CLINICAL STUDY DESIGNS:

An Overview

Dr Tejpal Gupta, MD, DNB

Officer-in-Charge, ECTU-CRS

Associate Professor, Radiation Oncology

ACTREC, Tata Memorial Centre, Mumbai

ICRO Teaching Course, 7-8th Sept' 13, Patna

Overview of the presentation

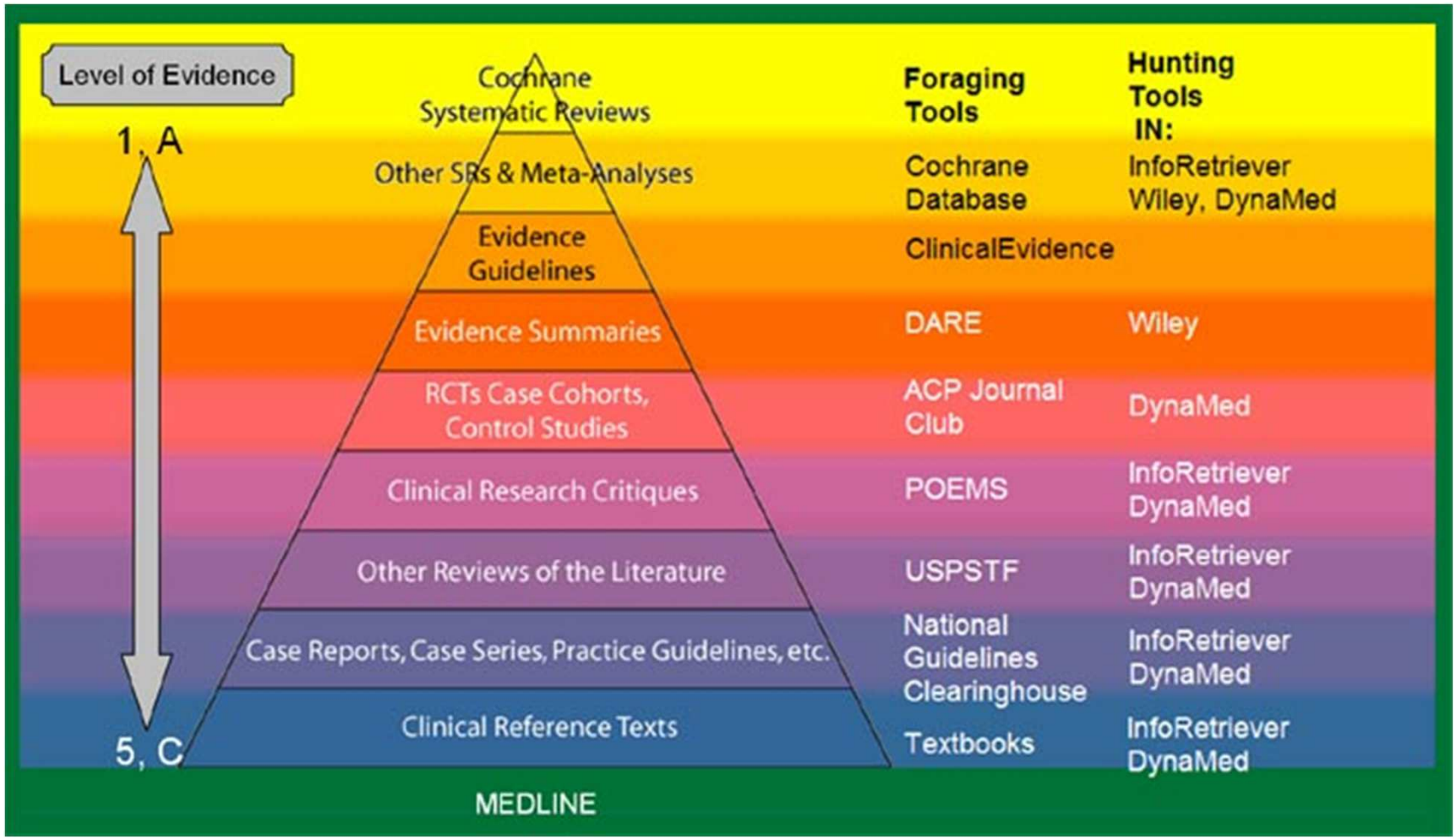
I. Introduction to clinical research

II. Types of clinical studies

III. Observational study designs

IV. Experimental study designs

The Evidence Pyramid



Important issues in Study Design

- **Validity: *Truth***
  - External Validity:
    - *The study can be generalized to the population*
  - Internal Validity:
    - *Results are not due to chance, bias or confounders*
  - Symmetry Principle:
    - *Groups are similar*

Important issues in Study Design

- Confounding: *distortion of the effect of one risk factor by the presence of another*
- Bias: *Any effect from design, execution, & interpretation that shifts or influences results*
  - **Confounding bias:** *failure to account for the effect of one or more variables that are not distributed equally*
  - **Measurement bias:** *measurement methods differ between groups*
  - **Sampling (selection) bias:** *design & execution errors in sampling*

Classification of Research Study Designs

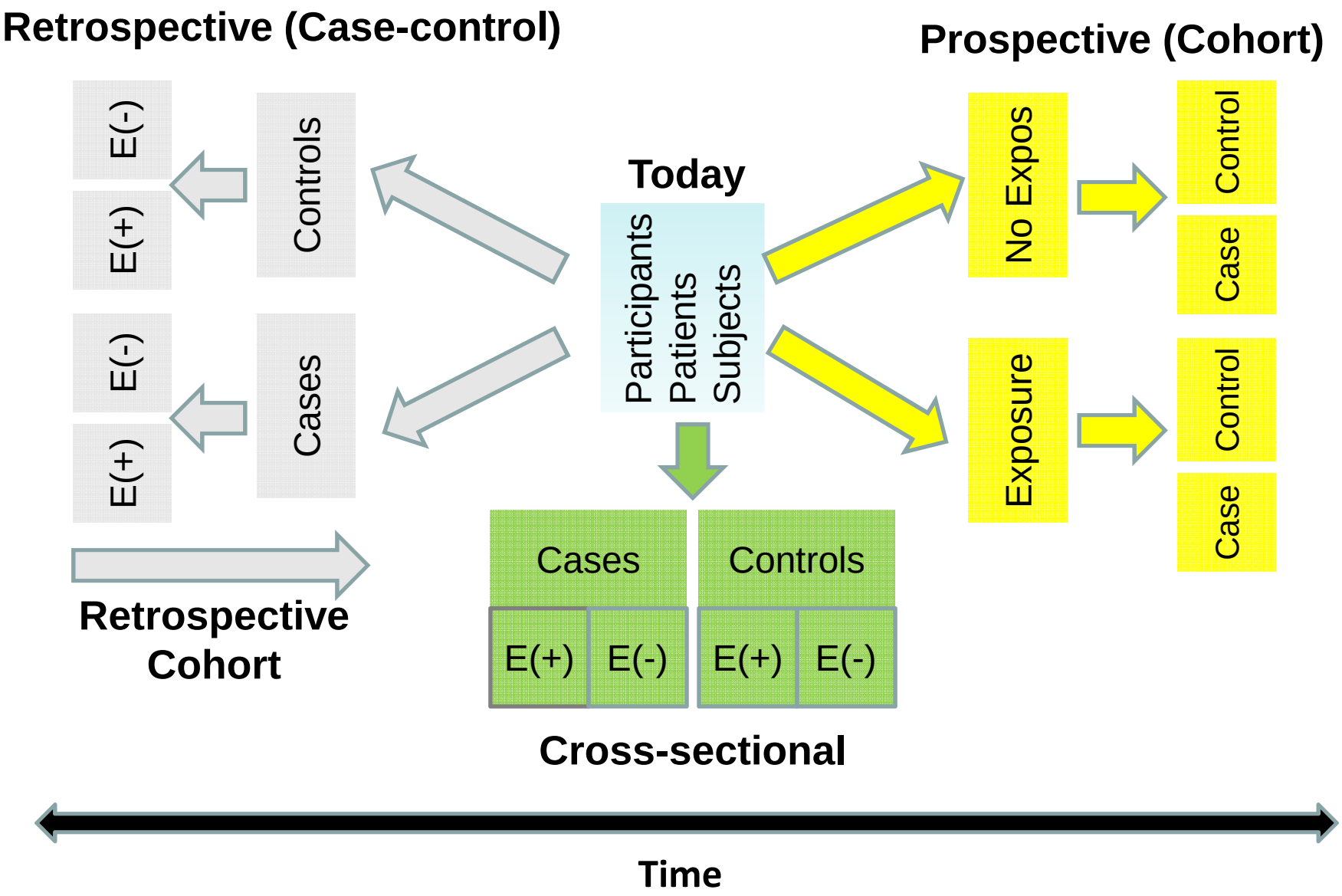
Observational studies

1. Case report/case-series
2. Case-control studies
3. Cross-sectional studies
4. Prospective (cohort)
5. Retrospective cohort

Interventional studies

1. Controlled trials
   a) Parallel designs
   b) Sequential designs
   c) External controls
2. Studies with no controls

Observational Designs



Aims of observational studies

- Evaluate the effect of a suspected risk factor (exposure) on an outcome (e.g. disease)

  → *define 'exposure' and 'disease'*

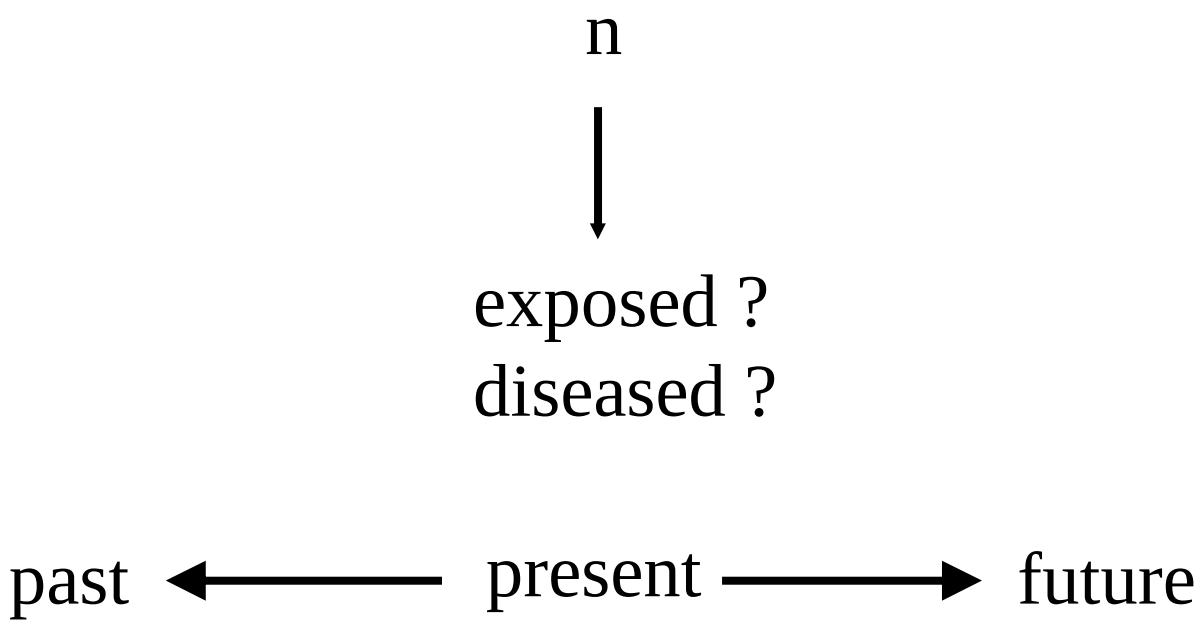
- Describe the impact of the risk factor on the frequency of disease in a population

Characteristics of observational studies

- No control over study units
- Can study risk factors that have serious consequences
- Study individuals in their natural environment
- Possibility of confounding

Cross-sectional study

- Exposure and disease measured once, i.e. at the same point in time



Cross-sectional study

- Random sample from population
  - *i.e. results reflect reference population*
- Estimates the frequencies of both exposure and outcome in the population
- Measures both exposure & outcome at one point in time
- Typical example is a survey

Cross-sectional study

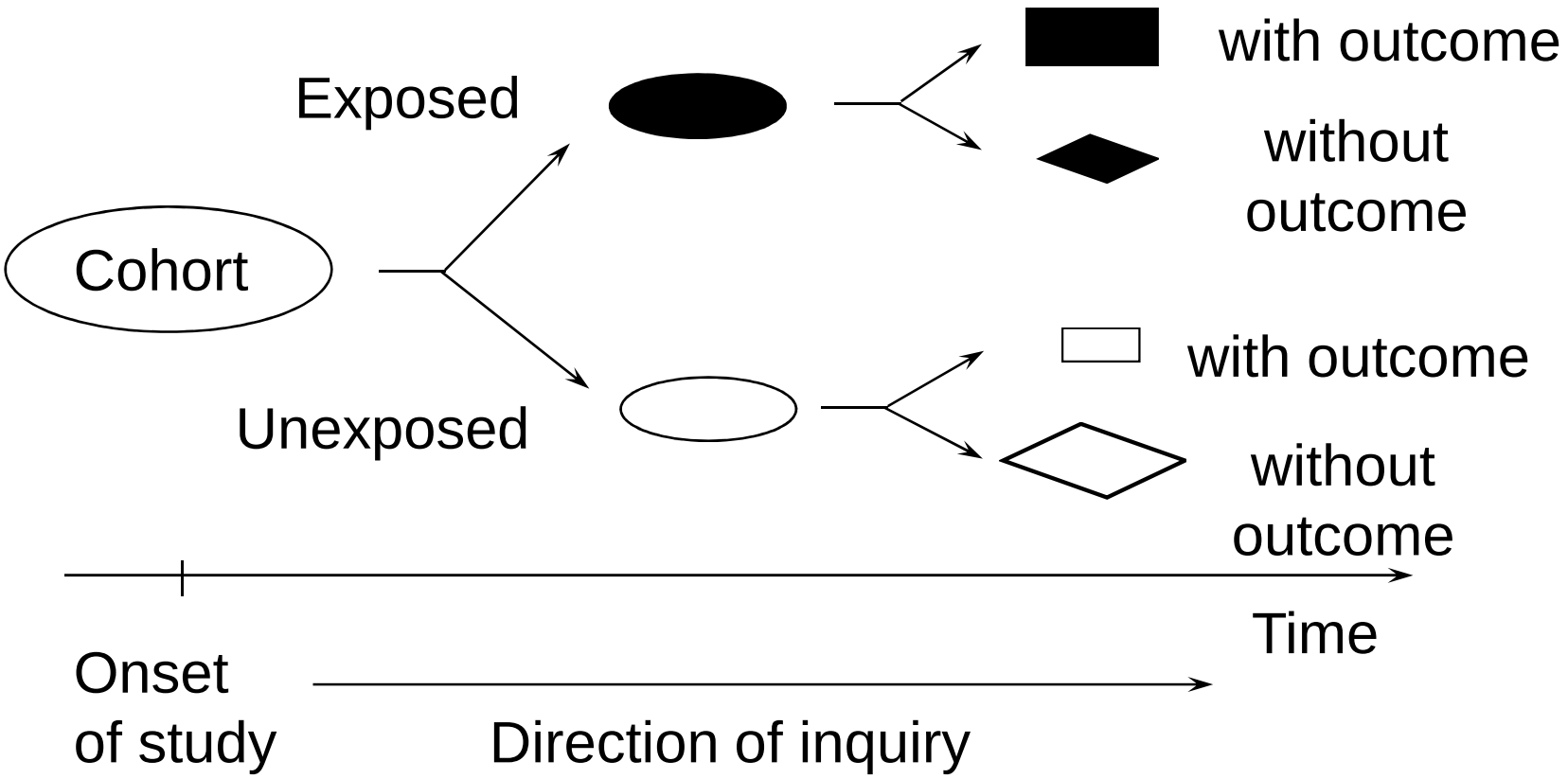
- Can study several exposure factors and outcomes simultaneously
- Determines disease prevalence
- Helpful in public health administration & planning
- Quick and low cost (e.g. mail survey)
- Limitations:
  - *Does not determine causal relationship*
  - *Not appropriate if either exposure or outcome is rare*

Cohort studies

- Follow-up studies: subjects selected on presence or absence of exposure & absence of disease at one point in time. Disease is then assessed for ***all*** subjects at another point in time

- Typically prospective but can be retrospective, depending on temporal relationship between study initiation & occurrence of disease

Prospective Cohort Study



Q: What will happen?

Cohort studies

- More clearly established temporal sequence between exposure & disease
- Allows direct measurement of incidence
- Examines multiple effects of a single exposure (nurses' health study, oral contraceptives & breast, ovarian cancers)

Cohort studies

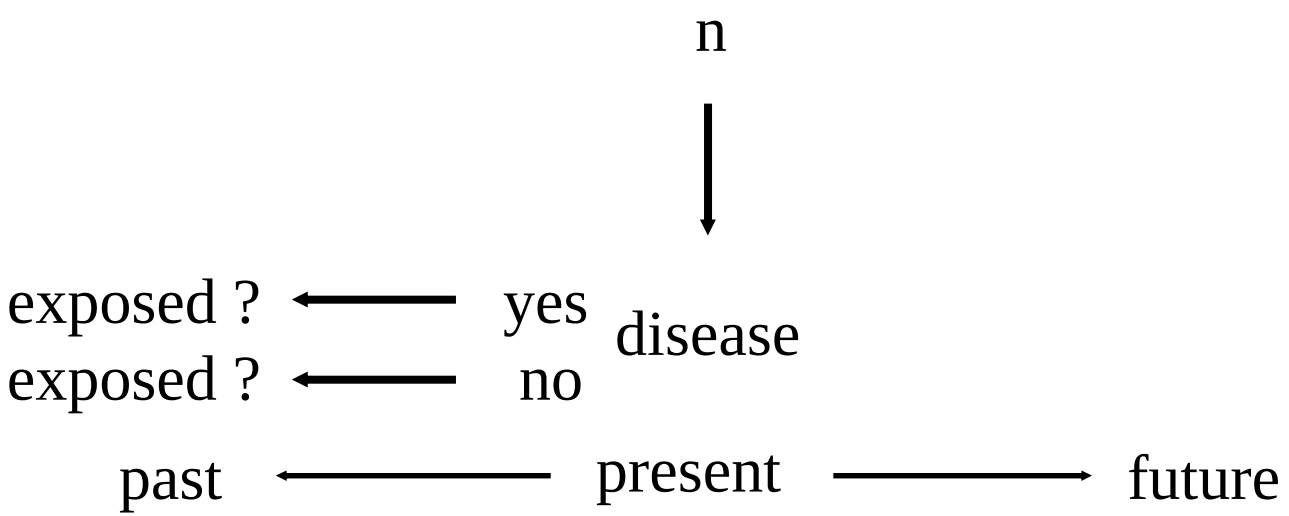
- *Limitations:*
  - time consuming and expensive
  - loss to follow-up & unavailability of data
  - potential confounding factors
  - inefficient for rare diseases

Case-Control Study

- Retrospective
  - *Can use hospital or health register data*
- First identify cases
- Then identify suitable controls
  - *Hardest part: who is suitable ??*
- Then inquire or retrieve previous exposure
  - By interview
  - By databases (e.g. hospital, health insurance)

Case-Control Study

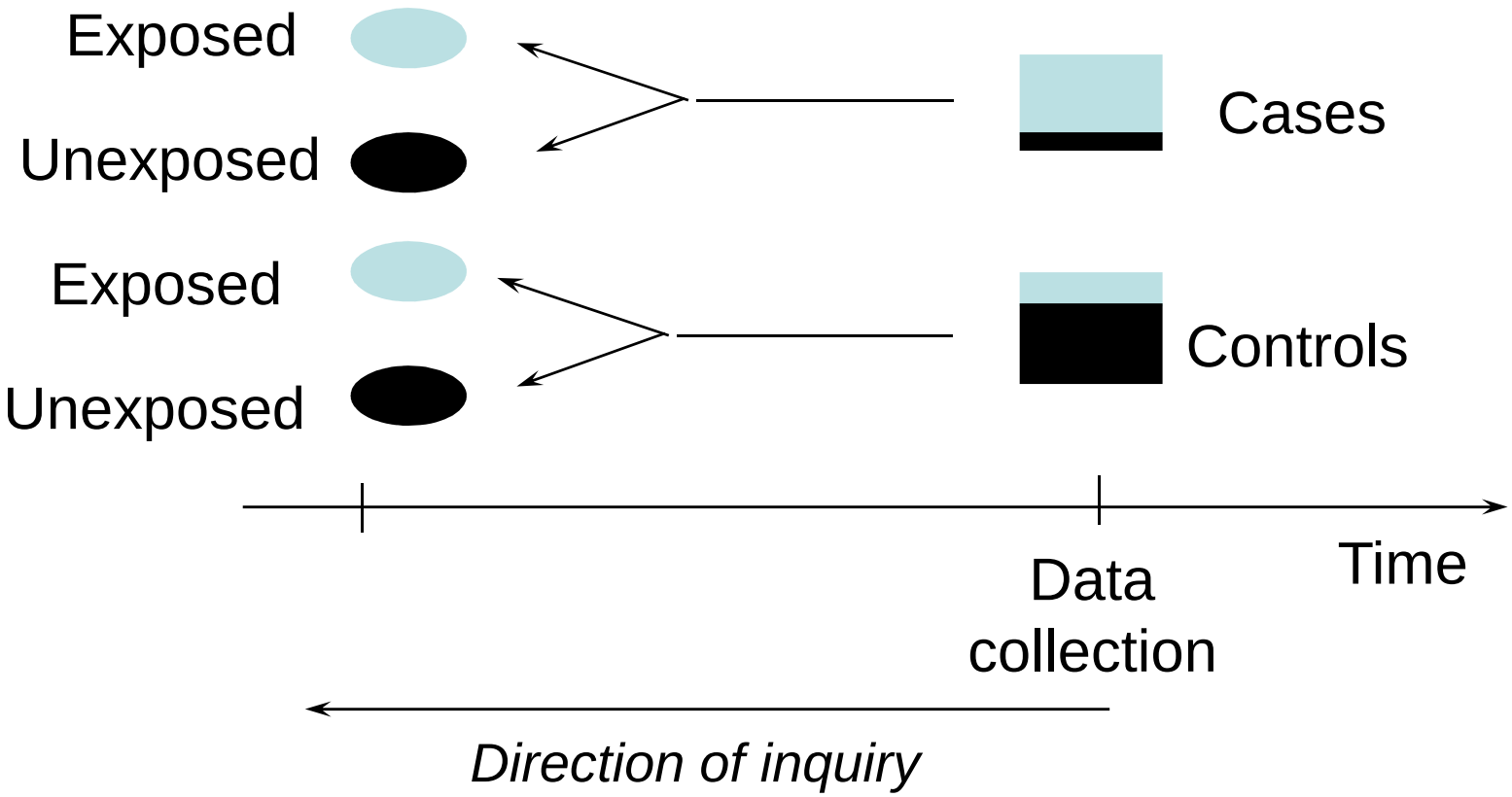
- Diseased and non-diseased individuals are selected first
- Then past exposure status is retrieved



Case-Control Study

- Good for rare disease (e.g. cancer)
- Can study many risk factors at the same time
- Usually low cost (though not always)
- Confounding likely
- Measures Odds Ratio (not Relative Risk !!)

Case-Control Study Design



Q: What happened?

Case Selection

- Define source population
- Cases
  - incident/prevalent
  - diagnostic criteria (sensitivity + specificity)
- Controls
  - selected from same population as cases
  - select independent of exposure status

Control Selection

- Random selection from source population
- Hospital based controls:
  - convenient selection
  - controls from variety of diagnostic groups other than case diagnosis
  - avoid selection of diagnoses related to particular risk factors
  - limit number of diagnoses in individuals

Summary of Observational Studies

Characteristic	Cross-sectional	Case-Control	Cohort
Sampling	Random sample: population	Purposive sample: diseased/non-diseased	Purposive sample: Exposed/non-exposed
Time	One point	Retrospective	Prospective
Causality	Statistical association	Screening for many risk factors	Testing one (or few) risk factors
Frequency measure	Prevalence	None	Incidence
Risk parameter	Prevalence (risk) ratio, odds ratio	Odds ratio	Relative risk, odds ratio

What is a clinical trial ?

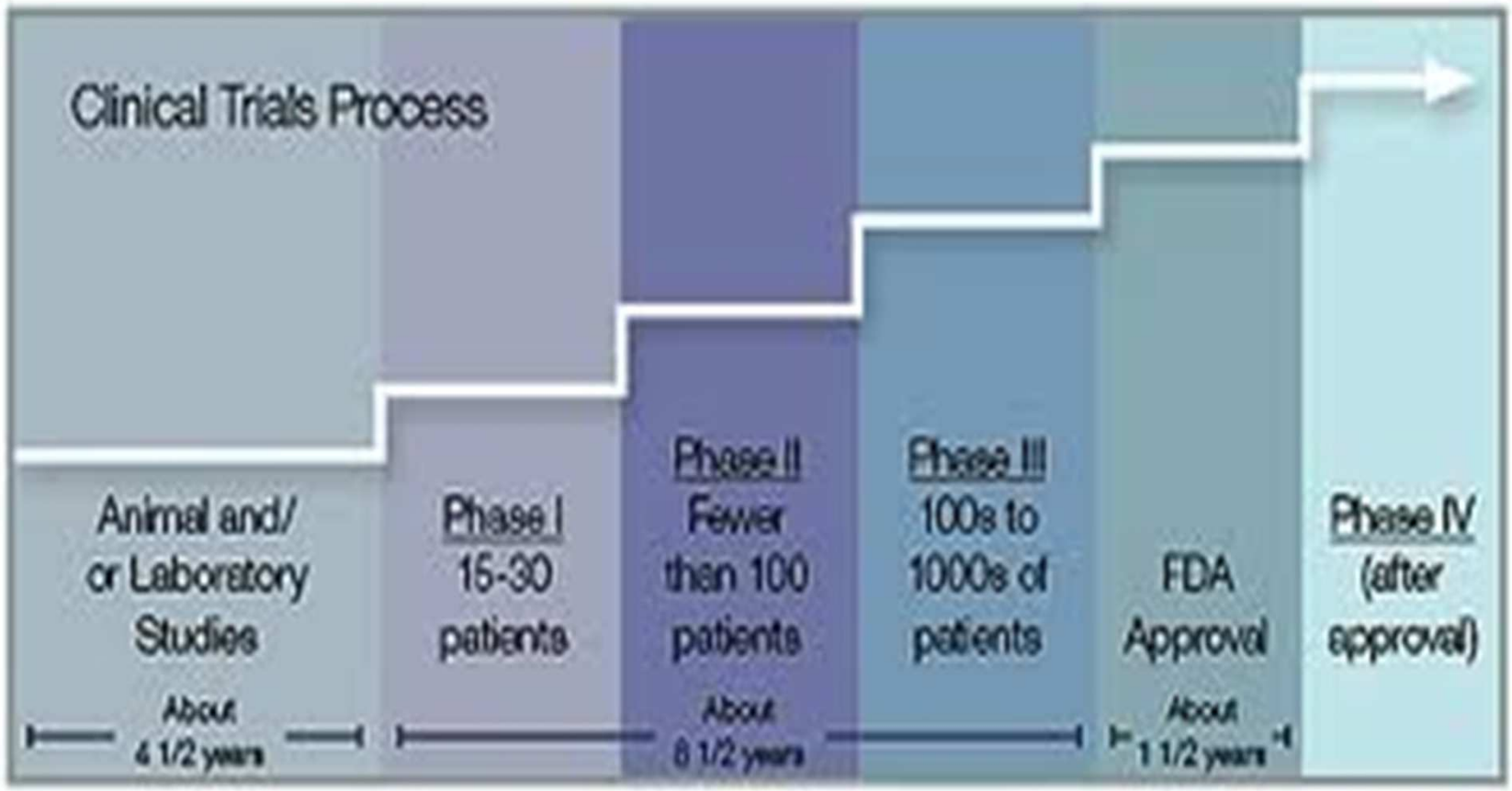
A clinical trial tests potential interventions in humans to determine if the intervention represents an advance and should be adopted for general use

FDA 2003

What do clinical trials test ?

- Clinical trials test research hypotheses
- Good clinical trials test specific research hypothesis
- A clinical research hypothesis is a carefully formulated assumption developed in order to test its logical consequences

Phases of Clinical Trials



What is a Phase I trial ?

- First evaluation of a new therapy in humans
- Classical Goals:

✓ Identify dose limiting toxicities (DLT)

✓ Identify maximum tolerated dose (MTD)

✓ Assess pharmacokinetics (PK)

✓ Assess pharmacodynamic (PD) endpoints

Dose Limiting Toxicity (DLT)

- Toxicity described by standardized grading criteria considered unacceptable because of severity or irreversibility
- What is considered unacceptable varies from situation to situation
- DLT is specified for each trial protocol

Classical examples of DLT

For intermittent therapy (x q 3-6 weeks)

- ANC < 500/µL for > 5 days
- ANC < 500/µL of any duration with fever >38.5 C
- Platelets <25,000
- Grade 3-5 non-hematologic toxicities
- Inability to retreat within 2 weeks of scheduled treatment

DLTs continued

For chronic therapy (daily or alternate days), the threshold for DLTs is lower

- Grade 3 or worse hematologic toxicities
- Grade 3 or worse hepatotoxicity (incl Enz)
- Grade 2 or worse non-hematologic toxicities
- 2-grade increase from baseline that is persistent despite supportive measures

Maximum Tolerated Dose (MTD)

- Dose associated with unacceptable toxicity (DLT) in a pre-specified proportion of patients (sometimes defined inconsistently)

  Generally the dose at which ≥ 33% patients experience unacceptable toxicity (eg. DLT in ≥ 2 out of 6 pts)

- MTD: 1 dose level below DLT

Principles of Phase I trial design

- Start with a safe starting dose
- Minimize number of patients treated at sub-therapeutic doses
- Escalate doses rapidly in absence of toxicity
- Escalate dose slowly in presence of toxicity

Selection of starting dose

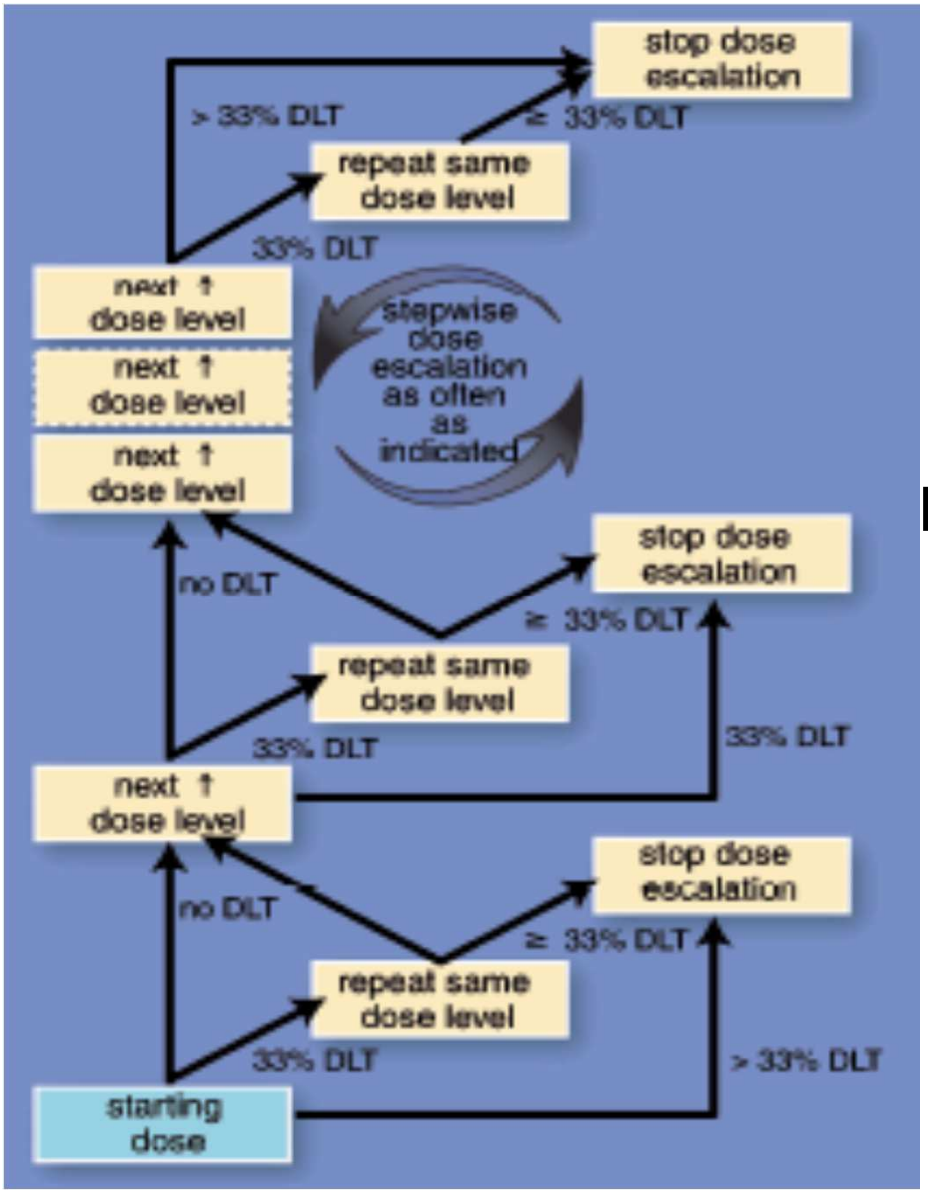
- Based on pre-clinical (animal) toxicity data

✓ 1/10th LD10 (0.1 MELD10) if mouse most sensitive species

✓ 1/ 3rd TDL (0.33 DETDL) if dog most sensitive species

- Conservative dosing based on other studies if data suggest human tissues more sensitive and PK data show a steep dose/toxicity curve

Classical Phase I Trial Design



Conventional 3 + 3 design

3-6 pts per cohort

Starting dose 0.1 MELD

Dose escalation using modified Fibonacci series

100%-66%-50%-40%-33%

MTD achieved

Previous dose level is RPTD

Problems with 3 + 3 design

- Empirical and traditional method
- Lacks solid statistical foundation
- May take too many or too few patients
- Takes a long time for dose escalation
- Starting dose may significantly under or overestimate ‘safe’ dose
- Early dose levels far away from optimal and/or therapeutic doses thus exposing substantial proportion of patients to low, suboptimal and ineffective doses

Possible Approaches

The variables in conventional 3 + 3 design

- **Starting dose**

  Should higher starting doses be used?

  If so, when and with what restrictions?

- **Number of patients per dose level**

  Is the entry of one patient per dose level ok?

  If so when, and with what restrictions?

- **Method/Rapidity of dose escalation**

Should novel dose escalation schemes be used?

Novel dose escalation designs

Pharmacologically guided dose escalation

- Rapid escalation (doubling) to target AUC (murine PK)
- Accelerated Titration designs
- **Optimal Biological Dose design for non-cytotoxics**

Statistically guided methods

- Continual Reassessment Methods
- Escalation with overdose control
- Logistic model to dose versus DLT
- Bayesian and /or regression modeling
- Markov modeling

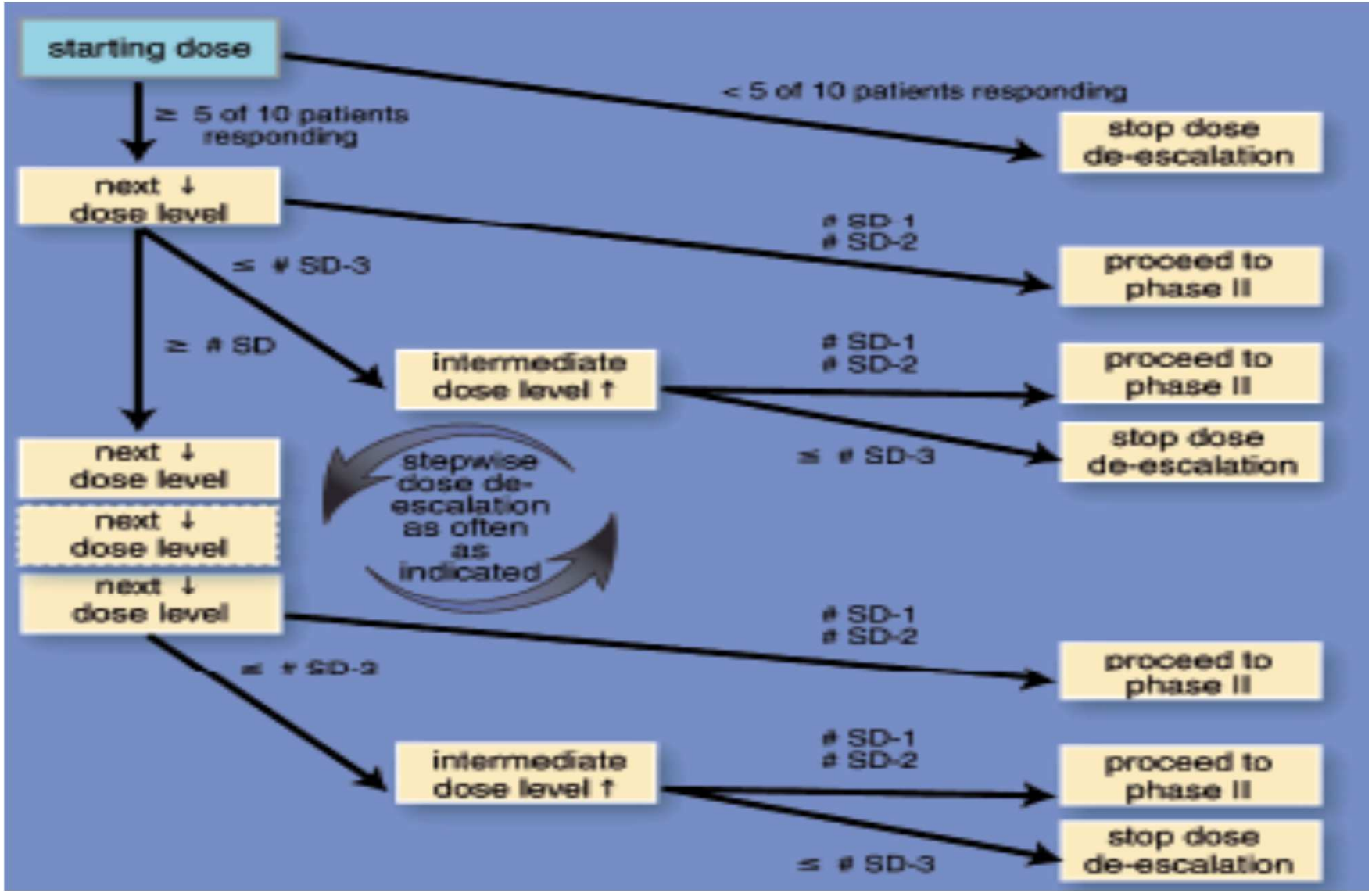
Trial design issues with non-cytotoxics

- Common AEs of conventional cytotoxics (e.g. myelosuppression) not seen
- Conventional response criteria based on tumour shrinkage may not apply
- Dose-toxicity curve may not predict a clinically efficacious dose
- DLTs or MTDs may not be reached / may not be applicable
- The most efficacious dose may be independent of the MTD

Optimal Biological Dose (OBD) or Biological Modulatory Dose (BMD) proposed as new paradigm for such agents

Uses pharmacodynamic information (hopefully relevant to drugs mechanism of action) including surrogate pharmacodynamic markers rather than the toxicity profile

Novel pharmacologically guided dose deescalation design for determining OBD



Dowlati et al CCR 2005

Summary of phase I trial designs

- Phase I clinical trials are the gateway for all new cancer drug development
- The traditional 3+3 design though safe, treats large numbers below therapeutic doses and takes a long time for escalation
- Alternative phase I designs enable a higher proportion of patients to be treated at or near MTD, but at an increased risk of Grade 3-4 toxicities within a shortened time span

What is a phase II trial ?

Primary goal is to establish ***efficacy / activity***

- ✓ “Proof-of-principle” for preliminary evidence of efficacy
- ✓ Generally relies on surrogate end-points
- ✓ Mostly single arm studies
- ✓ May be randomized (“pick-a-winner”)
- ✓ Suggests further potential (for phase III testing)
- ✓ Documents additional safety information

Objectives of phase II trials

1. To define antitumor ***activity***

2. To further demonstrate ***safety***

3. To gain new insights into the ***pharmacokinetics***, ***pharmacodynamics*** & ***metabolism*** of drugs

4. To evaluate ***biologic correlates*** which may predict response or resistance to treatment and/or toxicity

Phase II study designs

- *Frequentist*
  - Gehan 2-Stage
  - Simon 2-Stage Optimal
  - Simon 2-Stage Mini-max
  - Fleming 1-stage
  - Gehan-Simon 3-Stage
  - Randomized Phase 2
  - Constant Arc-Sine
  - Randomized Discontinuation
- *Bayesian*
  - Thall-Simon-Estey
  - 1-Stage Bayesian
  - 2-Stage Bayesian
    - Tan Machin
    - Heitjan
- *Adaptive*
- *Multiple Outcomes*

Two-stage Design

- Main objective is to minimize the number of patients treated with ineffective regimens
- A very commonly used 2 stage design is the Simon, which minimizes sample size based on a pre-specified response rate and an α and β error rates
  - Optimal: minimizes number of pts treated if the regimen is ineffective
  - Mini-max: minimizes the whole sample size
- RECIST criteria [CR+PR+SD] is generally utilized

Simon, Mini-max

- Treat ~12-18 patients at 1st stage
- Determine the “response rate”

- Less than that projected to indicate activity (p0): STOP!

- Sufficiently great to indicate activity: CONTINUE

- At the end of 2nd stage, declare intervention worthy of further evaluation if > x number of “responses” are observed (p1)

One-stage design

- When time-dependent endpoints are considered
  - *e.g., the proportion of patients free of progression at one year following initiation of treatment*

- Given the time period from initiation of treatment to the endpoint, two-stage designs are often impractical

- Early stopping rules are usually incorporated for obvious lack of efficacy or unacceptable toxicity

One-stage Design

- Fleming is a commonly utilized one-stage design
- You need to have a good grip on historical control data
- Mick Design: Compare time to treatment failure or progression on the new regimen [TTP2] with the individual patient's failure time or TTP1 observed with their prior regimen
  - *If the TTP2/TTP1 ratio is greater than 1.33, the regimen is considered effective and worthy of further study*

Randomized Phase II Trials

- Patients are randomized to receive one of two (or more) regimens differing by dose level, schedule, or agent
- It is not powered to make inferential comparisons between the treatment arms
- Pick the winner approach
- If both arms incorporate two-stage designs, you would have four specific decision points for determining efficacy

Randomized Discontinuation

- It incorporates time-dependent endpoints with disease response
- Stable disease patients are randomized to continuation with the agent or placebo (the discontinuation)
- Patients subsequently showing progression on placebo are then retreated with the agent to determine if stability can be regained
- This design is most appropriate in diseases where tumor growth rates are slow

Summary of Phase II studies

- Phase II trials are exploratory studies and rarely are definitive
- Efficient to exclude inactive therapies
- Results must be interpreted cautiously, in the context of the availability of other therapies
- Estimate clinical activity and provide further safety information – important in the “go/no go” decision
- Require confirmation in pivotal phase III trials

What is a phase III trial ?

Comparative Trials with or without controls

Primary goal is to establish ***actual clinical value***

- ✓ Survival (OAS, EFS,PFS) are primary endpoints
- ✓ Compares new treatment to current standard of care
- ✓ Randomized (with allocation concealment) to minimize bias
- ✓ May be sometimes placebo controlled and even blinded

Phase III Study Designs

- Comparative Studies
- Experimental Group vs Control Group
- Establishing a Control
  1. Historical
  2. Concurrent
  3. Randomized
- Randomized Control Trial (RCT) is the gold standard
  - *Eliminates several sources of bias*

Why Control Group

- To allow discrimination of patient outcomes caused by experimental intervention from those caused by other factors
  - *Natural progression of disease*
  - *Observer/patient expectations*
  - *Other treatment*
- Fair comparisons
  - *Necessary to be informative*

Type of Controls

- External
  - Historical
  - Concurrent, not randomized
- Internal and concurrent
  - No treatment
  - Placebo
  - Dose-response
  - Active (Positive) control
- Multiple
  - Both an Active and Placebo
  - Multiple doses of test drug and of an active control

Use of Placebo Control

- The “placebo effect” is well documented
- Could be
  - No treatment + placebo
  - Standard care + placebo
- Matched placebos are necessary so patients and investigators cannot decode the treatment
- E.g. Vitamin C trial for common cold
  - Placebo was used, but was distinguishable
  - Many on placebo dropped out of study
  - Those who knew they were on vitamin C reported fewer cold symptoms and duration than those on vitamin who didn't know

Historical Control Study

- A new treatment used in a series of subjects
- Outcome compared with previous ‘historical’ series of comparable subjects
- Non-randomized, non-concurrent
- Rapid, inexpensive, good for initial testing of new therapy
- Two sources of historical control data
  - *Literature*
  - *Data base*

Historical Control Study

- Vulnerable to bias

- Publication bias in literature-based controls

- Changes in outcome over time

- May come from change in:

  - *underlying patient populations*
  - *criteria for selecting patients*
  - *patient care and management peripheral to treatment*
  - *diagnostic or evaluating criteria*
  - *quality of data available*

Historical Control Study

- Tend to exaggerate the value of a new treatment
- Literature controls particularly poor
- Even historical controls from a previous trial in the same institution or organization may still be problematic
- Adjustment for patient selection may be made, but all other biases will remain

Concurrent Controls

- Not randomized controls
- Patients are compared concurrently, treated by different strategies during same the period
- Advantage
  - *Eliminates time trend*
  - *Data of comparable quality*
- Disadvantage
  - Selection Bias
  - Treatment groups not comparable
- Covariance analysis not adequate

Randomized Control Trial

- Patients assigned at random to either standard arm (control arm) or treatment arm (experimental) arm
- Equal chance of getting randomized to either arm in 1:1 randomization
- Neither patient nor physician can influence this chance
- Eliminates several known & unknown biases
- Considered to be ***"Gold Standard"***

Comparing Treatments

- **Fundamental principle**
  - Groups must be alike in all important aspects and only differ in the treatment each group receives
  - In practical terms, “comparable treatment groups” means “alike on the average”

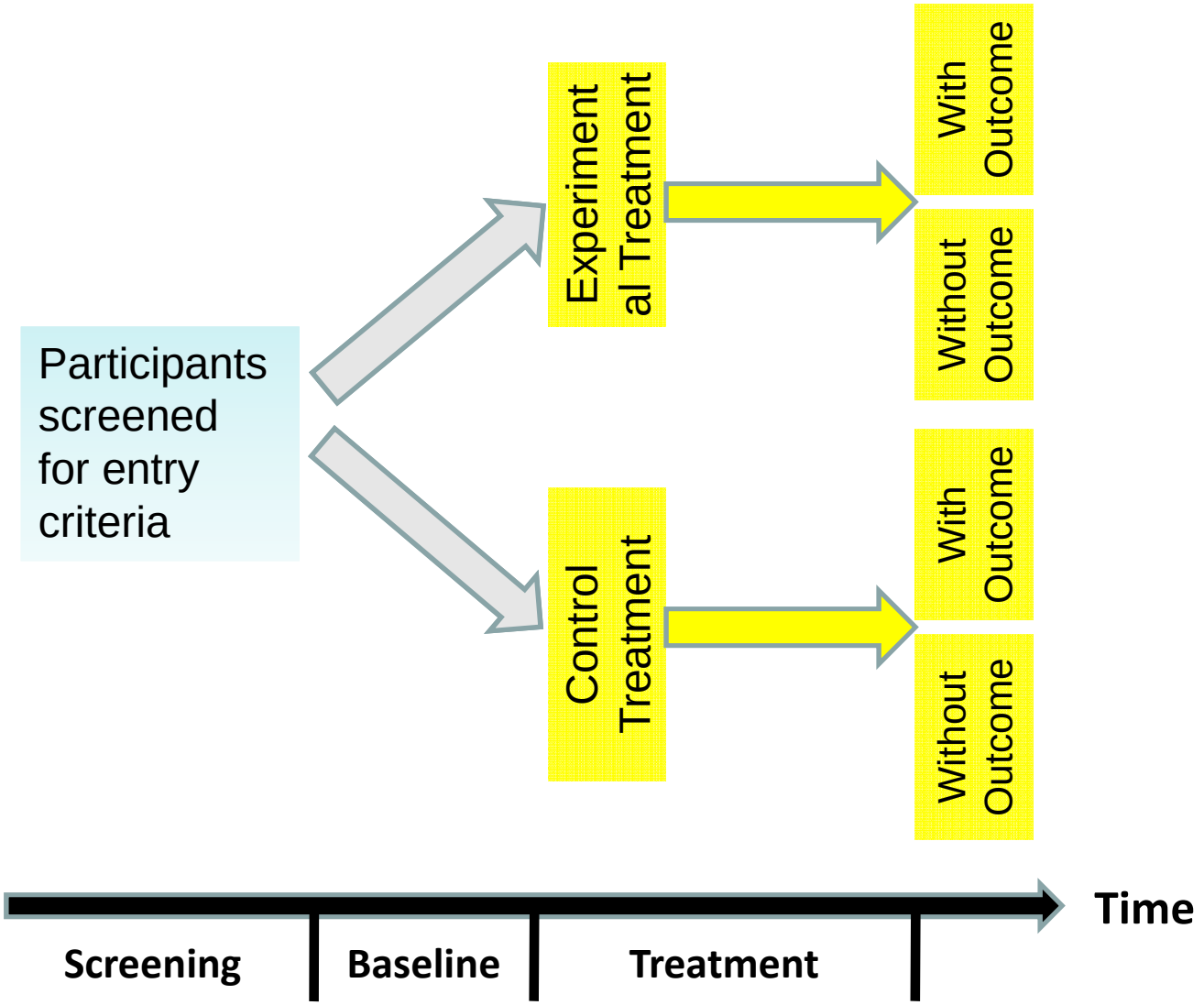
- **Randomization**
  - Each patient has the same chance of receiving any of the treatments under study
  - Allocation of treatments to participants is carried out using a chance mechanism so that neither the patient nor the physician know in advance which therapy will be assigned

- **Blinding**
  - Avoidance of psychological influence
  - Fair evaluation of outcomes

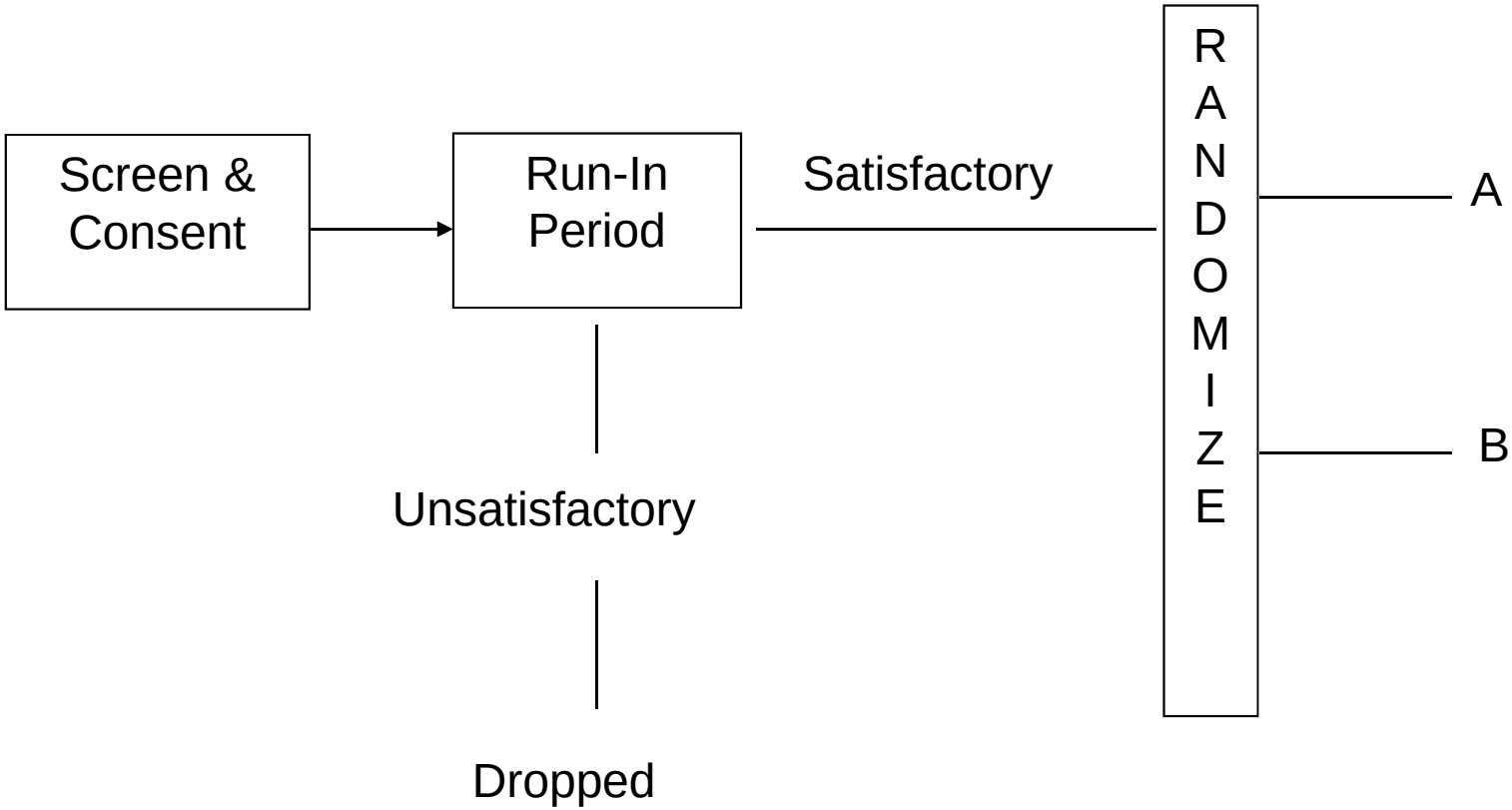
Commonly Used Phase III Designs

- Parallel Group
- Run-in designs
- Withdrawal
- Group/Cluster
- Randomized Consent
- Cross Over
- Factorial
- Large Simple
- Equivalence/Non-inferiority
- Sequential

Parallel Group Design



Run-In Design



<u>Note</u>: It is assumed that all patient entering the run-in period are eligible and have given consent

Run-In Design

Problem:

- Non-compliance by patient may seriously impair efficiency and possibly distort conclusions

Possible Solution:

- Placebo run-in for drug trials: Assign all eligible patients a placebo to be taken for a “brief” period of time. Patients who are “judged” compliant are enrolled into the study. This is often referred to as the “Placebo Run-In” period

Cluster Randomization Designs

- **Groups (clinics/communities) are the unit of randomization**
  - *TMH-NIH Visual Inspection of Acetic Acid (VIA) study*
  - *Breast self-examination programs in clinics*
  - *Smoking cessation intervention trial*

- **Advantages**
  - *Sometimes logistically more feasible*
  - *Avoid contamination*
  - *Allow mass intervention, thus "public health trial"*

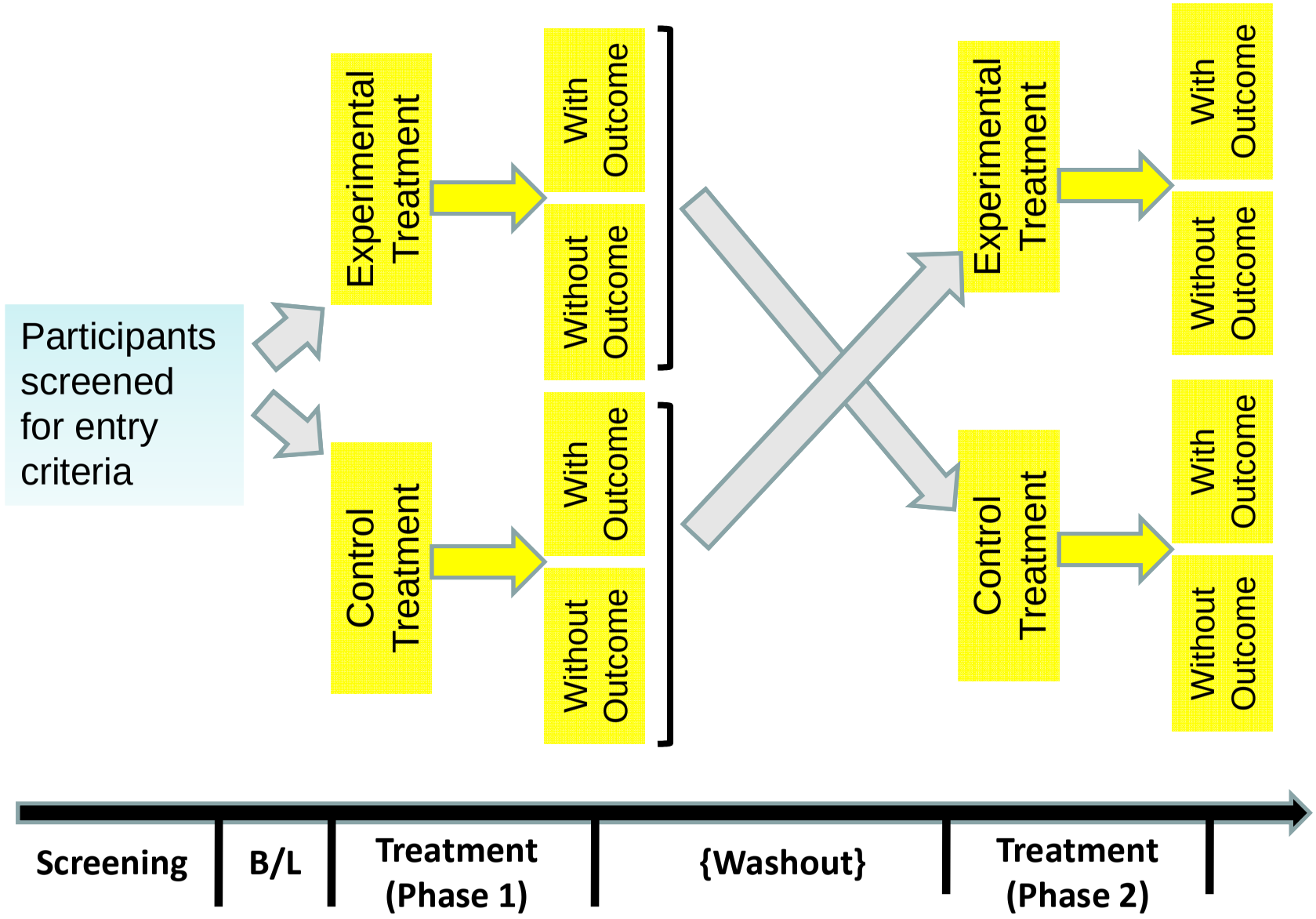
- **Disadvantages**
  - *Effective sample size less than number of subjects*
  - *Many units must participate to overcome unit-to-unit variation, thus requires larger sample size*
  - *Need cluster sampling methods*

Cross Over Design
H_0: A vs. B

		Period	
	Group	I	II
AB	1	TRT A	TRT B
BA	2	TRT B	TRT A

- Advantage
  - Each patient their own control
  - Smaller sample size
- Disadvantage
  - Not useful for acute disease
  - Disease must be stable
  - Assumes no period carry over
  - If carryover, have a study half sized (Period I A vs. Period I B)

Crossover Design



Factorial Design

- Schema

		Factor I	
		Placebo	**Trt B**
Factor II	**Placebo**	**N/4**	**N/4**
	Trt A	**N/4**	**N/4**

A vs. Placebo

B vs. Placebo

Factorial Design

- Advantages
  - Two studies for one
  - Discover interactions

- Disadvantages
  - Test of main effect assumes no interaction
  - Often inadequate power to test for interaction
  - Compliance

- Examples
  - Physicians' Health Study (PHS) *NEJM* 321(3):129-135, 1989.
  - Final report on the aspirin component
  - Canadian Cooperative Stroke Study (1978) *NEJM* p. 53

Superiority vs. Non-Inferiority

Superiority Design:

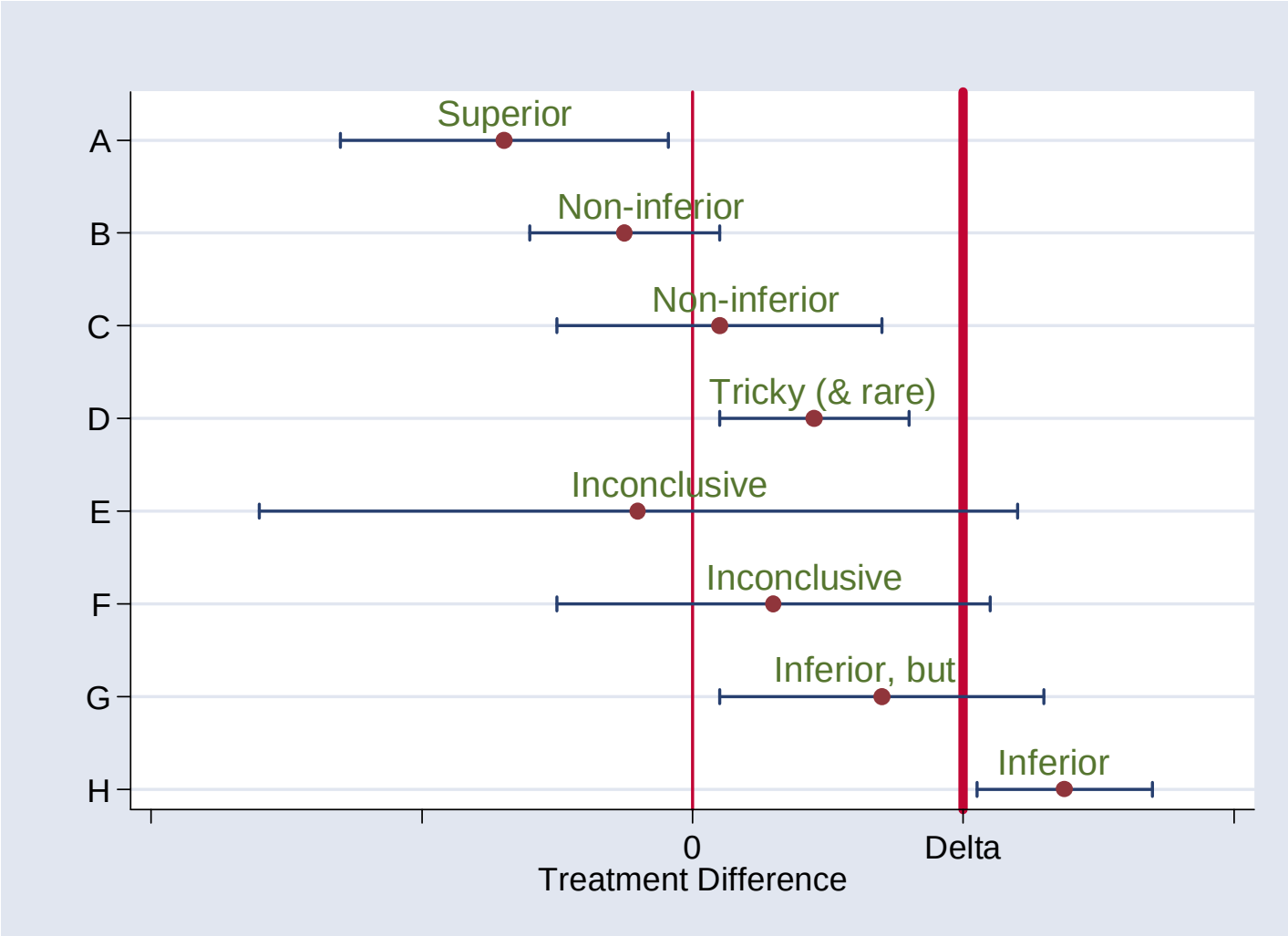
shows that new treatment is better than the control or standard (maybe a placebo)

Non-inferiority Design:

shows that the new treatment

a) is not worse than standard by more than some margin
b) Would have beaten placebo if a placebo arm had been included (regulatory)

Possible outcomes in a non-inferiority trial (observed difference & 95% CI)



← New Treatment Better New Treatment Worse →

Phase IV trials

- Post-marketing surveillance studies
- Assess long-term toxic effects & risk-benefit ratio
- Optimize use of the drug / device / intervention
- Study specific patient population i.e children

Research: A Big Challenge

- **CONCEIVE**
- **CONDUCT**
- **COMMUNICATE**

Understanding of research study designs & principles helps

MISSION

SERVICE

EDUCATION

RESEARCH