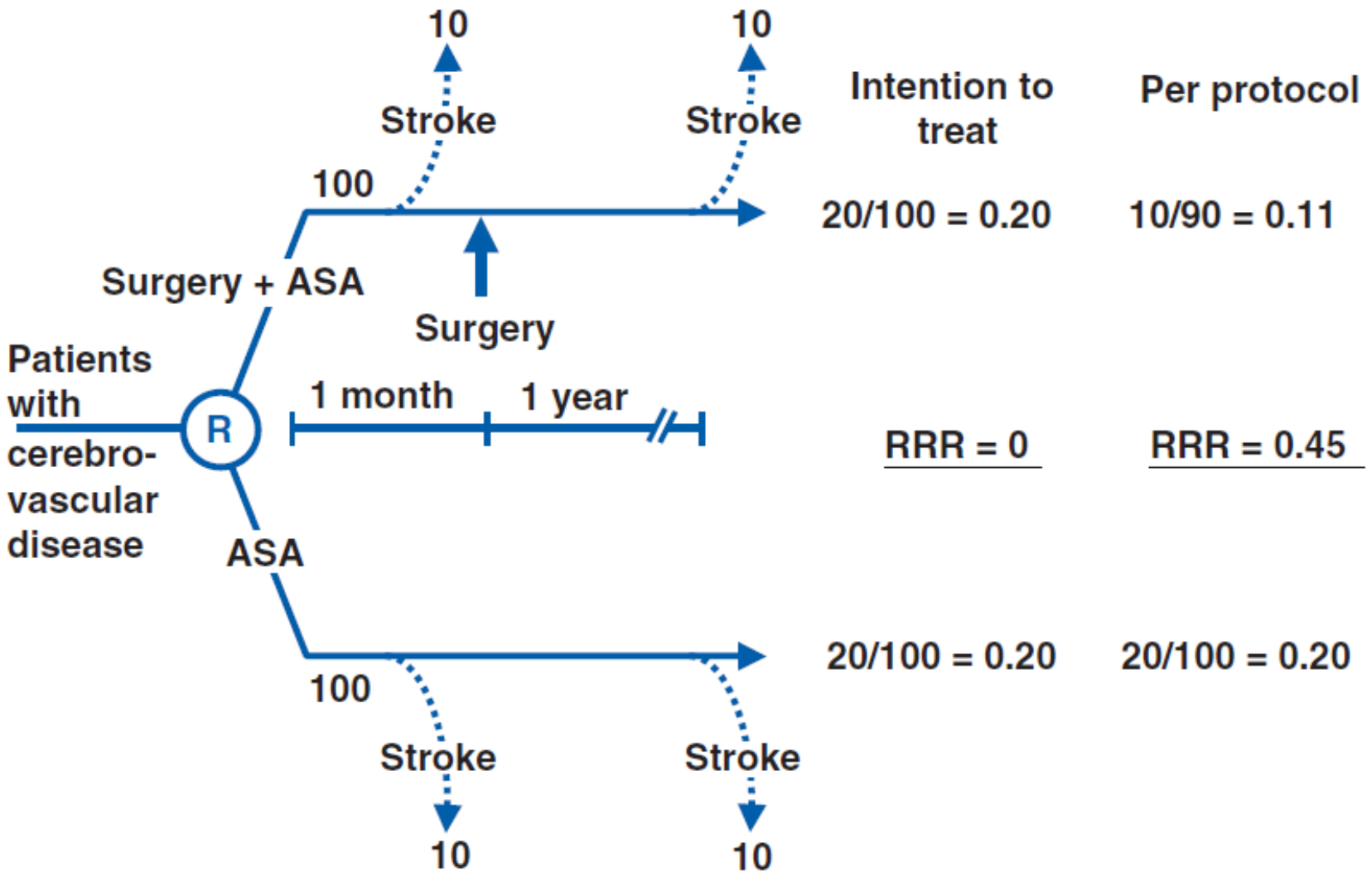
[a]The worst-case scenario assumes that all patients allocated to the treatment group and lost to follow-up died and all patients allocated to the control group and lost to follow-up survived.

Users' Guides for an Article About Therapy

Are the results valid?

- Did intervention and control groups start with the same prognosis?
  - Were patients randomized?
  - Was randomization concealed?
  - Were patients in the study groups similar with respect to known prognostic factors?
- Was prognostic balance maintained as the study progressed?
  - To what extent was the study blinded?
- Were the groups prognostically balanced at the study's completion?
  - Was follow-up complete?
  - Were patients analyzed in the groups to which they were randomized?
  - Was the trial stopped early?

Results of a Hypothetical Trial of Surgical Therapy in Patients With Cerebrovascular Disease



Abbreviations: ASA, acetylsalicylic acid; R, randomization; RRR, relative risk reduction.

Users' Guides for an Article About Therapy

Are the results valid?

- Did intervention and control groups start with the same prognosis?
  - Were patients randomized?
  - Was randomization concealed?
  - Were patients in the study groups similar with respect to known prognostic factors?
- Was prognostic balance maintained as the study progressed?
  - To what extent was the study blinded?
- Were the groups prognostically balanced at the study's completion?
  - Was follow-up complete?
  - Were patients analyzed in the groups to which they were randomized?
  - Was the trial stopped early?

Was the trial stopped early?

- Was There a Preplanned Stopping Rule?
- Did the Rule Involve Few Interim Looks and a Stringent P Value?
- Were There a Large Number of Events?

Users' Guides for an Article About Therapy

Are the results valid?

- Did intervention and control groups start with the same prognosis?
  - Were patients randomized?
  - Was randomization concealed?
  - Were patients in the study groups similar with respect to known prognostic factors?
- Was prognostic balance maintained as the study progressed?
  - To what extent was the study blinded?
- Were the groups prognostically balanced at the study's completion?
  - Was follow-up complete?
  - Were patients analyzed in the groups to which they were randomized?
  - Was the trial stopped early?

What are the results?

- How large was the treatment effect?
- How precise was the estimate of the treatment effect?

How can I apply the results to patient care?

- Were the study patients similar to my patient?
- Were all patient-important outcomes considered?
- Are the likely treatment benefits worth the potential harm and costs?

www.cebm.net

Author:	Ref:

	Description	Numbers	
Question	**P**atients		
	Intervention		
	Comparator		
	Outcomes	CER (%)	IER (%)
	1		
	2		
Appraisal	**R**andomized		
	Ascertainment		
	Measures		

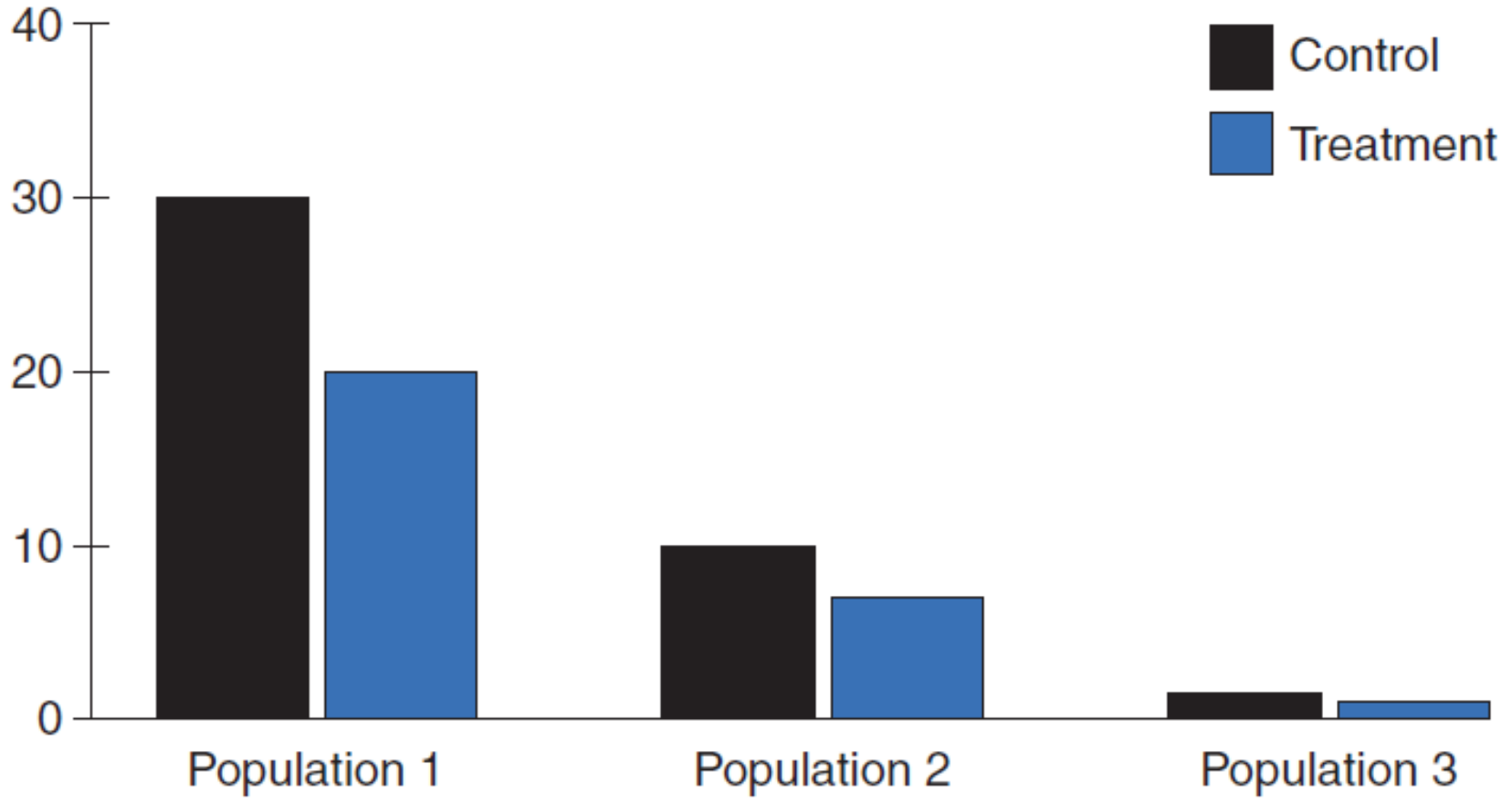
Outcomes	RDifference	CER – EER	**ARR: Absolute RR**
	RRR	RD/CER	**CI: Confidence Interval**
	NNT	1/RD	

Clinical Bottom-line:

Further Actions:

- ER: event rate
- C: control
- E: experimental
- RD: rate difference
- RRR: relative risk reduction
- NNT: number needed to treat
- CI: confidence interval
- ARR: absolute risk reduction

Constant Relative Risk With Varying Risk Differences

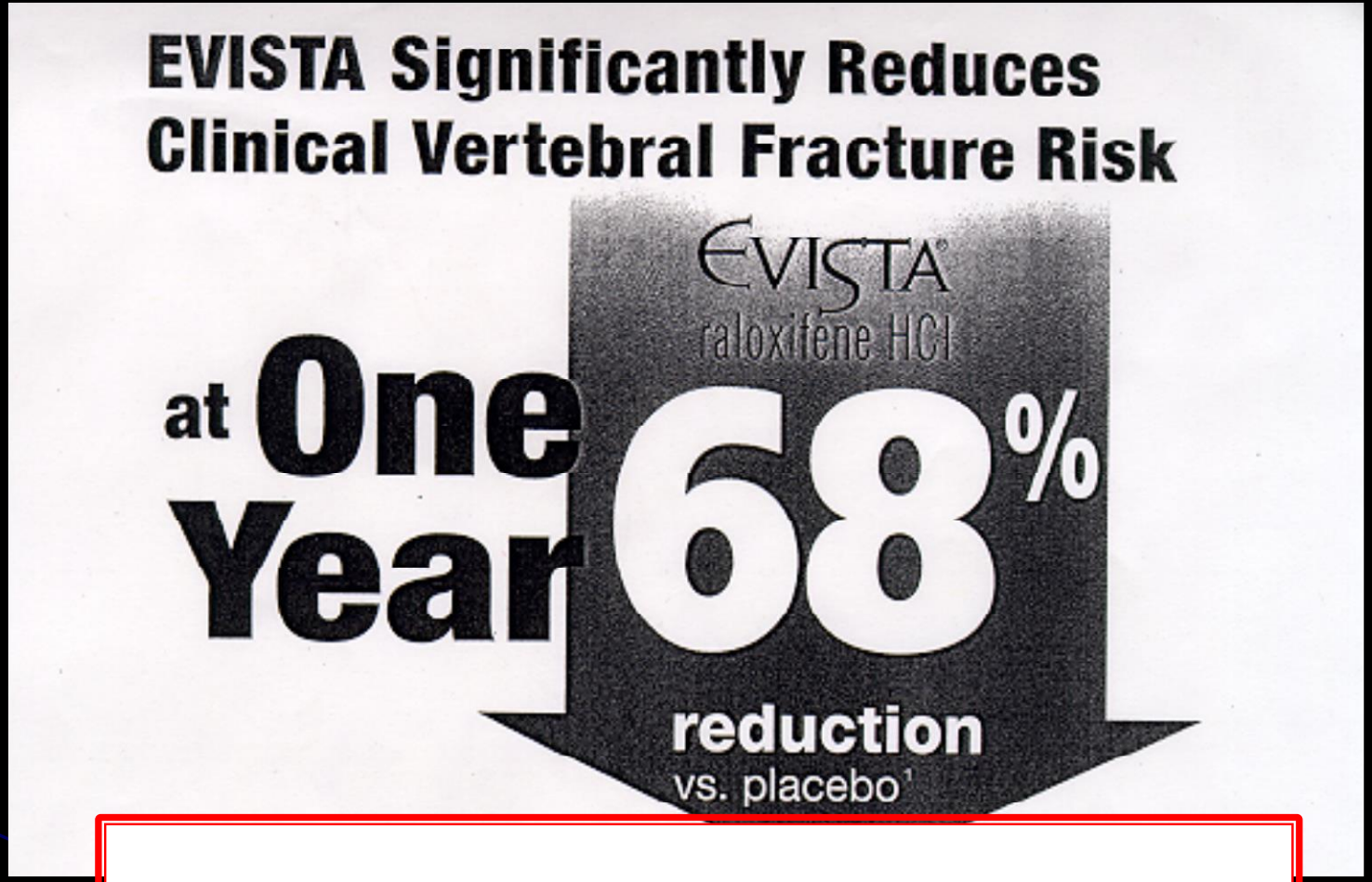


Absolute risk reduction & NNT

Relationship Among the Baseline Risk, the Relative Risk Reduction, and the Number Needed to Treat[a]

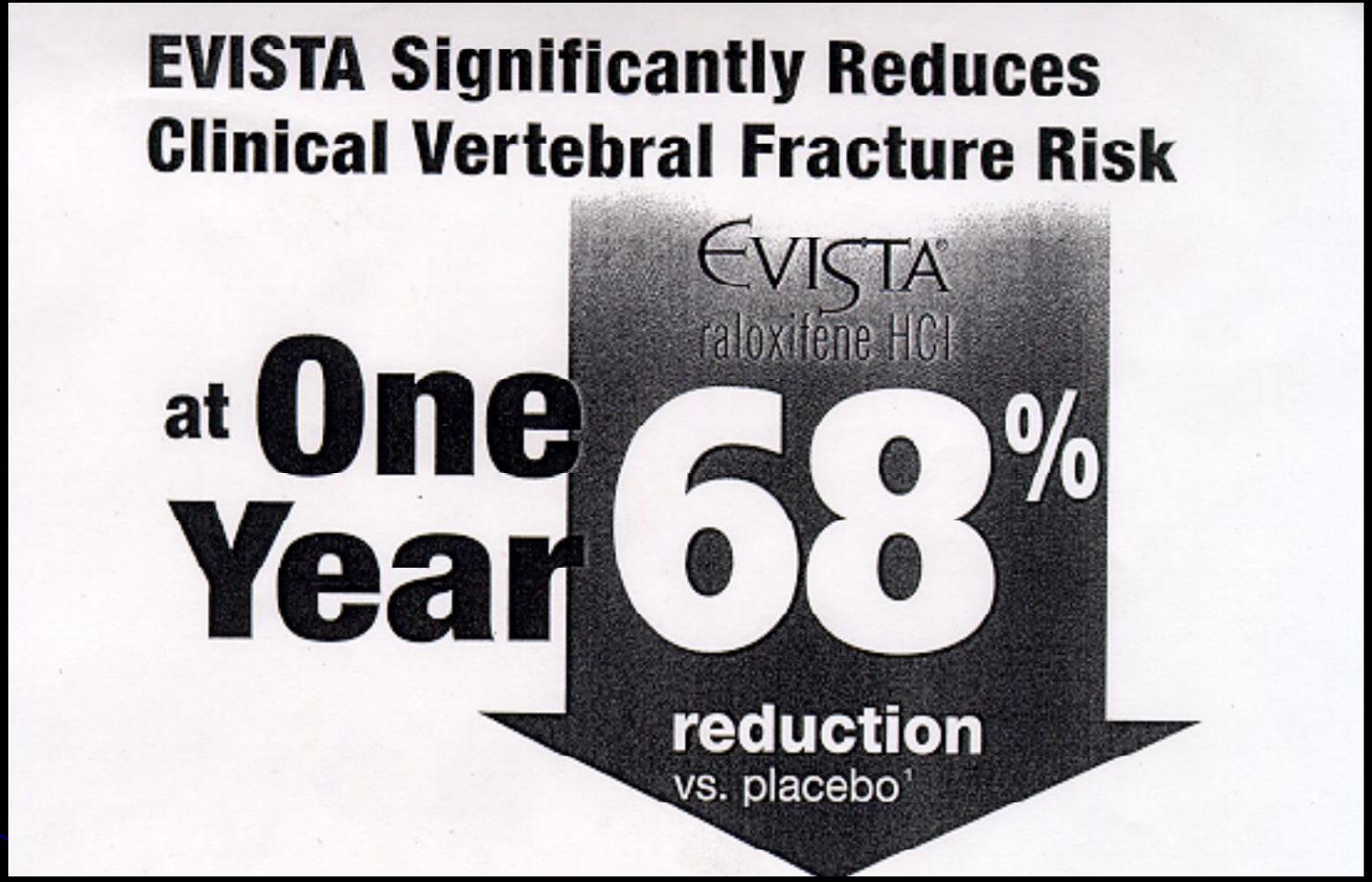
Control Event Rate	Intervention Event Rate	Relative Risk, %	Relative Risk Reduction, %	Risk Difference	Number Needed to Treat
0.02	0.01	50	50	0.01	100
0.4	0.2	50	50	0.2	5
0.04	0.02	50	50	0.02	50
0.04	0.03	75	25	0.01	100
0.4	0.3	75	25	0.1	10
0.01	0.005	50	50	0.005	200

Changes in Risk

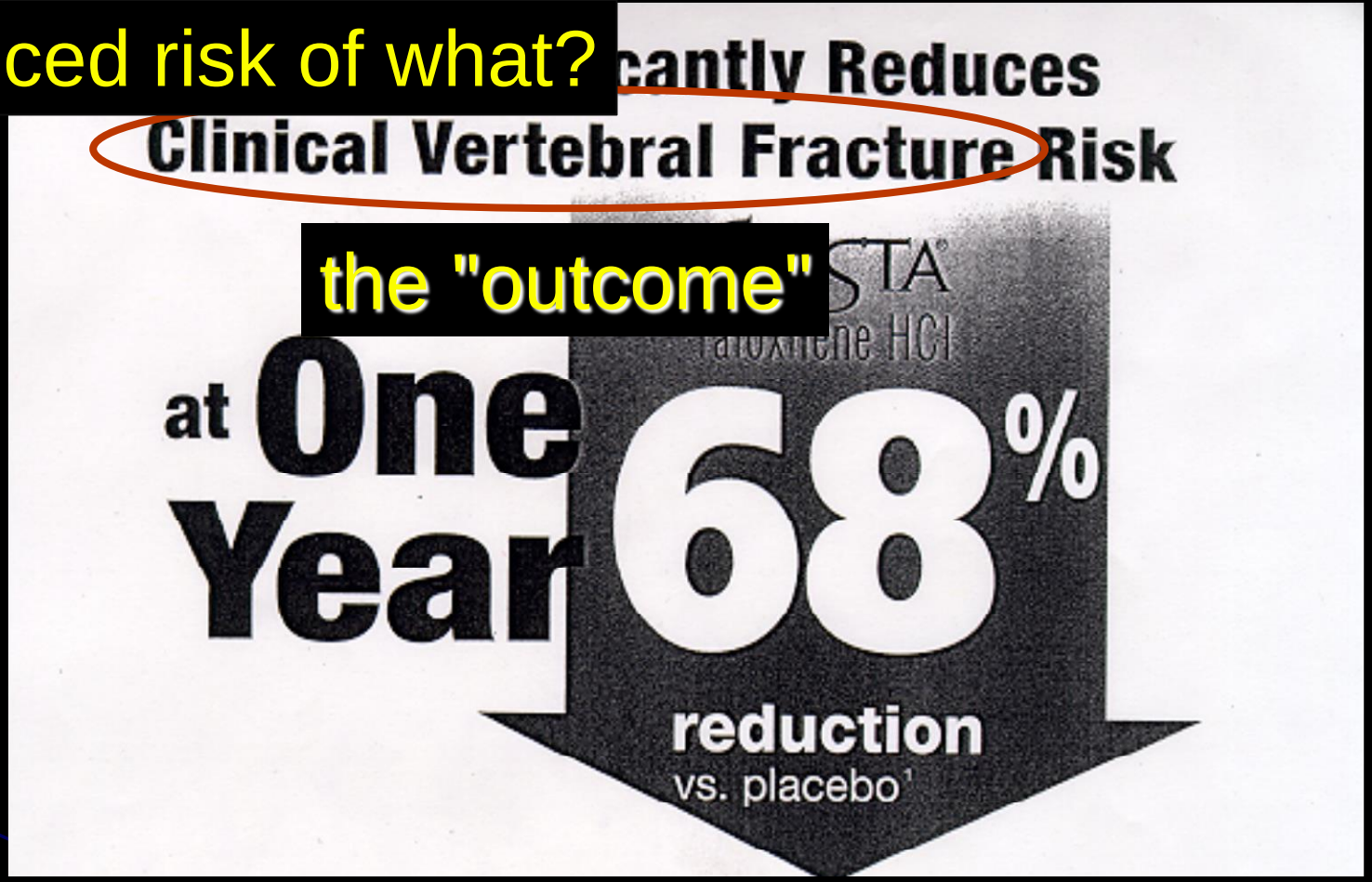


1. **How would you describe the benefit of EVISTA?** (check one)

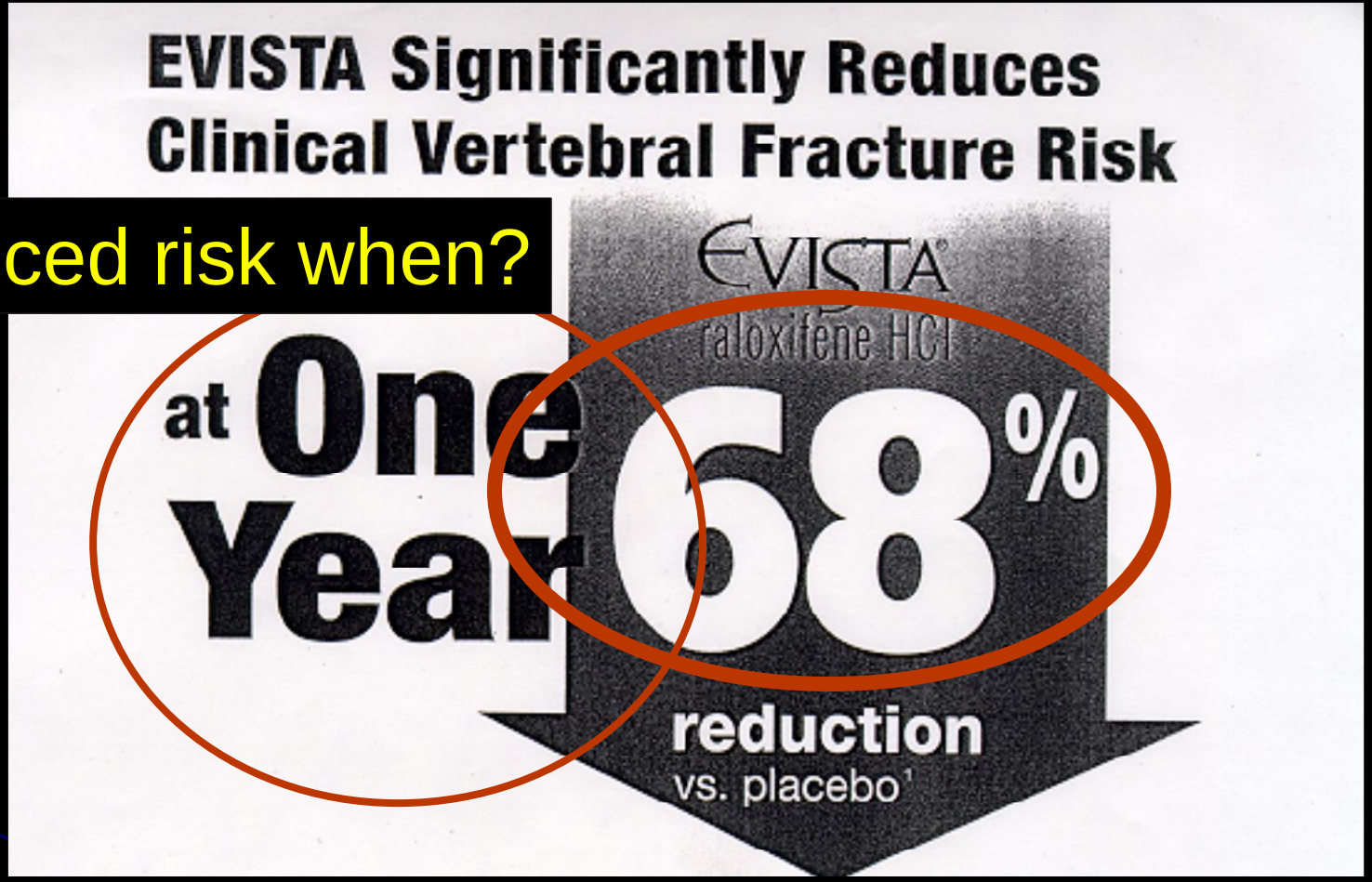
- ● Very big
- ● Big
- ● Moderate
- ● Small
- ● Very small


EVISTA Significantly Reduces
Clinical Vertebral Fracture Risk
at One Year
EVISTA®
raloxifene HCl
68%
reduction
vs. placebo[1]

Reduced risk of what?



Reduced risk when?



68% of what??

It's like a sale

Would you go if selected items were

Things like TV's, washing machines

save $100's

Things like a pack of gum?

save pennies

"68% of what" matters!

Know the REGULAR price!

Let's work through a simple example…..

How good is the sale?

How much money do you save??

Savings = $500 ice $450 rice = $50

How does the sales price compare to the original pric

$50 lower

The sales price is 10% lower than the regular price.

Let's do the same thing for EVISTA…..

How good is the sale Vertebral fracture risk

with Placebo

???

with EVISTA

???

68%OFF

How much "risk" do you save??

The science behind the ad

*What is the primary **outcome**?*

Vertebral Fracture

*Describe the distinct **exposure** in each study group*

Group A **EVISTA**

*Quantify the **outcome***

% fracture

Describe who is in the study

4,551 post-menopausal women with osteoporosis

Randomize

How long was the follow-up?

1 year (days)

Group B **Placebo**

% fracture

The chance of having a clinical vertebral fracture over 1 year in the PLACEBO group

PLACEBO

0.83%

The REGULAR price!

The "base rate"

Absolute risk in the placebo group

Event rate in the placebo group

The chance of having a clinical vertebral fracture over 1 year in the INTERVENTION group

PLACEBO	EVISTA
0.83%	0.27%

The REGULAR price!

The SALES price!

What is the effect of Evista?

How good is the sale?

PLACEBO	EVISTA
0.83%	0.27%

How much do you save?

Absolute risk reduction = 0.83% − 0.27% = 0.56%

If 1000 women took EVISTA instead of placebo for 1 year,

What is the effect of Evista?

How good is the sale?

PLACEBO	EVISTA
0.83%	0.27%

$$\text{Relative Risk} = \frac{\text{Risk of outcome (interventic}}{\text{Risk of outcome (contro}}$$

What is the effect of Evista?

How good is the sale?

PLACEBO	EVISTA
0.83%	0.27%

$$\text{Relative Risk} = \frac{???\%}{???\%} = 0.32$$

Clunky !!

The 1 year risk of [illegible] fracture for women taking EVISTA [illegible] 0.32 times that of women taking placebo

A less clunky way to describe the effect of EVISTA

So finally....this is how you get to the 68% "relative risk reductio

$$\text{Relative Risk} = \frac{0.27\%}{0.83\%} = 0.32$$

"% Lower" format =1 - =1 - .32 =.68

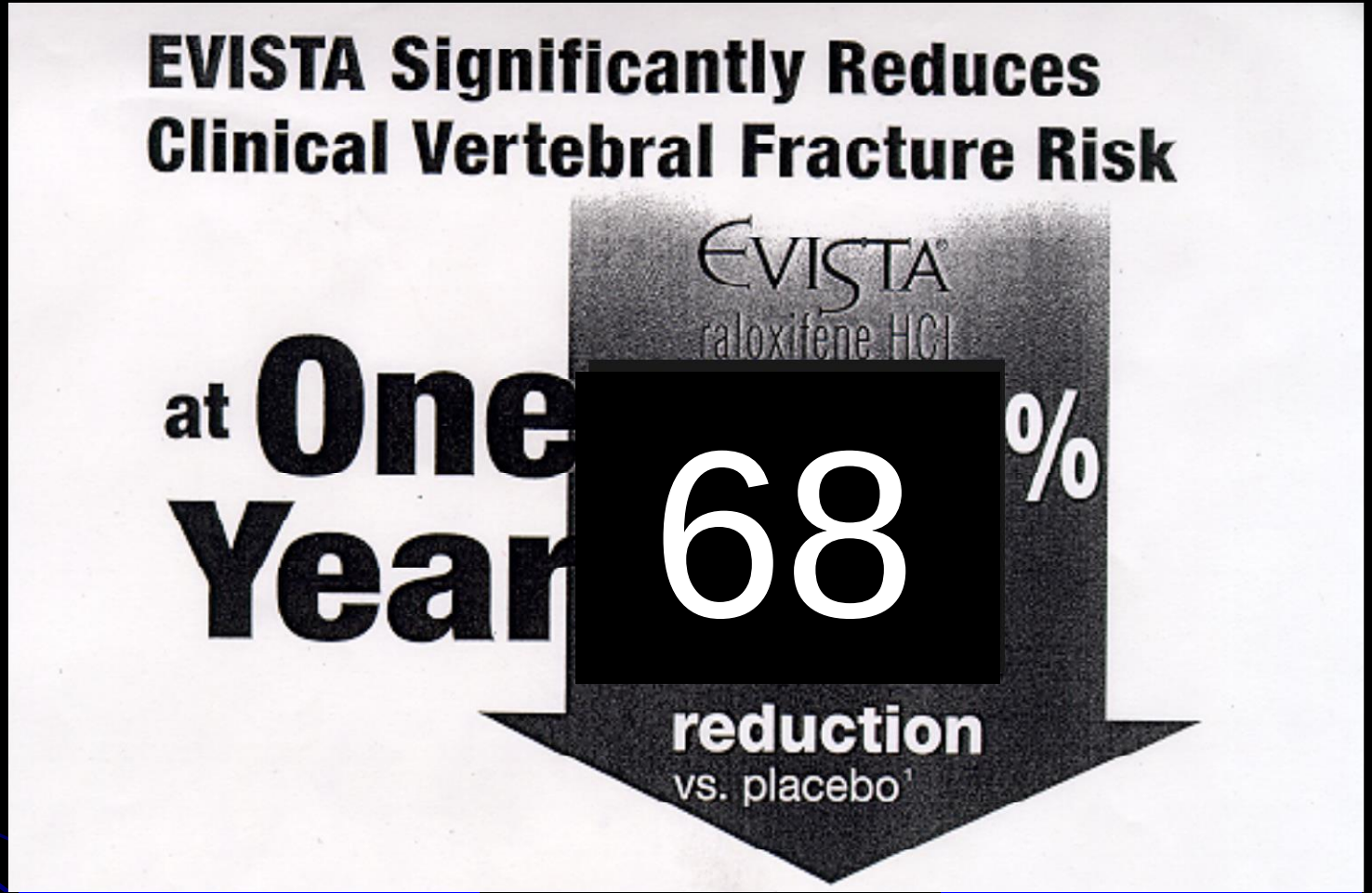
The 1 year risk of vertebral fracture for women taking EVISTA was 68 percent lower than that of women taking

This is called framing

The same information feels very different when you see the absolute risks.

Well-described finding that relative risk reduction appears more impressive than corresponding absolute risk reduction.

The EVISTA ad agency knew what they were doing....



Placebo		Evista
~0.8%	vs	~ 0.3%

Get the absolute risks for both groups!

In patients with multiple risk factors for heart disease,

Lipitor

reduces risk *of* heart attack by 36%*

3% of patients taking a sugar pill had heart attack compared to 2% of patients taking Lipitor.

If ... his ... HDL ('good' ch... terol) or smoking.

*That means in a large clinical study, 3% of patients taking a sugar pill or placebo had a heart attack compared to 2% of patients taking Lipitor.

LIPITOR®
atorvastatin calcium
tablets

Relative vs absolute risk reductions

"% off"

"savings"

Absolute risk of death at 1

Risk reductio

Group A Placebo	Group B DRUG	Relative (1-B/A)	Absolute (A-B)
30%	10%	67%	20%
3%	1%	67%	2%
0.003%	0.001%	67%	0.002%

ATTENTION!

We are NOT saying "RELATIVE RISKS" are bad.

They are good!

Efficient way to summarize 2 numbers to make comparisons – either across studies or across tr

eg, RR of death was 0.9 for drug X
0.7 for drug Y
0.3 for drug Z

ATTENTION!

We are NOT saying "RELATIVE RISKS"

Our point:
Unless the absolute risks are given, relative risks are incomplete information about effect size.

Relative risks exaggerate the magnitude of effects -- particularly when the base rate is small

.3 for drug Z

(NNT)

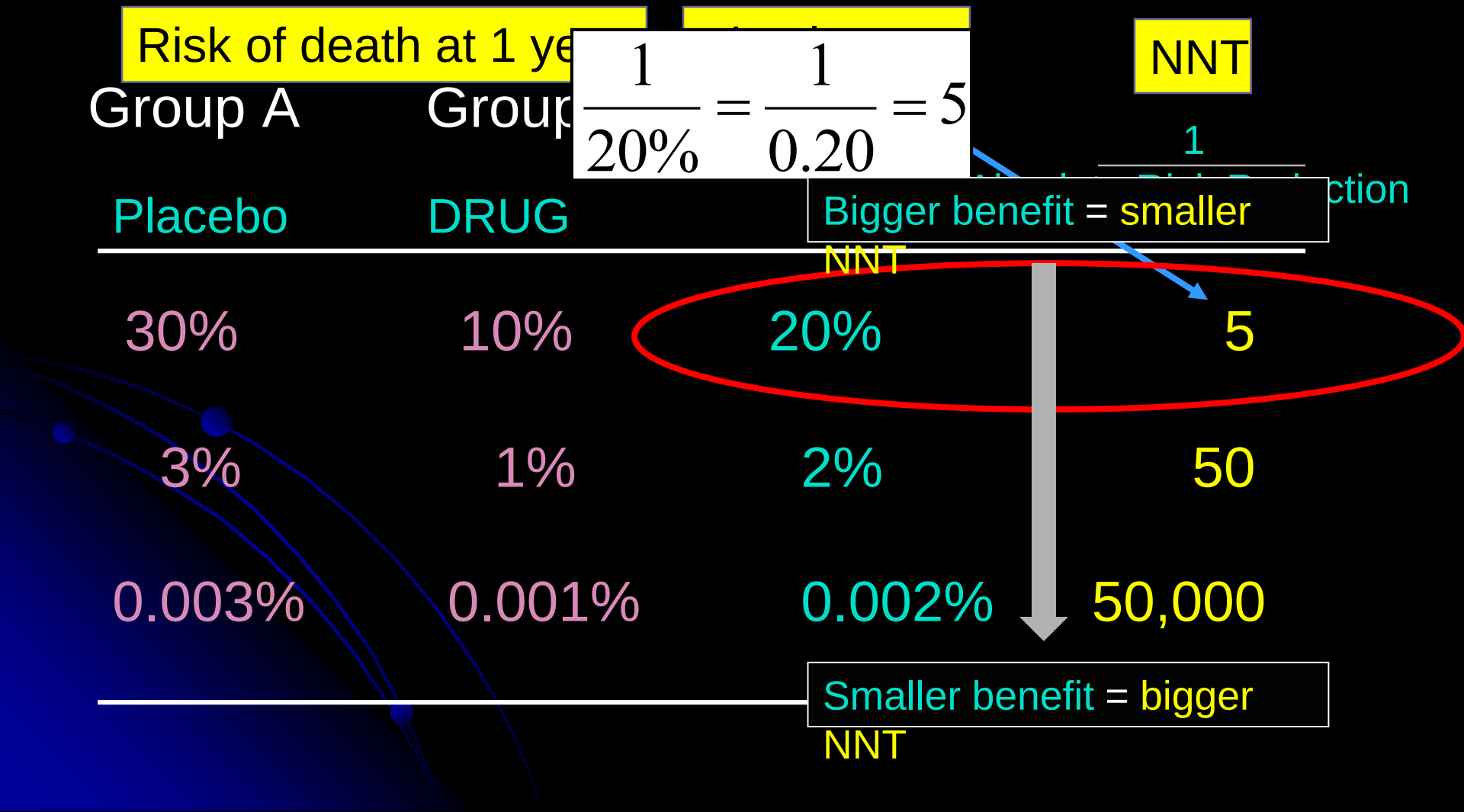
(yet one more way to talk about the same data)

$$NNT = \frac{1}{\text{Absolute risk reduction}}$$

$$= \frac{1}{(0.83\% - 0.27\%)} = 178$$

178 women would have to take Evista for a year to prevent 1 clinical vertebral fracture

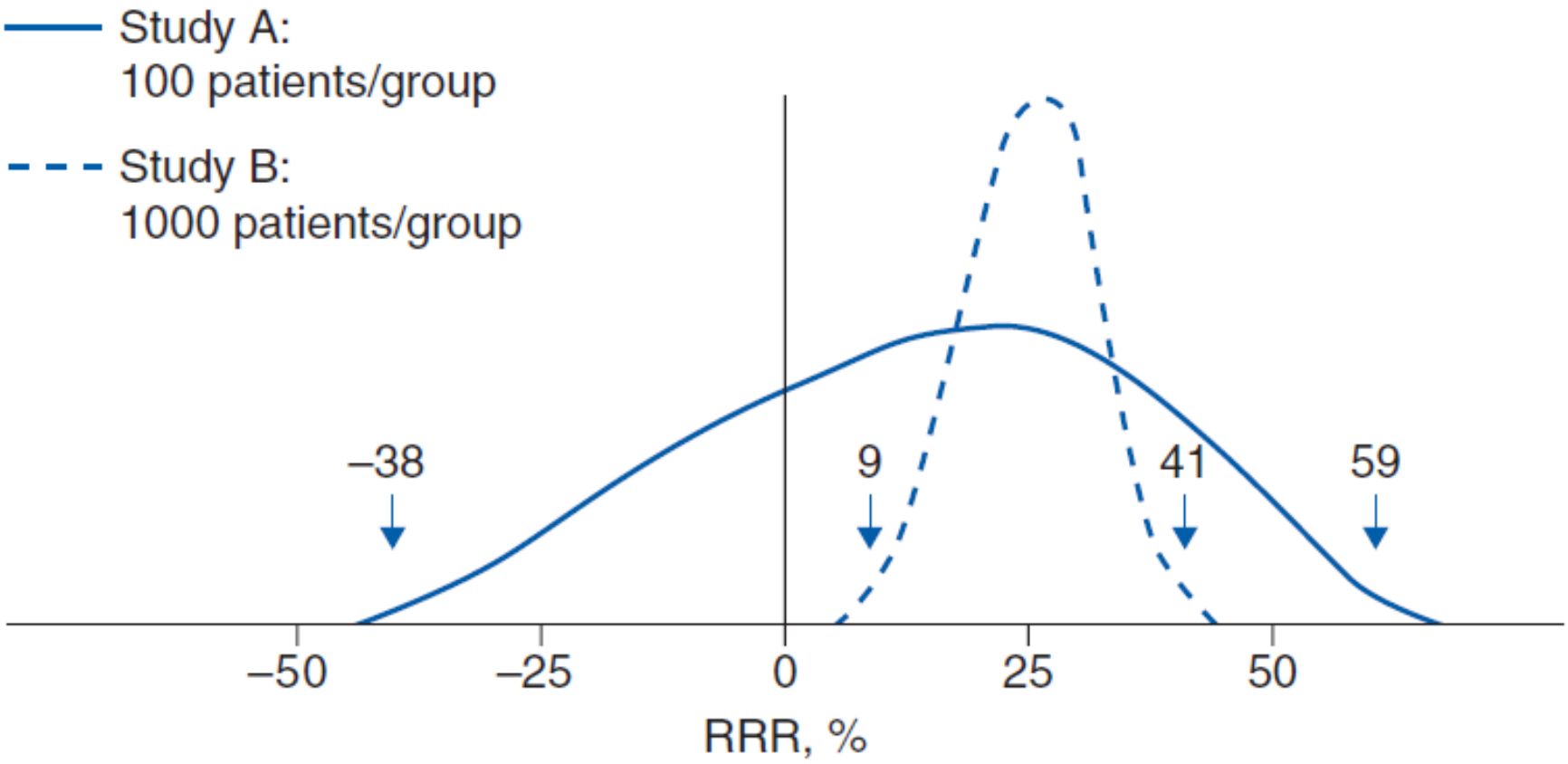
Number needed to treat



Summary
Changes in risk

- Relative risk reductions (or relative risks) are not meaningful unless you provide the base rate
  - e.g. 68% less….need to know 68% of what
- Comparing absolute risks is most helpful
  - e.g. 8 in 1000 (placebo) vs. 3 in 1000 (EVISTA)
- Relative risks are helpful to compare effect sizes of different exposures.

Confidence Intervals in Trials of Various Sample Size



Abbreviations: CI, confidence interval; RRR, relative risk reduction.

Two studies with the same point estimate, a 25% RRR, but different sample sizes and correspondingly different CIs. The x-axis represents the different possible RRR, and the y-axis represents the likelihood of the true RRR having that particular value. The solid line represents the CI around the first example, in which there were 100 patients per group, and the number of events in active and control was 15 and 20, respectively. The broken line represents the CI around the second example in which there were 1000 patients per group, and the number of events in active and control was 150 and 200, respectively.

CI & different sample size

Confidence Intervals Around the Relative Risk Reduction for the Hypothetical Results of 5 Successively Larger Trials

Control Event Rate	Treatment Event Rate	Relative Risk, %	Relative Risk Reduction (RRR), %	Intuitive Confidence Interval, %	Calculated 95% Confidence Interval Around the RRR, %
2/4	1/4	50	50	–50 to 90	–174 to 92
10/20	5/20	50	50	–20 to 90	–14 to 79.5
20/40	10/40	50	50	0 to 90	9.5 to 73.4
50/100	25/100	50	50	20 to 80	26.8 to 66.4
500/1000	250/1000	50	50	40 to 60	43.5 to 55.9

Users' Guides for an Article About Therapy

Are the results valid?

- Did intervention and control groups start with the same prognosis?
  - Were patients randomized?
  - Was randomization concealed?
  - Were patients in the study groups similar with respect to known prognostic factors?
- Was prognostic balance maintained as the study progressed?
  - To what extent was the study blinded?
- Were the groups prognostically balanced at the study's completion?
  - Was follow-up complete?
  - Were patients analyzed in the groups to which they were randomized?
  - Was the trial stopped early?

What are the results?

- How large was the treatment effect?
- How precise was the estimate of the treatment effect?

How can I apply the results to patient care?

- Were the study patients similar to my patient?
- Were all patient-important outcomes considered?
- Are the likely treatment benefits worth the potential harm and costs?

Take Home Messages

- 如何批判性阅读治疗学文献
- 绝对值和相对值

Critical Appraial: 3 easy steps

1. Are the results valid?

2. What are the results?

3. Will they help me look after my patients?