

Critical Appraisal

Dr. Nada A. Al-Yousefi

**Assistant Professor & Consultant
Family Medicine Unit
Department of Family and Community
Medicine**

**College of Medicine(KSU)
nalyousefi@ksu.edu.sa**

Feb 27, 2014





What's A Paper on Therapy?

- Clinical Trial (Controlled) Compares

INTERVENTION
with
CONTROL

Clinical Trial Compares

- **INTERVENTION**

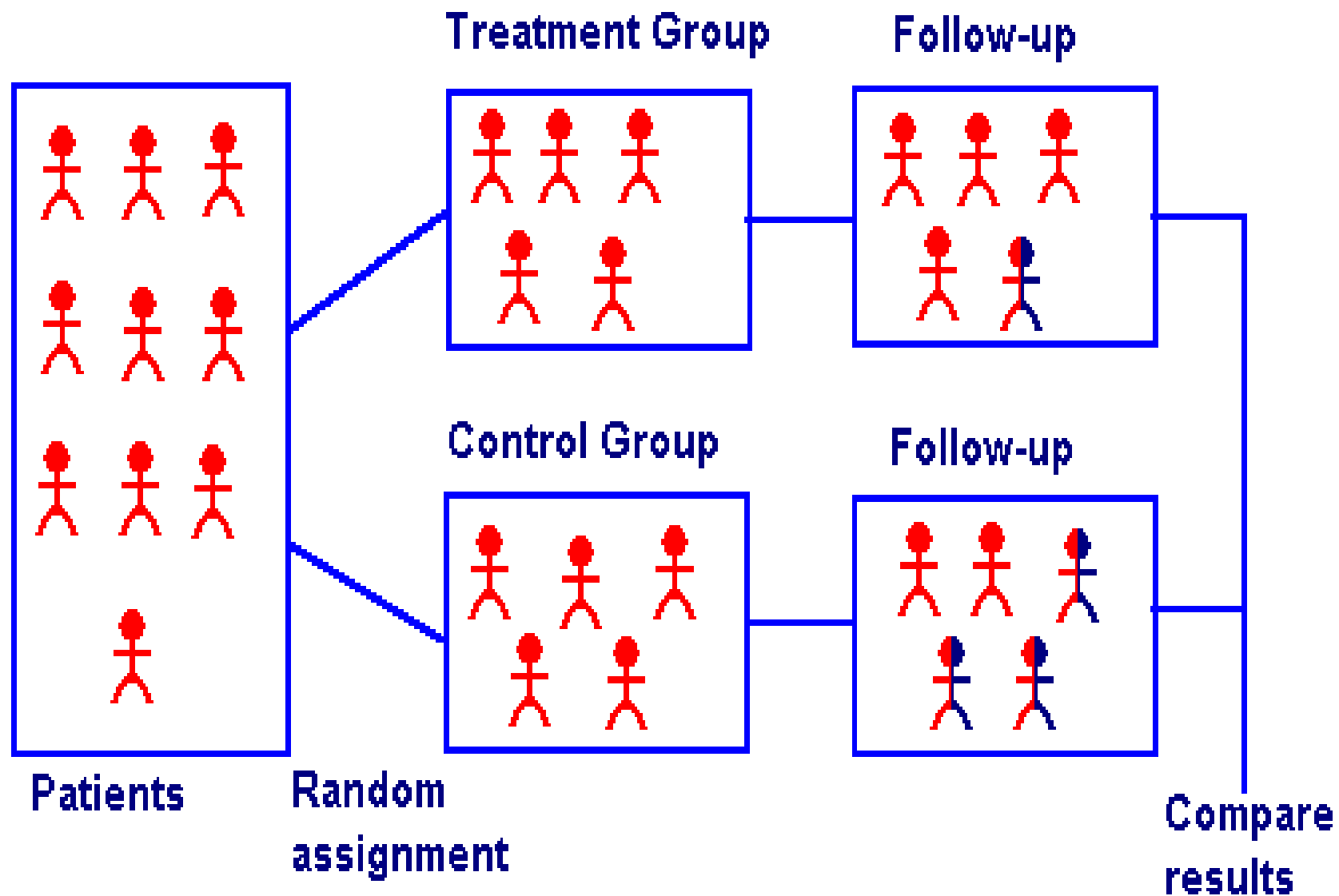
- Drug (New)
- Structured exercise program (e.g. osteoporosis)
- Surgical procedure

- **CONTROL**

- Placebo, old drug or old intervention
- Usual regular advise given (osteoporosis)
- Another surgical procedure / No surgery

process of RCTs

- ✓ Preparation: Randomization, Computer generated list
- ✓ Eligibility assessment (Inclusion/exclusion)
- ✓ Consent
- ✓ Allocation to study arms (Concealment)
- ✓ Baseline assessment
- ✓ Initiation of intervention (Blind)
- ✓ Follow-up
- ✓ Outcome assessment
- ✓ Data analysis



Appraise the Evidence

- Assess **validity**? Correctness (likely to be true)
- What are the **results**? Clinically important
- Can we **apply** the results to our patient? Applicable in and useful for my patients

USERS' GUIDES TO THE MEDICAL LITERATURE

A Manual for Evidence-Based Clinical Practice

The Evidence-Based
Medicine Working Group

Edited by
Gordon Guyatt, MD
Drummond Rennie, MD
Robert Hayward, MD (*interactive guides*)

AMA
press

JAMA
&
ARCHIVES
JOURNALS
American Medical Association

APPLICABILITY

RESULTS

VALIDITY

VALIDITY

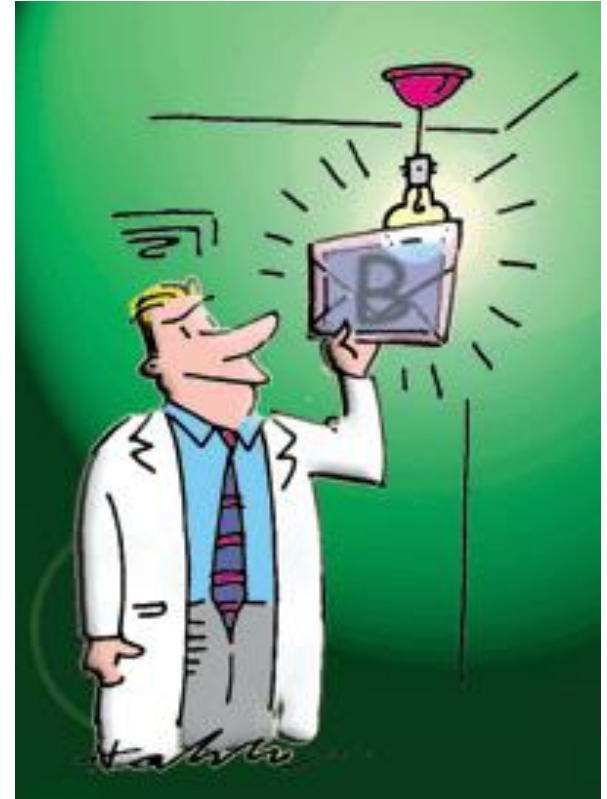
- Randomization.
- Concealment.
- Blindness.
- Follow up complete.
- Intention to treat.
- Similar groups at start.
- Both groups treated equally.

Randomization

- Randomisation = similar groups at baseline
- Equal (50%) chance to be in either group
- How was it randomized?
- Was randomization concealed?
 - selection
 - allocation

concealed allocation

- Did investigators know to which group the potential subject would be assigned **before** enrolling them?
- Trials with unconcealed allocation consistently overestimate benefit by ~40%



Selection bias

Reduced by:

- ✓ centralised randomisation
- ✓ on-site computer system with group assignments in a locked file
- ✓ sequentially numbered, sealed, opaque envelopes
- ✗ Not: alternation, dates of birth, day of week.

Blindness

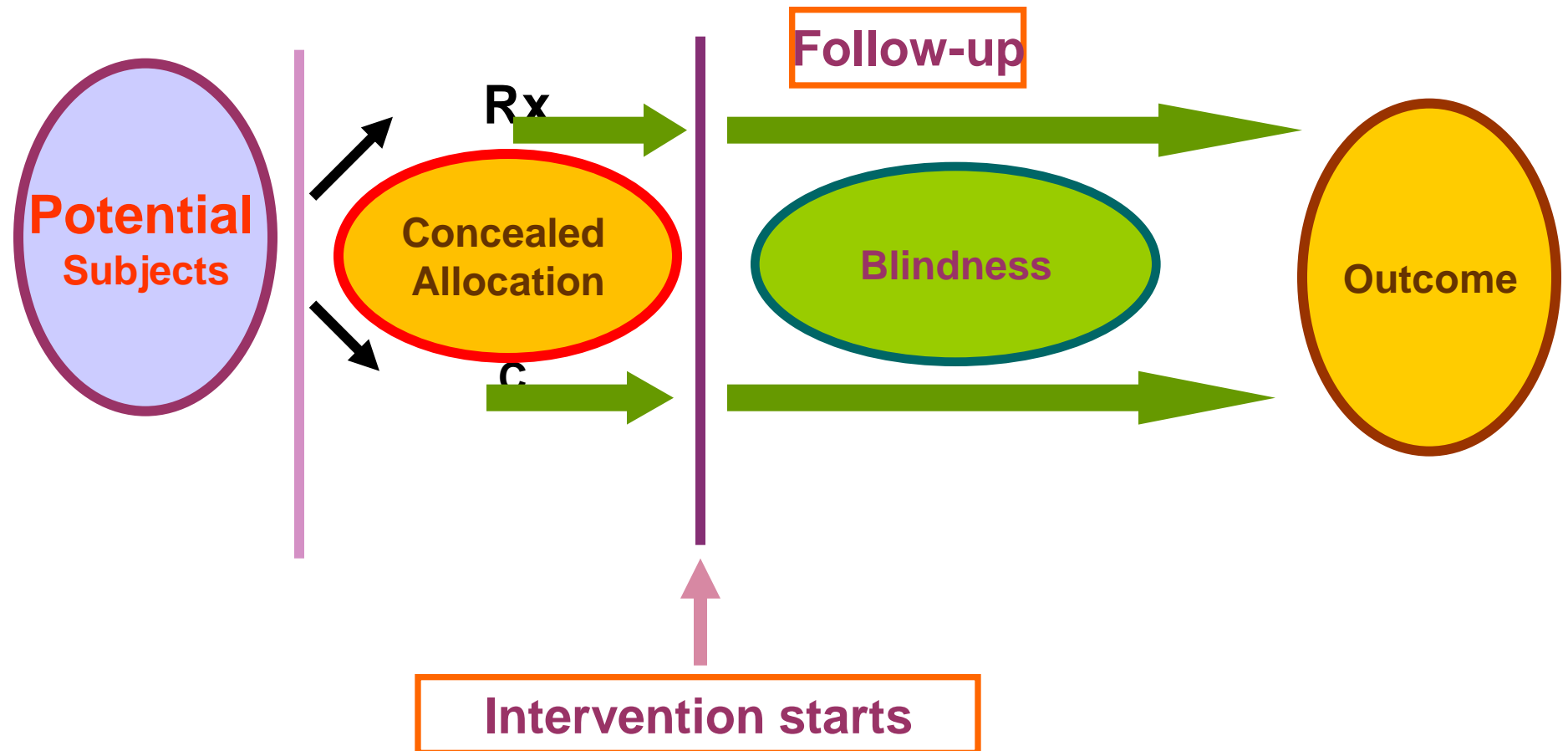
- Who is Blind?
- - Physicians-Nurses-Patients-Data gathering staff- data analyzers.
- - Single, double...

Blindness

- If patient knows: ***Placebo effect*** Those who are on effective treatment perform better than those who receive Placebo
- If Physician knows: ***Overestimate Treatment effect*** (More care, Co-intervention)

Selection bias•

Performance bias•



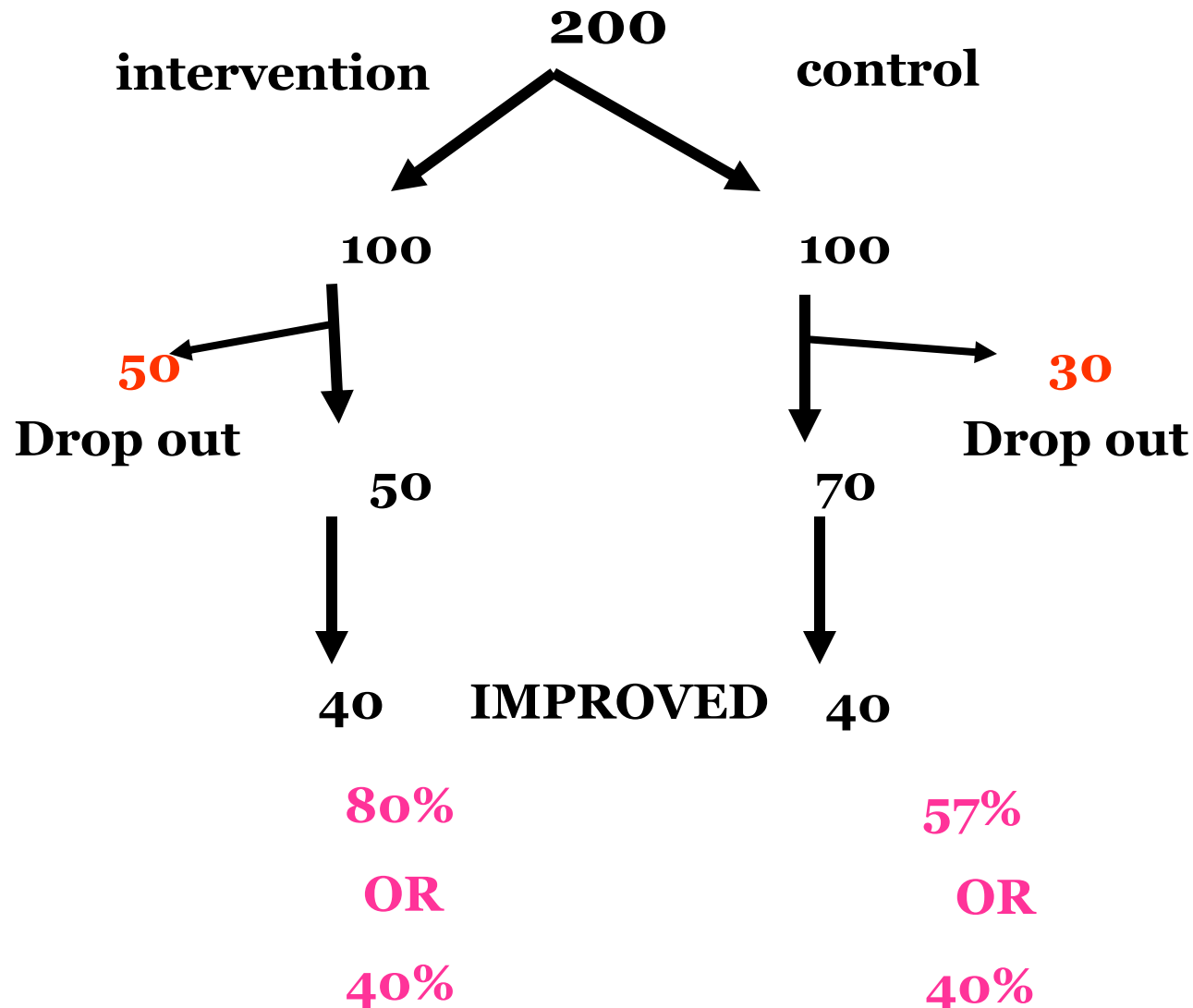
Follow up

- duration of study.
- drop out < 20%.

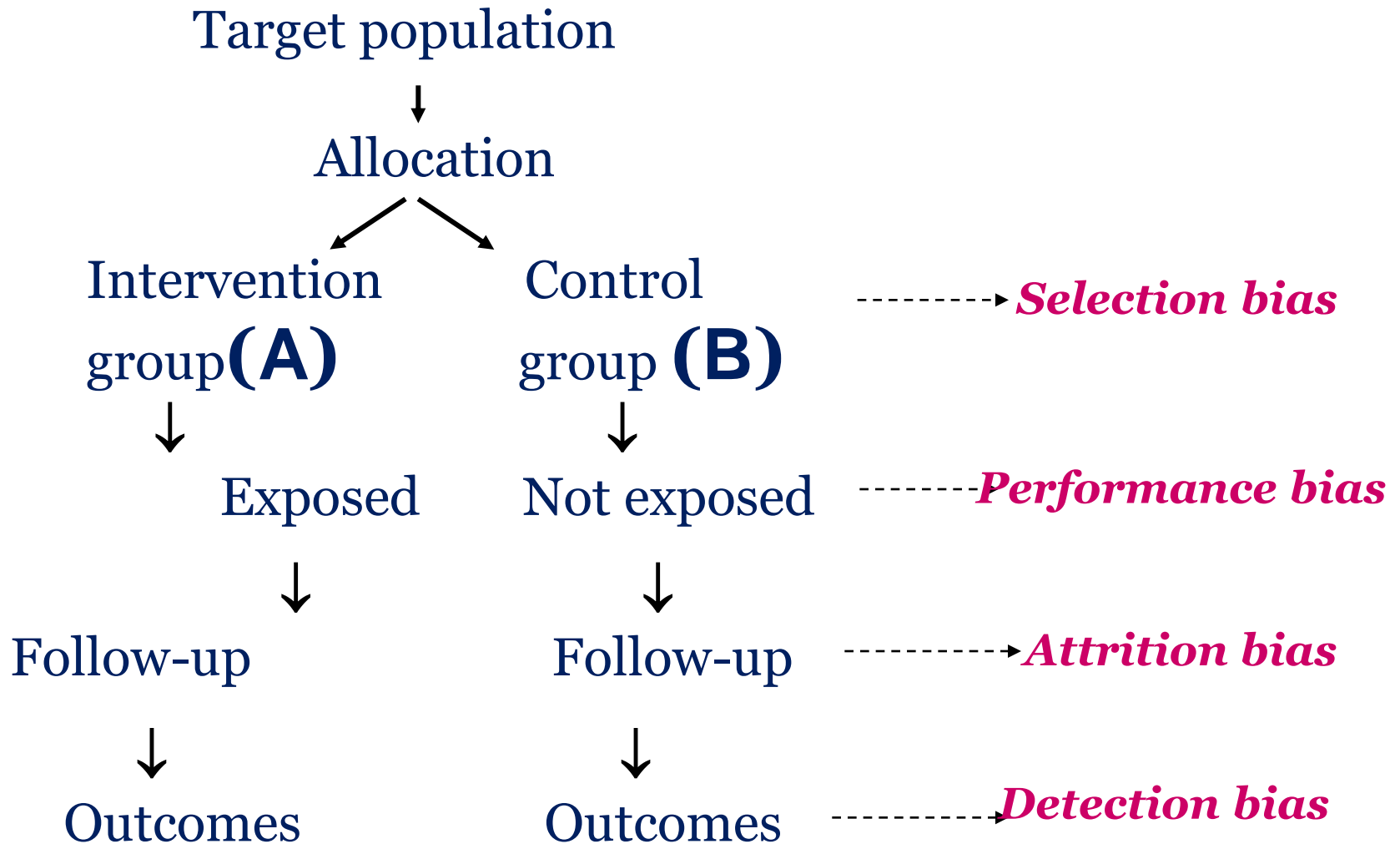
INTENTION TO TREAT

All patients analyzed in the groups
to which they were allocated

INTENTION TO TREAT (ITT)



Sources of bias in trials



How RCTs differ from other designs

Two balanced groups:

- **Start Balanced:** All prognostic factors are equally distributed at the start (**Concealed Randomization**)
- **Run Balanced:** All prognostic factors are maintained balanced throughout the study (**Blindness and the 3C**)
- **End Balanced:** All prognostic factors are maintained balanced at the end of the study (**ITT**)

Intervention

Figure 1. Profile of the Estrogen Plus Progestin Component of the Women's Health Initiative

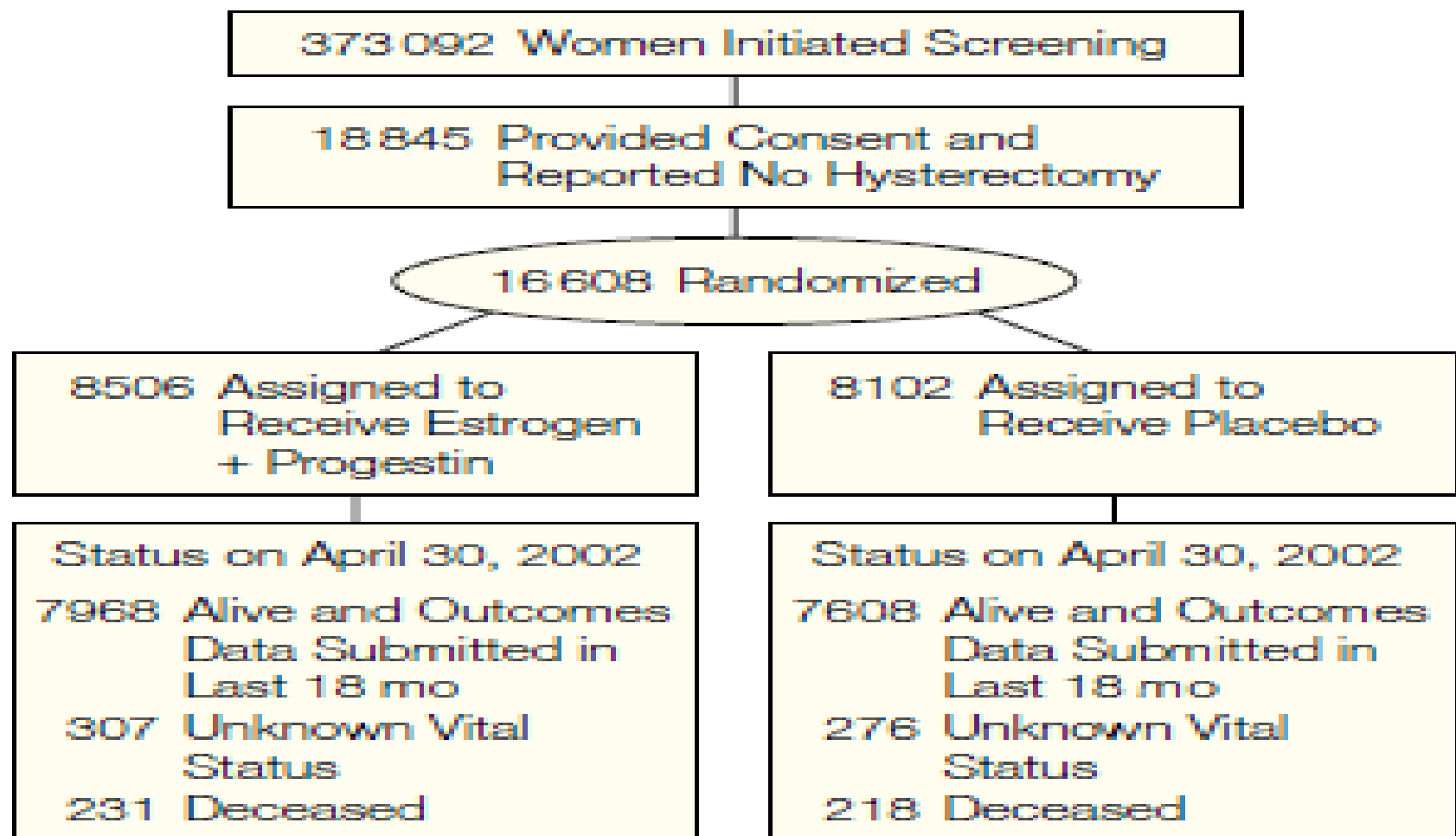
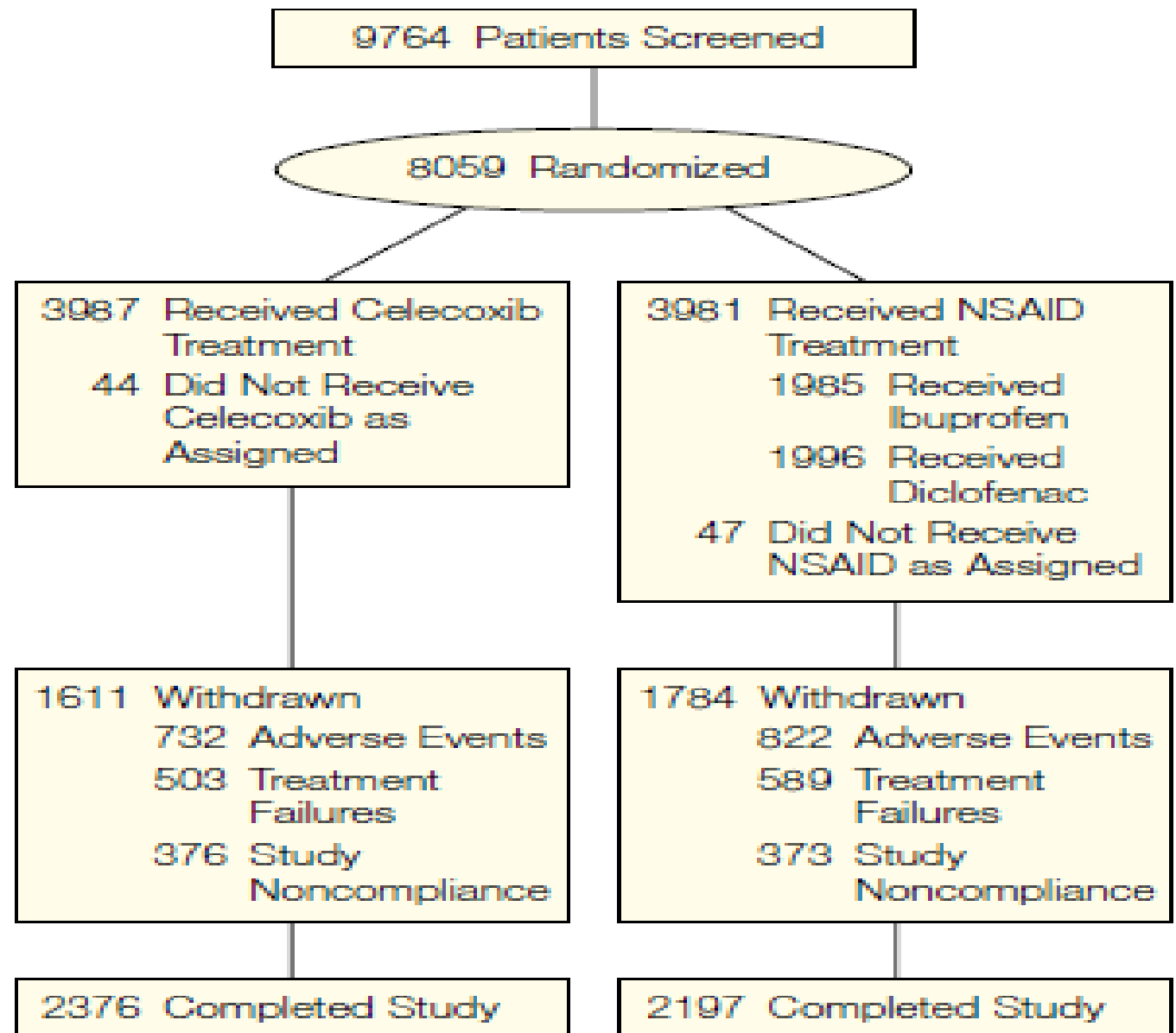


Figure 1. Flowchart of Patient Disposition at 6 Months



Menerba® Phase 3 Advanced Clinical Trial Design

PI: Wulf Utian, M.D., Ph.D. D.Sc.

Founder and President Emeritus of the North American Menopause Society (NAMS)

40 U.S. Clinical Sites Approved and Trained

Eligibility:

- Postmenopausal women ages 40-65
- ≥ 7 moderate to severe hot flashes/day

Randomization:

1200 patients to 12 weeks of treatment

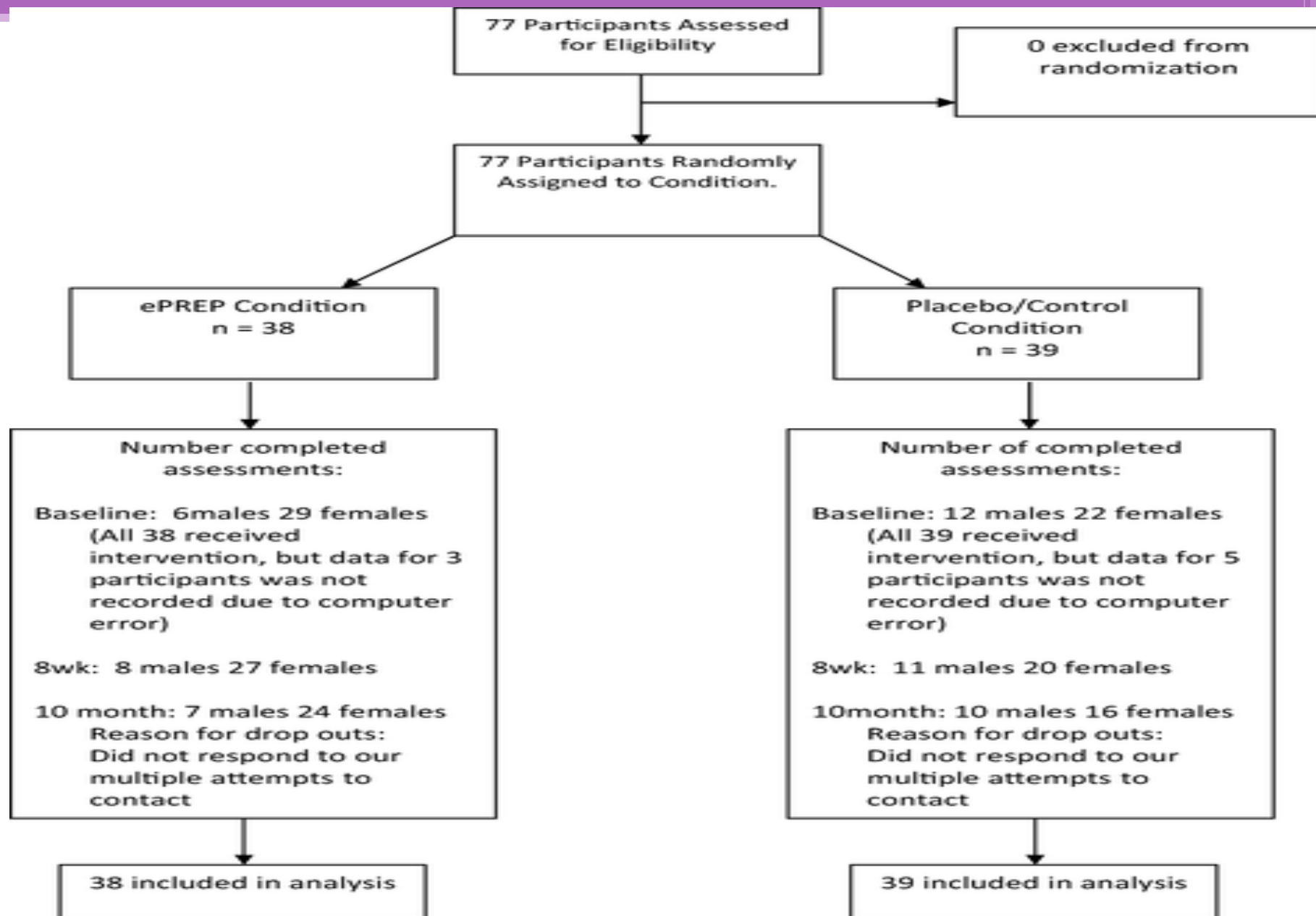
400 Participants
5g/day

400 Participants
15 g/day

400 Participants
Placebo

Primary Endpoint:

Change in frequency of moderate to severe hot flashes at 12 weeks



© Original Artist
Reproduction rights obtainable from
www.CartoonStock.com



Search ID: dc00185

"The police called, we're taking you out of the clinical trial and putting you in a criminal trial."

WHAT DO WE LOOK FOR?

- VALIDITY
- IMPORTANCE
- APPLICATION

VALIDITY

A series of horizontal lines of varying lengths and shades of purple and white, extending from the left edge of the slide towards the right, positioned below the word 'VALIDITY'.

Are the results of this single preventive or therapeutic trial valid?

Was the assignment of patients to treatments **randomised**?

Was the randomisation list concealed?

Was follow-up of patients sufficiently **long** and **complete**?

Were all patients **analysed** in the groups to which they were randomised?

Were patients and clinicians kept "**blind**" to treatment?

Were the groups **treated equally**, apart from the experimental treatment?

Were the groups **similar at the start** of the trial?

Was the assignment of patients to treatments **randomised**?

Was the randomisation list concealed?

Ensuring Allocation Concealment

BEST – most valid technique

- Central computer randomization



DOUBTFUL

- Envelopes, etc

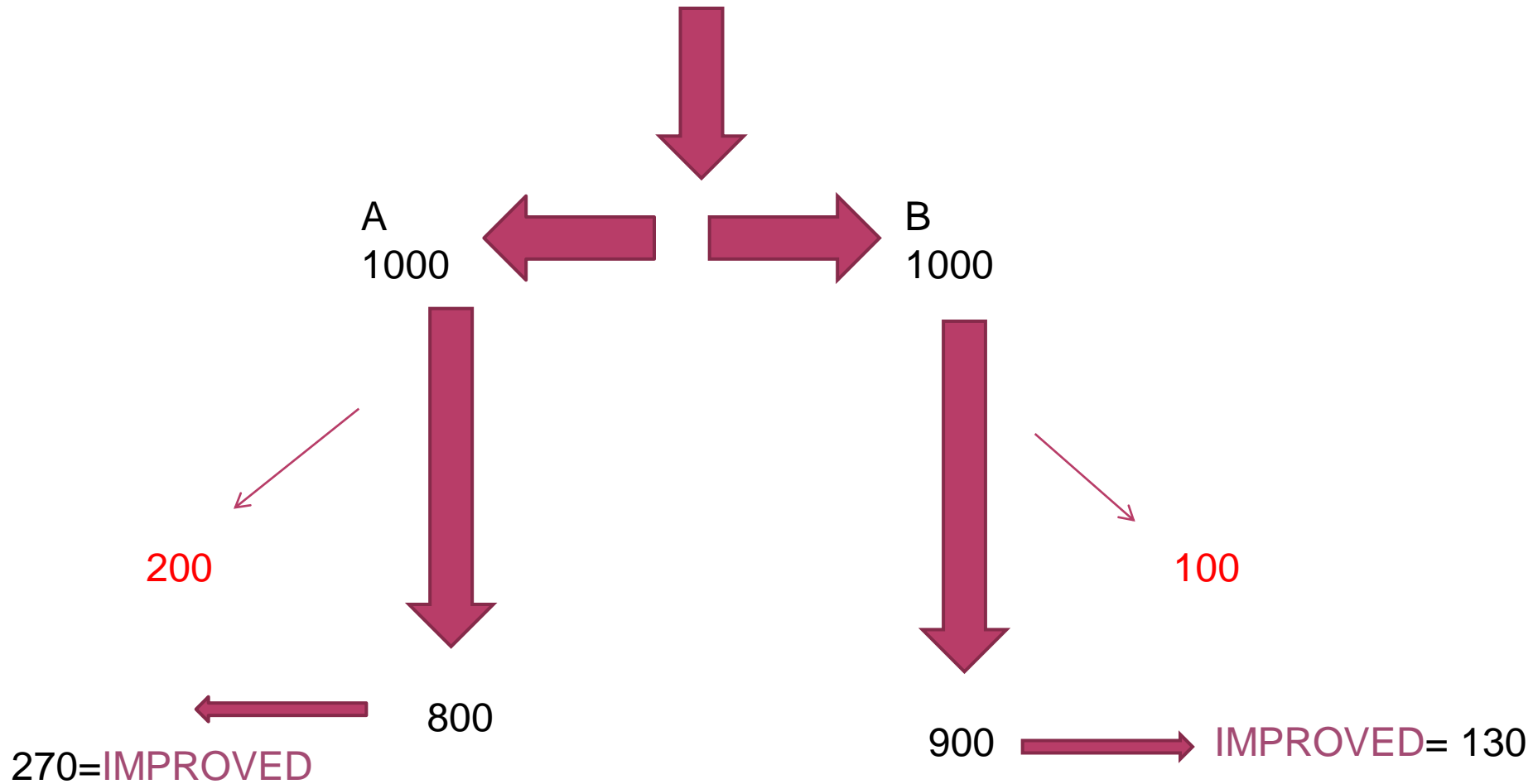


NOT RANDOMIZED

Date of birth, alternate days, etc ■

Was follow-up of patients
sufficiently long and complete?

2000 RANDOMIZED



EER= 270/?
CER=130/?

Losses-to-follow-up

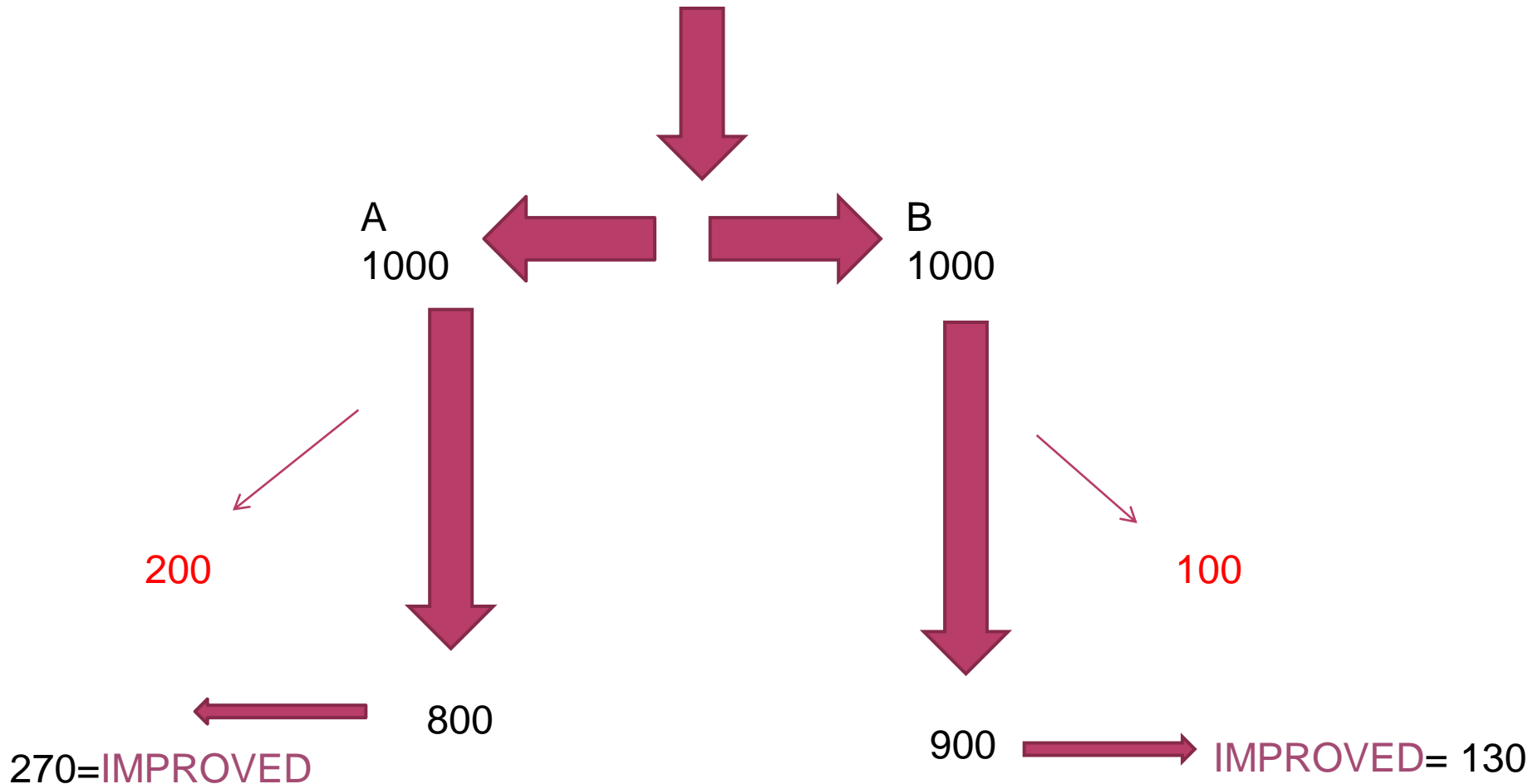
How many is too many?

“5-and-20 rule of thumb”

- 5% probably leads to little bias
- >20% poses serious threats to validity

Were all patients analysed in the groups to which they were randomised?

2000 RANDOMIZED



EER= 270/?

CER=130/?

Intention-to-Treat Principle

Maintaining the randomization

Principle:

Once a patient is randomized, s/he should be analyzed in the group randomized to - even if they discontinue, never receive treatment, or crossover.

Exception: If patient is found on BLIND reassessment to be ineligible based on pre-randomization criteria.

Were patients and clinicians kept
"blind" to treatment?

Measurement Bias - minimizing differential

- Blinding – Who?
 - Participants?
 - Investigators?
 - Outcome assessors?
 - Analysts?
- Most important to use "blinded" outcome assessors when outcome is **not objective!**
- Papers should report **WHO** was blinded and **HOW** it was done



Figure 1: The authors: double blinded versus single blinded



Figure 2: The authors blinded and masked

Schulz and Grimes. Lancet,
2002

Best RCTs: Double Blind

- Subject doesn't know which he's getting.
- Researcher doesn't know which he's giving.



- Exit poll to see if patients could guess if they were in the placebo group

Active tablet



Placebo capsule



Active capsule



Placebo tablet



Were the groups **treated equally**,
apart from the experimental
treatment?

Were the groups similar at the
start of the trial?

IMPORTANCE

MEASURES OF ASSOCIATION

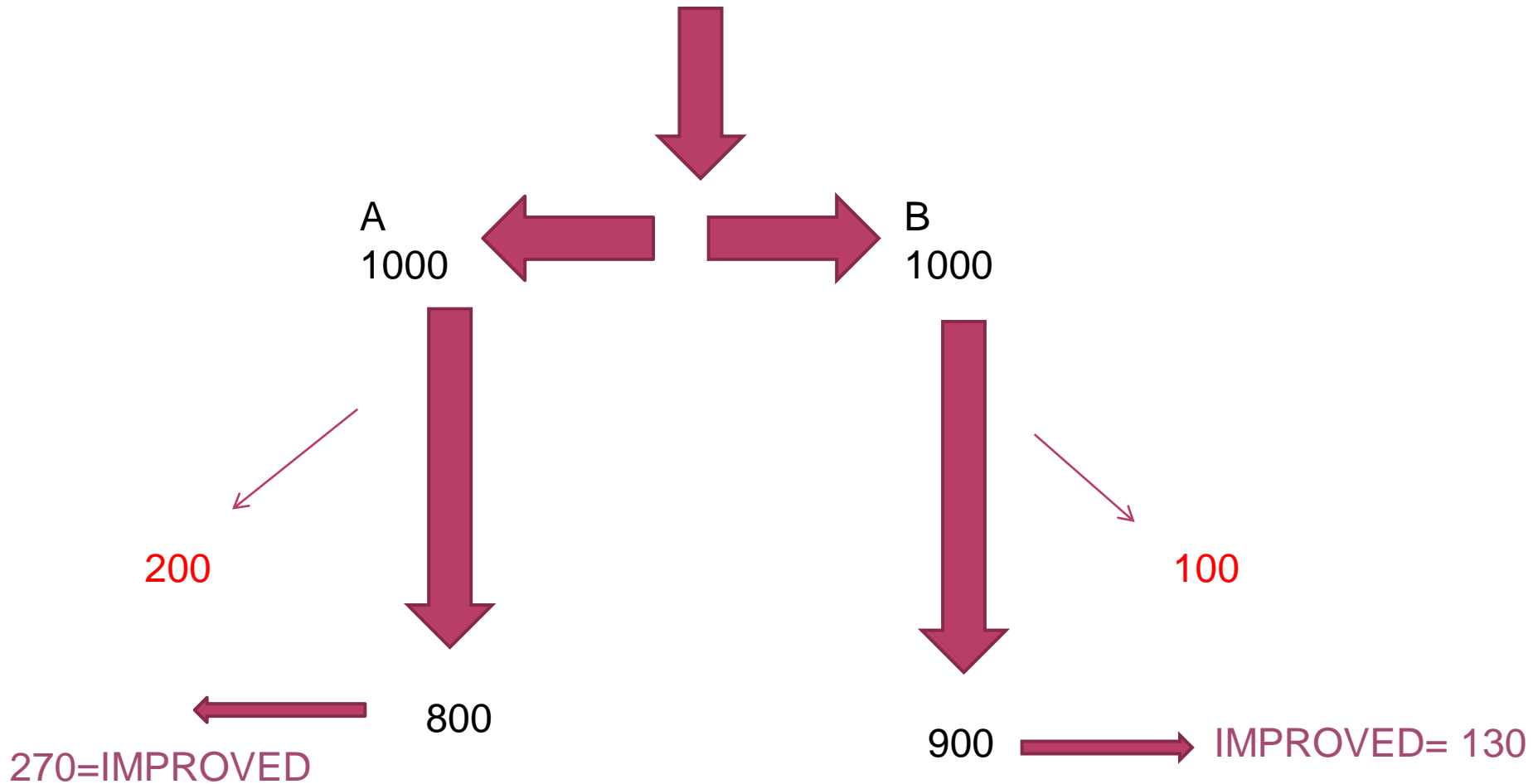
A series of horizontal lines of varying lengths and colors (purple, white, and light purple) extending from the left edge of the slide towards the right, positioned below the title.

Definition

- **Number Needed to Treat (NNT):**
 - Number of persons who would have to receive an intervention for 1 to benefit.

$$\text{NNT} = 1/\text{ARR}$$

2000 RANDOMIZED



EER= 270/?
CER=130/?

$$\text{EER}=270/800 = 33\% = 0.33$$

$$\text{CER}= 130/900=14 \, \%=0.14$$

$$\text{ARR}= 0.33-0.14= 0.19$$

$$\text{NNT}=1/0.19= 5.2=6$$

$$\text{EER}=270/1000=27\%=0.27$$

$$\text{CEER}=130/1000= 13\%=0.13$$

$$\text{ARR}=0.27-0.13=0.14$$

$$\text{NNT}=1/0.14=7$$



NUMBER NEED TO HARM(NNH)

WHEN THE OUTCOME IS UNFAVOURABLE

Therapy

- ARR
- RR
- RRR
- NNT
- CI

Result

- **Result Experimental Event Rate (EER)**
Risk (or chance) of outcome event in experimental group
- **Results control event rate (CER)**
Risk (or chance) of outcome event in control group.

Result

Relative Risk (RR)

- A measure of the chance of the event occurring in the experimental group relative to it occurring in the control group.
- **$RR = EER / CER$**

Relative Risk Reduction (RRR):

- **$RRR = CER - EER / CER$**
- **A RRR of 25% means that the new treatment reduced the risk of death by 25% relative to that occurring among control patients; the greater the relative risk reduction, the more effective the therapy.**

Absolute Risk Reduction (ARR)

- The absolute difference between the risk of the event in the control and experimental groups.
- $ARR = CER - EER$

Number needed to treat (NNT)

- Measure of clinical significance
- How many pat's have to be treated with intervention in order one patient Would expected to benefit.
- **$NNT=1/ARR$**
- Conversely, can do number needed to harm
 - uses harmful outcomes, eg death, weight gain

Magnitude (treatment effect):

- Absolute effects (ARR & NNT)
- Relative effects (RR, RRR)

Precision:


- **P value.**
- Confidence interval?

Result Tabulation

	Event + Ve	Event - Ve	Total
Experimental	a	b	a+b
Control	c	d	c+d

EER = Experimental Event Rate ($a/a+b$) •
CER = Control Event Rate ($c/c+d$) •

Result Tabulation



	Bleeding present	Bleeding Absent	Total
Drug A	20	80	100
Drug B	40	60	100

EER-A (Risk A) = $20/100 = 20\%$ (0.2) •

CER-B (Risk B) = $40/100 = 40\%$ (0.4) •

Calculations

- **$ARR = CER - EER$**
- **$NNT = 1 / ARR$**
- **$RR = EER / CER$ (Risk A/Risk B)**
- **$RRR = 1 - RR$**

Result Tabulation

	Bleeding present	Bleeding Absent	Total
Drug A	20	80	100
Drug B	40	60	100

$$ARR = CER - EER$$

$$NNT = 1 / ARR \bullet$$

$$RR = EER / CER$$

$$RRR = 1 - RR \bullet$$

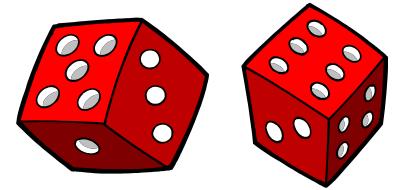
Calculations

- ▶ $ARR = CER - EER = 0.4 - 0.2 = 0.2 \text{ (20\%)}$
- ▶ $NNT = 1 / ARR = 1/0.2 = 5$
- ▶ $RR = EER/CER = 0.2/0.4 = 0.5$
- ▶ $RRR = 1 - RR = 1 - 0.5 = 0.5 \text{ (50\%)}$

Confidence intervals?

- The range within which the likelihood of a true value is expected to be within a given degree of certainty, usually evaluated at 95% CI.
- Precision

Confidence Intervals (Estimation) - in DVT study



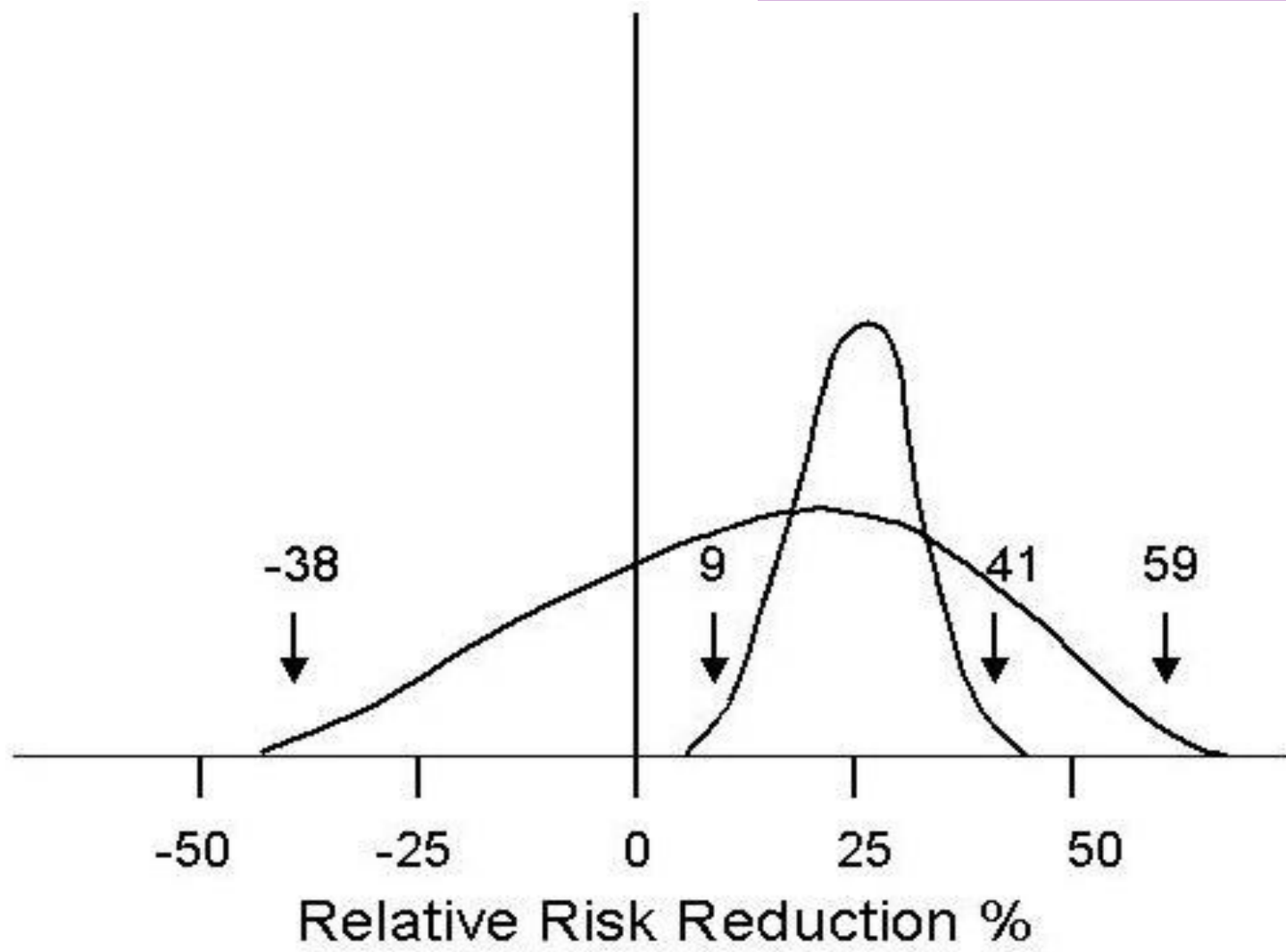
- Incidence of DVT
 - Stocking group - 0
 - No Stocking group - 0.12

Risk difference = $0.12 - 0 = 0.12$

(95% CI, 0.058 - 0.20)

The true value could be as low as 0.058 or as high as 0.20 - *but is probably closer to 0.12*

Since the CI does not include the 'no effect' value of '0' → the result is statistically significant



APPLICABILITY

A series of horizontal lines in various shades of purple and white, extending across the width of the slide below the title.

CAN I APPLY THESE VALID, IMPORTANT RESULTS TO MY PATIENT?

- **Do these results apply to my patient?**
 - **IS OUR PATIENT SO DIFFERENT?**
 - **IS THE TREATMENT FEASIBLE?**
 - **POTENTIAL BENEFITS AND HARMS**
- **Are my patient's values and preferences satisfied by the intervention offered?**

Summary

- Validity – is the paper likely to be true
- Importance – size of effect
 - NNT
 - Percision
- Applicability – can it work for me/my setting



Evidence Based Medicine Toolkit

Home

About EBM ▶

Domains ▶

Practice Guidelines ▶

Systematic Review ▶

Economic Analysis ▶

Glossaries ▶

JAMA User's Guide

Links ▶

[Home](#)>Domains-Therapy/Prevention-Appraisal Guide

Therapy/Prevention Article Appraisal Guide

Questions to Ask Yourself

Are the results valid?

1. Was the assignment of patients to treatment [randomized](#)?

2. Were all patients who entered the trial properly accounted for and attributed at its conclusion?

➤ Was [follow-up](#) complete?

➤ Were patients analyzed in the groups to which they were randomized ?

➤ [Intention to treat analysis](#)?

Toolkit

Home

About EBM ▶

Domains ▶

Practice Guidelines ▶

Systematic Review ▶

Economic Analysis ▶

Glossaries ▶

JAMA User's Guide

Links ▶



"This is a collection of tools for identifying, assessing and applying relevant evidence for better health care decision-making. The appraisal tools are adapted from the Users' Guides series prepared by the Evidence Based Medicine Working Group and originally published in JAMA"

Comments or suggestions for improvement?

Contact:

jeanette.buckingham@ualberta.ca

bruce.fisher@ualberta.ca

duncan.saunders@ualberta.ca



Thank You!

