Types of study designs: from descriptive studies to randomized controlled trials

October 13, 2016
Derek Stephens
Dalla Lana School of Public Health

Objectives

- To understand the difference between descriptive and analytic studies
- To identify the hierarchy of study designs, and the strengths and weaknesses of each design
- To be able to apply different study designs to the same research question

TWO STUDIES TO BE USED TO ILLUSTRATE DIFFERENT DESIGNS

- UCSF medical school students
- Femur Fracture @ sickkids (multicenter)

Research Question

Is the regular consumption of Red Bull *associated* with improved academic performance among U.S. medical students?

Femur Fracture study

- Research Question: Does one type of surgery produce better healing in Femural fractures?

  (Femur = thigh bone)

- Outcome: Mal-union of Femur/Healing

- Treatments/Surgical interventions

  X-fix

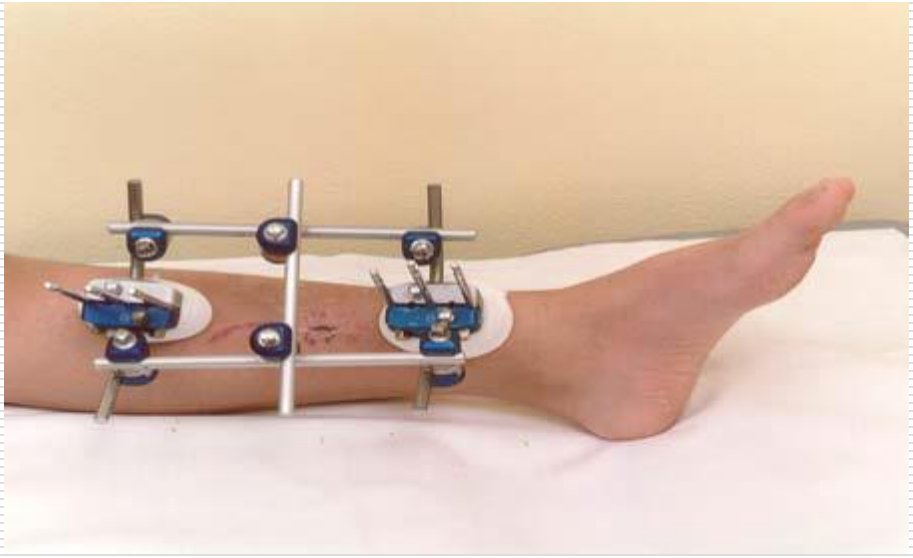
  External fixation is a method of holding broken bones in place. An ***external fixator*** has screws that are inserted into the bone

  Spica

  - A ***hip spica*** cast is used to treat hip and thigh problems, most often in young children. The ***hip spica*** cast can be cumbersome etc…

Hip Spica

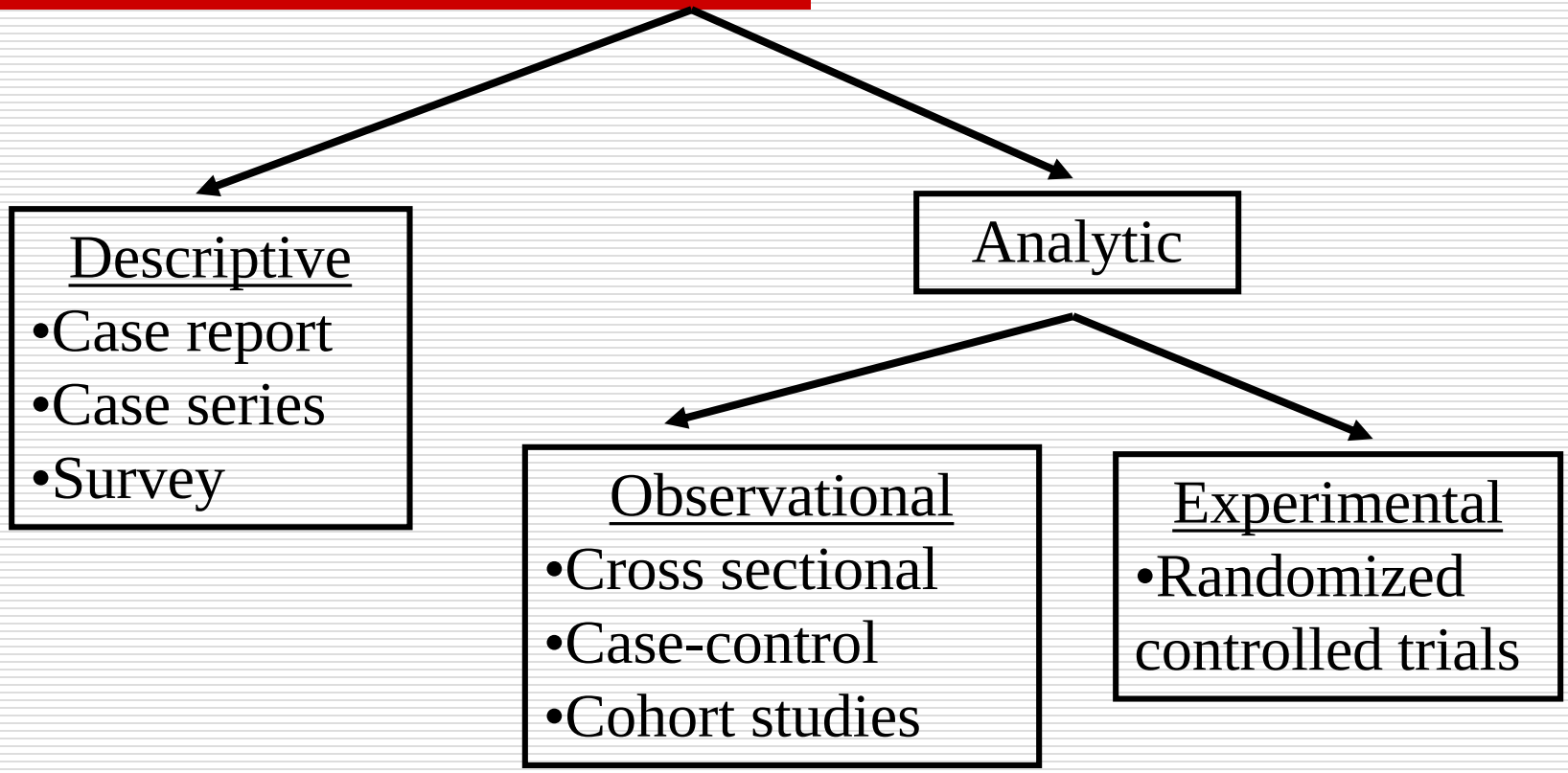
External Fixator



Types of Studies

- Descriptive Studies
- Observational Analytic Studies
  - Cross Sectional studies
  - Case Control studies
  - Cohort studies
- Experimental Studies
  - Randomized controlled trials

Hierarchy of Study Types



Strength of evidence for causality between a risk factor and outcome

Descriptive studies

- Getting a "lay of the land"
  - Surveys
    - How many types of surgery were done for femur fractures and respective success rates?
    - Reporting of proportions, means, and measures of spread
- Describing a novel phenomena
  - Case reports or case series
    - Sudden Infant deaths (SIDS)

Descriptive studies

- Cannot establish causal relationships
- Still play an important role in describing trends and generating hypotheses about novel associations
  - Association between retinal hemorrhage and Shaken Baby Syndrome
- Boxplots are a good way to show the central values and spread of the data.

Analytic Studies

- Attempt to establish a *causal* link between a predictor/risk factor and an outcome.

- You are doing an analytic study if you have any of the following words in your research question:
  - *greater than, less than, causes, leads to, compared with, more likely than, associated with, related to, similar to, correlated with*

Research Question

Is the regular consumption of Red Bull *associated* with improved academic performance among U.S. medical students?

Red Bull

- "Functional drink" designed for periods of mental and physical exertion.
  - reported to increase performance, concentration, memory, vigilance, and emotional balance
  - decrease reaction time…etc.
- Taurine + glucuronolactone + caffeine

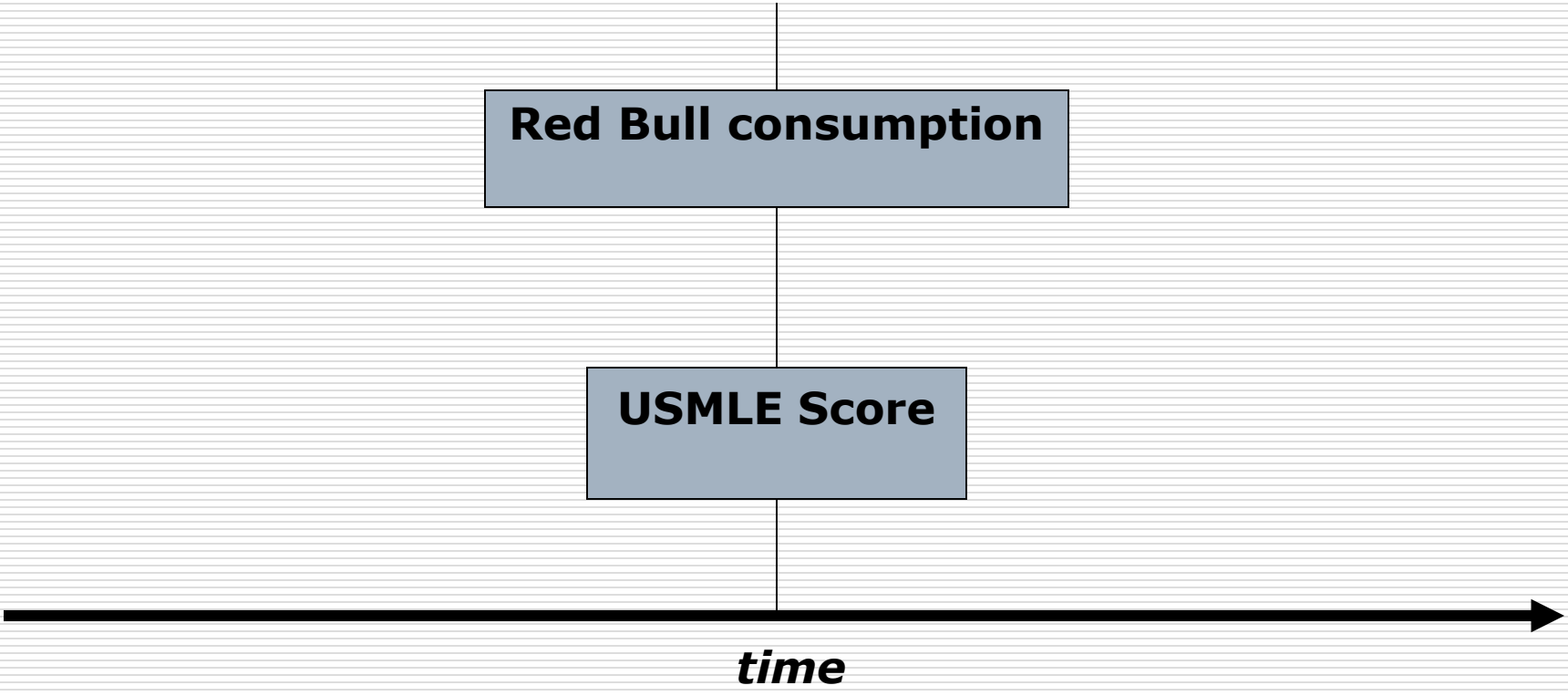
Great idea, but how do you get started….

- Interesting, novel, and relevant, but…
- You only have a very *limited budget* to start investigating this question.
- What (design) is feasible?

Study Design #1

- Cross-sectional study of UCSF medical students taking USMLE Step 2
- Questionnaire administered when registering for USMLE 2
  - Primary predictor: self-report of >3 cans Red Bull per week for the previous year
  - Factors/Covariates: Age, sex, undergraduate university, place of birth
- Outcome: Score on USMLE Step 2

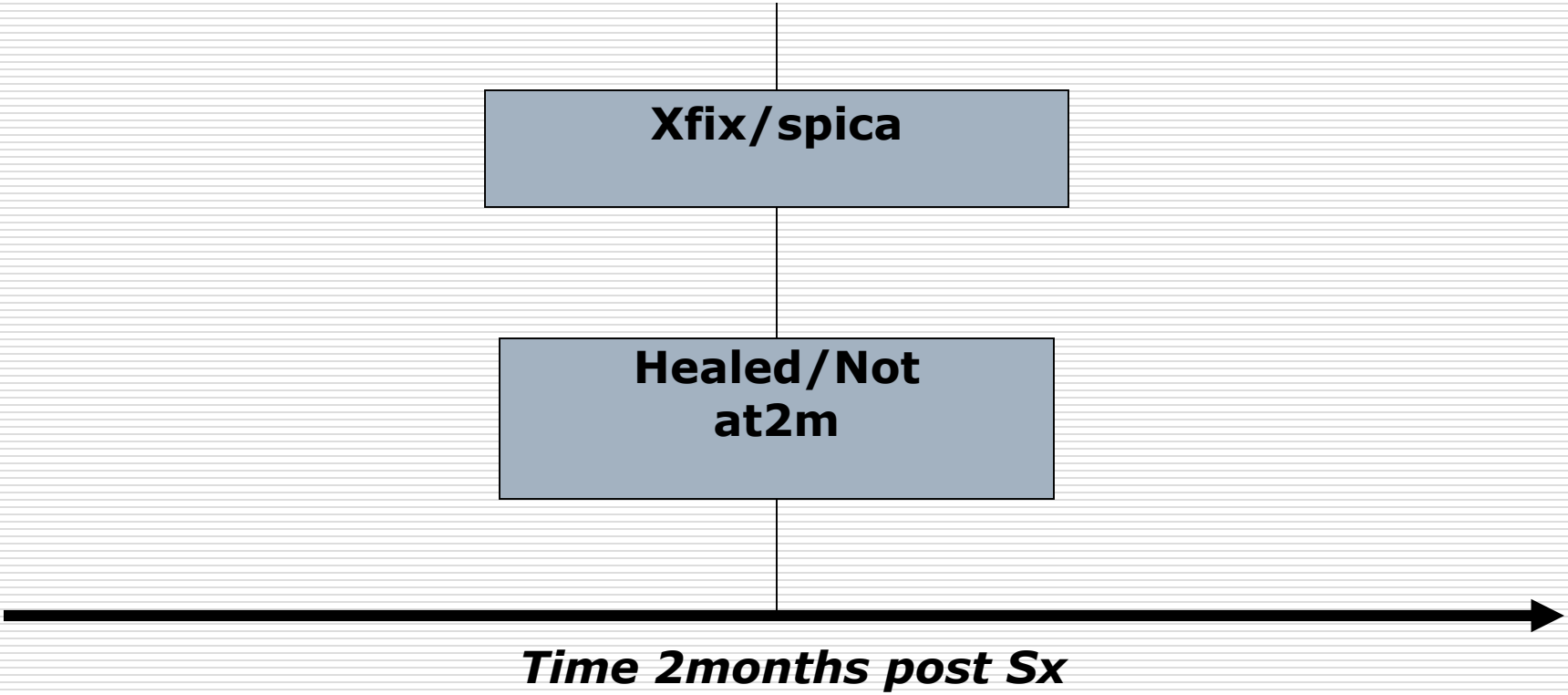
Cross-sectional study: structure



Study 2 (Femur Fracture)

- Cross-sectional study of patients at clinic visit 2 months
- Questionnaire administered when registering for clinic visit at 2months postSx
  - Primary predictor: Type of surgery
  - Factors: Age, sex, centre
- Outcome: Femur healed/not healed

Cross-sectional study: structure



Cross-sectional Study:

- Descriptive value:
  - How many patients had type x-fix and of those how many healed at 2 months.
  - How many patients had type spica and of those how many healed at 2 months.
  - What is the age and sex distribution of patients having xfix and spica ?
- Analytic value:
  - *Can the question of whether there is a difference in the 2 surgical treatments be answered using a cross-sectional design?*

Cross-sectional Study: Pluses

+ **Prevalence**

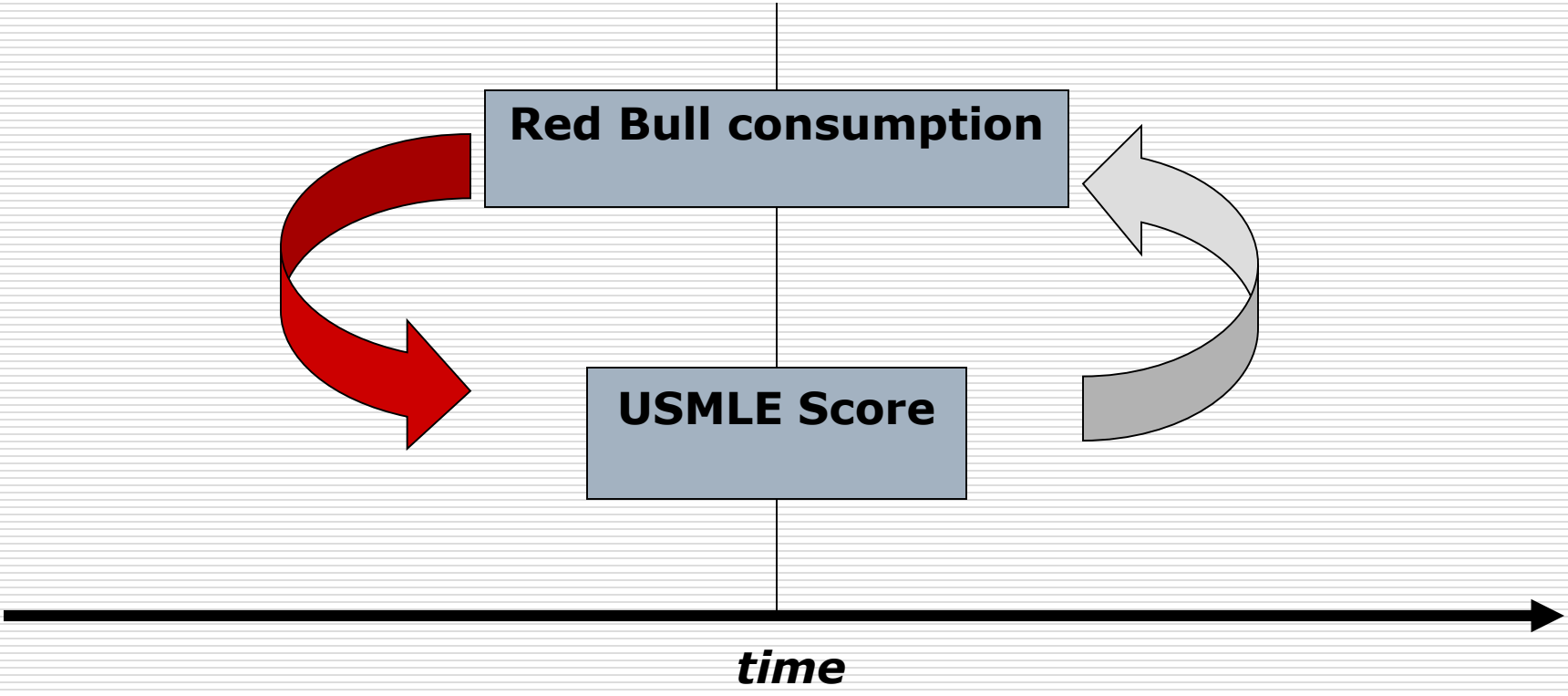
+ **Fast/Inexpensive - no waiting!**

+ **No loss to follow up**

+ **Associations can be studied**

Cross-sectional study: minuses

- **No time element**



Cross-sectional study: minuses

- **Cannot determine causality**

- **Cannot study rare outcomes**

What if you are interested in the rare outcome?

- The association between regular Red Bull consumption and…
  - A perfect score on the USMLE – Step 2
  - Graduating top 1% of the medical school class
  - Acceptance into a highly selective residency

(all of the above are examples of rare outcomes)

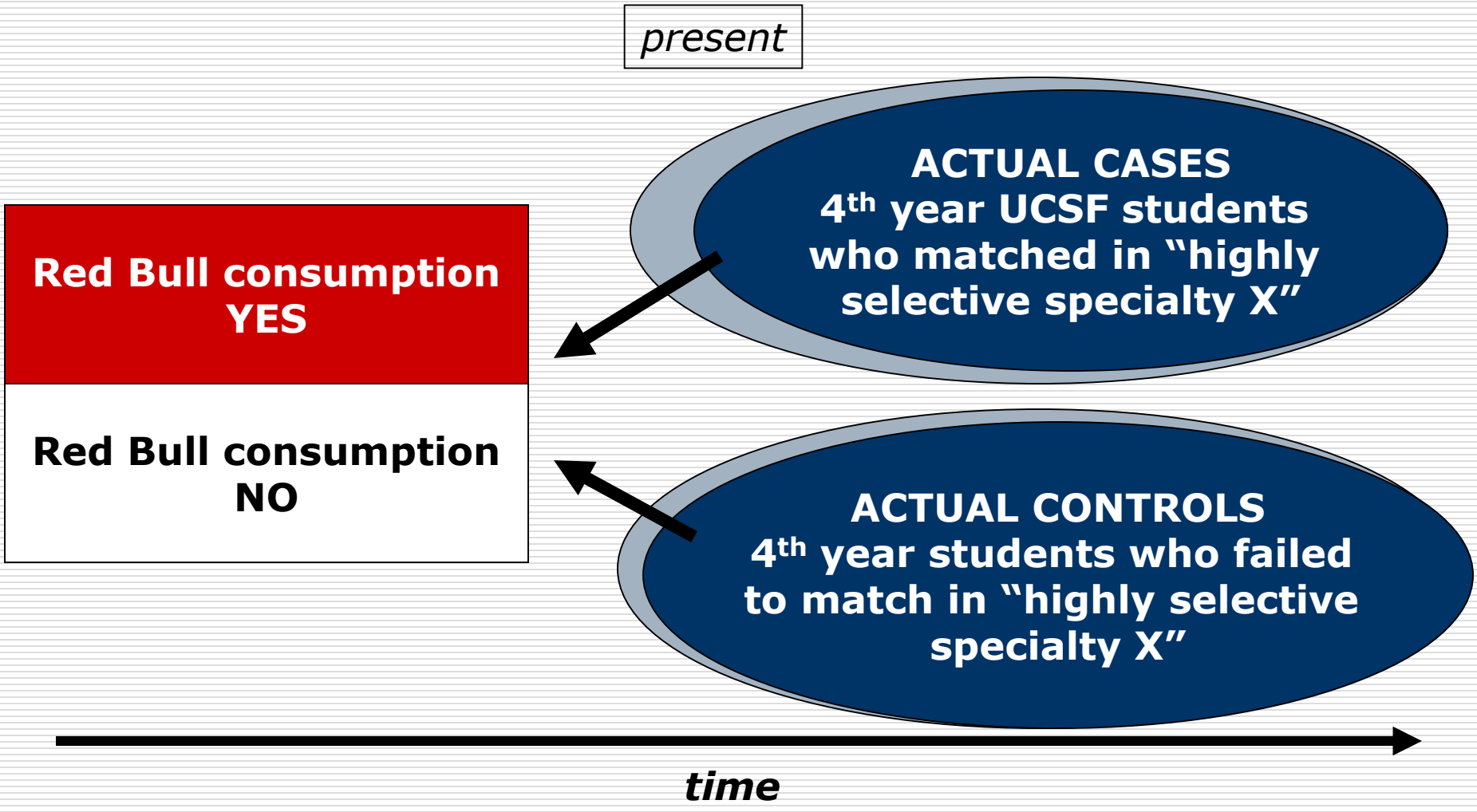
Leads us to Study Design #2

- A case-control study
- Cases: 4th year med students accepted to residency in "highly selective specialty X".
- Controls: 4th year med students who applied but were not accepted.
- Predictor: self-reported regular Red Bull consumption
  - Additional covariates (age, sex, medical school, undergraduate institution)

Case control studies (main feature)

- **Investigator works "backward" (from outcome to predictor)**
- **Sample chosen on the basis of outcome (cases), plus comparison group (controls)**

Case-control study structure



Case control studies

- **Can estimate the strength of the association between each predictor variable and the presence or absence of disease**
- **Cannot yield estimates of incidence or prevalence of disease in the population (why?)**
- **Odds Ratio is statistic used to measure association**

Case-control Study: pluses

+ **Rare outcome**

+ **Inexpensive and efficient: may be only feasible option**

+ **Establishes association (Odds ratio)**

+ **Useful for generating hypotheses (multiple risk factors can be explored)**

Case-control study-minuses

- **Causality still difficult to establish**
  - **Selection bias (appropriate controls)**
    (controls should be matched as closely as possible to cases except on the exposure)
  - **Recall bias: sampling (retrospective)**
  - **Cannot tell about incidence or prevalence**

Recall bias

- *Recall or memory bias*.
  - subjects recall past events.
  - Often a person recalls positive events more than negative ones.
  - Alternatively, certain subjects may be questioned more vigorously than others, thereby improving their recollections.

Selection Biases

- Selection biases.
  - Occurs when the groups to be compared are different. These differences may influence the outcome. Common types of sample (subject selection) biases include volunteer or referral bias, and nonrespondent bias.

Selection of subjects in Case-Control designs

- Controls must fulfil all the eligibility criteria defined for the cases apart from those relating to diagnosis of the disease.
- In case–control studies, controls should represent the population from which the cases are drawn, i.e., they should provide an estimate of the exposure prevalence in the population from which the cases arise. If not, the results of the study are likely to be distorted because of selection bias.
- In a nested case–control study, it is relatively straightforward to ensure that the cases and controls are drawn from the same study population, since both will arise from a clearly defined population–the cohort. In general, all the cases arising as the cohort is followed prospectively become the 'cases' in the case–control study, while a sample of unaffected members of the cohort become the 'controls'.

Measures of association

		Disease: Yes	Disease: No
Exp	Yes	**a**	**b**
	No	**c**	**d**

Odds ratio: ad / bc

- In the previous table a+c = number of cases. Of these a = number of the cases with the exposure. So an estimate of the proportion of exposed in disease = a/(a+c) and proportion of unexposed in disease = c/(a+c).
- So Odds of exposure in diseased = [a/(a+c)]/[c/(a+c)] = a/c
  (note odds (of an event) = p/(1-p))
- Similarly odds of exposure in controls =b/(b+d)/d/(b+d) = b/d
- So odds ratio of exposure in diseased = ratio of the above odds = ad/bc
- A measure of the "likelihood" of exposure given disease

Measures of association. Relative Risk

Risk Factor	Disease: Yes	Disease: No
Yes	A	B
No	C	D

Risk ratio (relative risk)

$$\frac{\frac{A}{A+B}}{\frac{C}{C+D}}$$

Forward direction

Equivalence of OR and RR

- When is ad/bc approximately equal to [a/(a+b)]/[c/(c+d)]
- If a << b then a+b ~ b
- If c << d then c+d ~ d

- Then [a/(a+b)]/[c/(c+d)] = [a/b]/[c/d]
- Which is ad/bc
- Equivalence holds when prevalence is low

Where are we in answering research question #1

- (Say) Preliminary results from our cross-sectional and case-control study suggest an association between Red Bull consumption and improved academic performance among medical students

- What's missing? - strengthening evidence for a causal link between Red Bull consumption and academic performance

- Use results from our previous studies to apply for funding for a prospective cohort study!

Study design #3

- Prospective cohort study of UCSF medical students Class of 2009
- All entering medical students surveyed regarding beverage consumption and variety of other potential covariates
- Survey updated to record changes in Red Bull consumption
- Outcomes: USMLE score, match in first choice residency

Cohort studies

- A cohort (follow-up, longitudinal) study is a *comparative, observational* study in which subjects are *grouped by their exposure status*, i.e., whether or not the subject was exposed to a suspected risk factor

- The subjects, exposed and unexposed to the risk factor, are followed forward in time to determine if the disease occurs
  - Subjects should be disease free on entry(i.e. USMLE score is not known)
  - No new subjects allowed in after initial recruitment

- The rates of disease incidence among the exposed and unexposed groups are determined and compared.

Elements of a cohort study

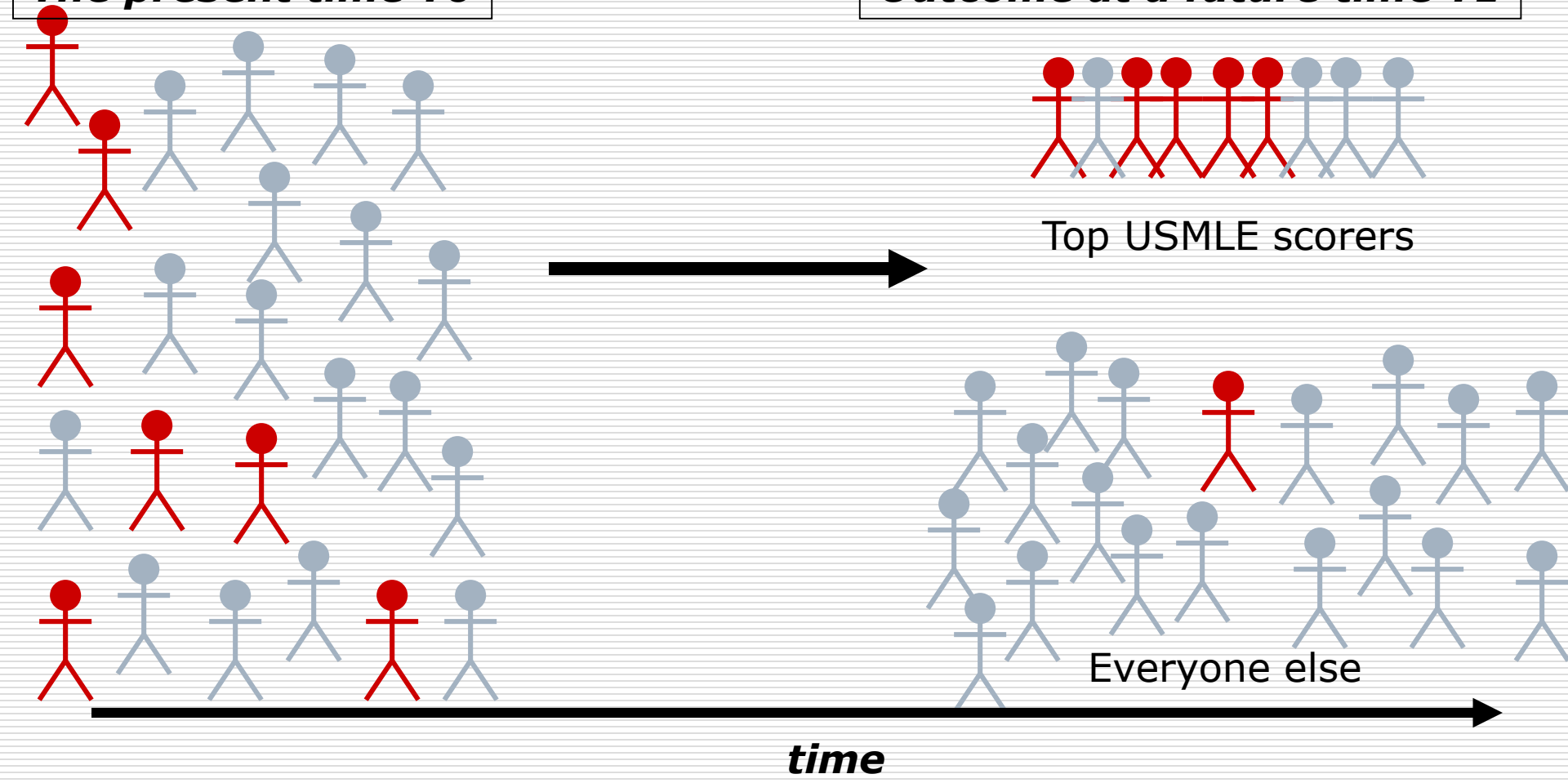
- Selection of sample from population
  - Problems: self select, physician selected
- Measures of predictor variables in sample
- Follow population for period of time
  - For prospective/retrospective there is a time element, Forward/backward, respectively)
  - Same follow up time for all groups being followed
- Measure outcome variable

Prospective cohort study structure

Red represents Red Bull consumption

The present time T0

Outcome at a future time T1



Strengths of cohort studies

- Know that predictor variable was present before outcome variable occurred (some evidence of causality)
- Directly measure *incidence* of a disease outcome
- Can study multiple outcomes of a single exposure (RR is measure of association)

Weaknesses of cohort studies

- Expensive and inefficient for studying rare outcomes
- Often need long follow-up period or a very large population
- Loss to follow-up can affect validity of findings

Other types of cohort studies

- Retrospective cohort
  - Identification of cohort, measurement of predictor variables, follow-up and measurement of outcomes have all occurred in the past
  - Much less costly than prospective cohorts since often a chart review.
  - Investigator has minimal control over study design

Other types of cohort studies

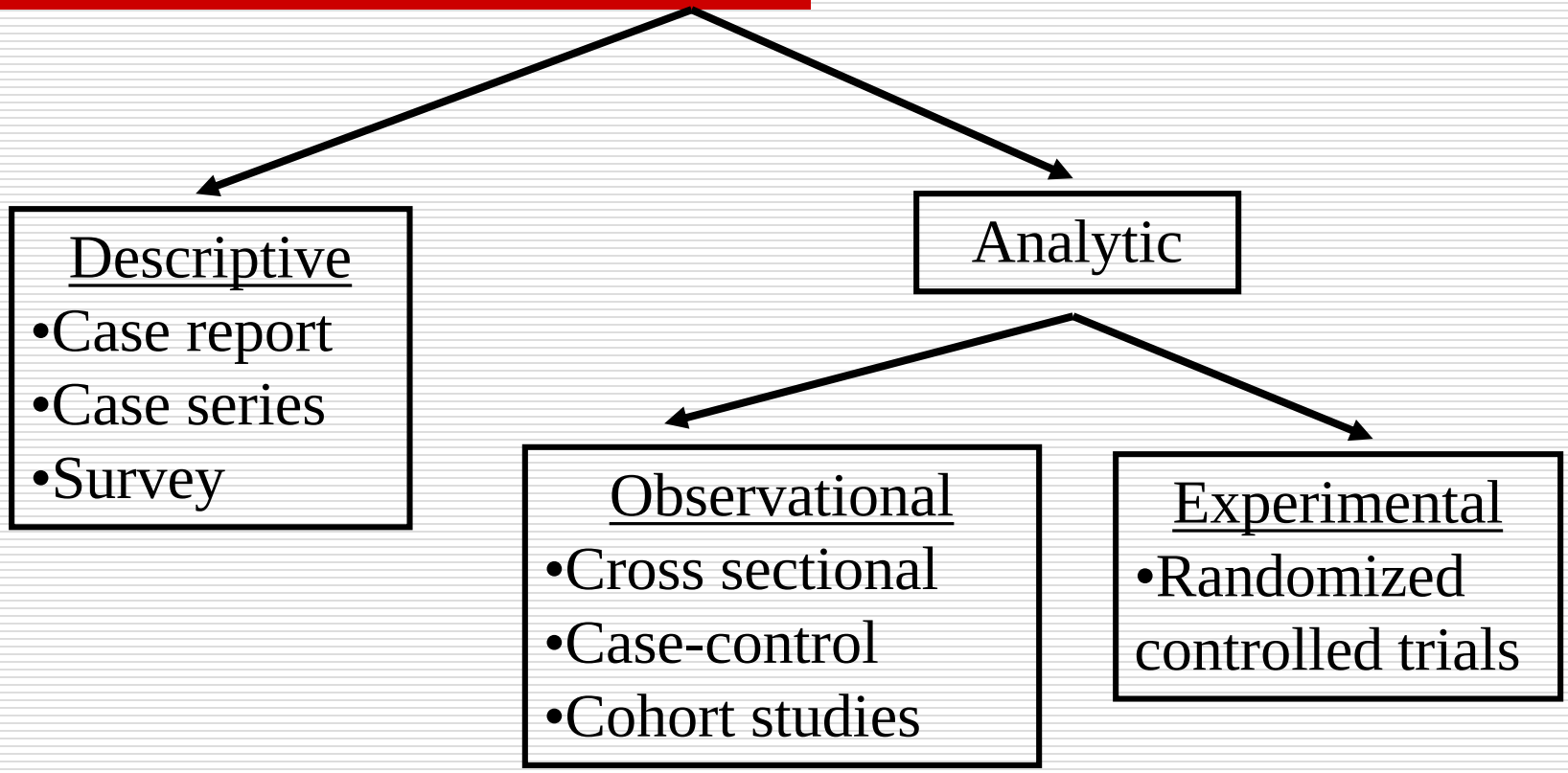
- Nested case-control study
  - Case-control study embedded in a cohort study offers impressive reductions in costs and efforts of data collection and analysis compared with the full cohort approach, with relatively minor loss in statistical efficiency
  - Controls are drawn randomly from study sample
  - Matching controls to cases within the cohort could make the design more efficient.

Other types of Cohort Designs

- Crossover designs
- Each patient receives both(all) treatments
  - Usually requires fewer subjects variance is smaller than independent samples design
  - $[\sigma^2(d) = \sigma^2(1) + \sigma^2(2) - 2\rho\sigma(1)\sigma(2)]$

    What happens if the correlation ρ is < 0?
  - Serious problem is the carryover effect of the treatments, if sufficient *Washout* is not built in.
  - If carryover effect is observed then can only use first time point measurements
  - Power decreases

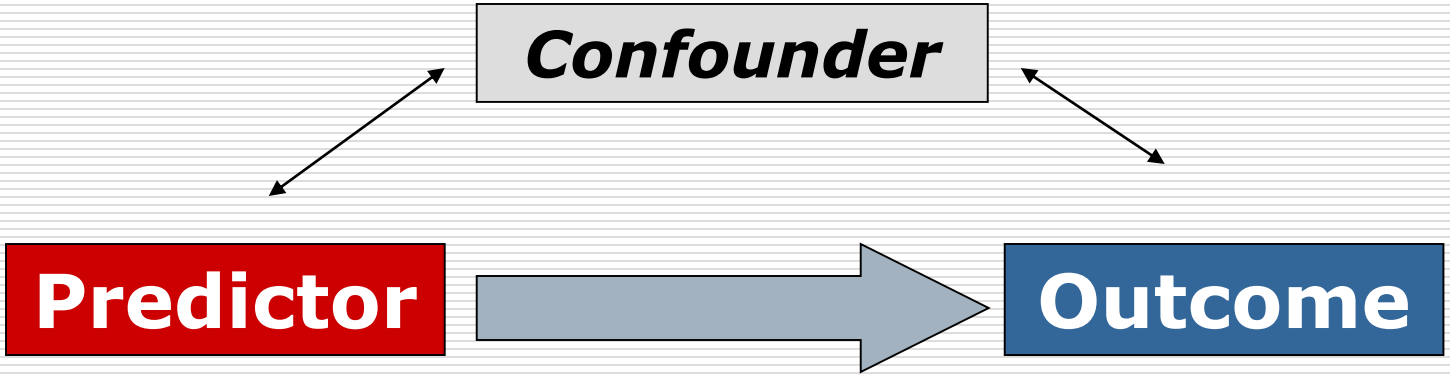
Hierarchy of Study Types



Strength of evidence for causality between a risk factor and outcome

What distinguishes observational studies from experiments?

- Ability to control for confounding



Examples:

sex (men are more likely to drink red bull and men are more likely to match in neurosurgery)

Undergraduate institution (students from northwest school are more likely to drink red bull and also more likely to score higher on USMLE)

But we measured all of the potential confounders…….

- In an observational prospective cohort study you can (maybe) measure all potential *known* confounders, but…

- You can't control for unanticipated or unmeasured confounders. (i.e. *unknown variables)*

Brings us to Study Design # 4

The RCT

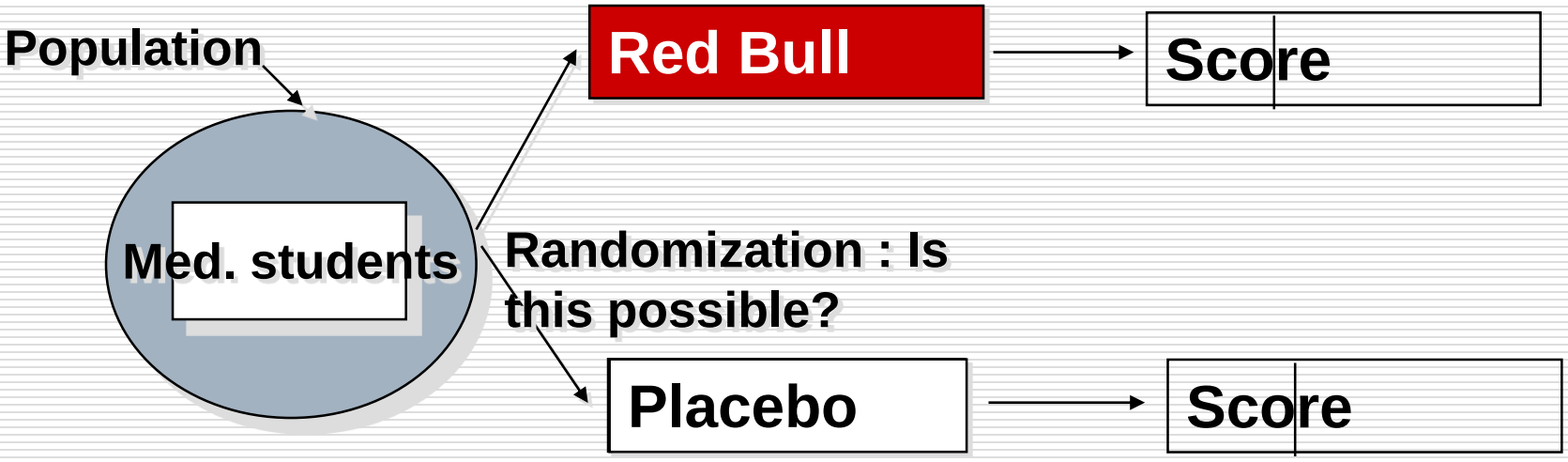
- Randomized controlled trial of daily Red Bull consumption among entering UCSF medical students Class 2009
- Randomly allocate students to daily consumption of Red Bull vs. daily consumption of placebo (*May not be feasible how can this be done for the Red Bull question*)

- Outcome: USMLE Step 1 score, USMLE Step 2 score, match in first choice residency

Randomized Controlled Trials

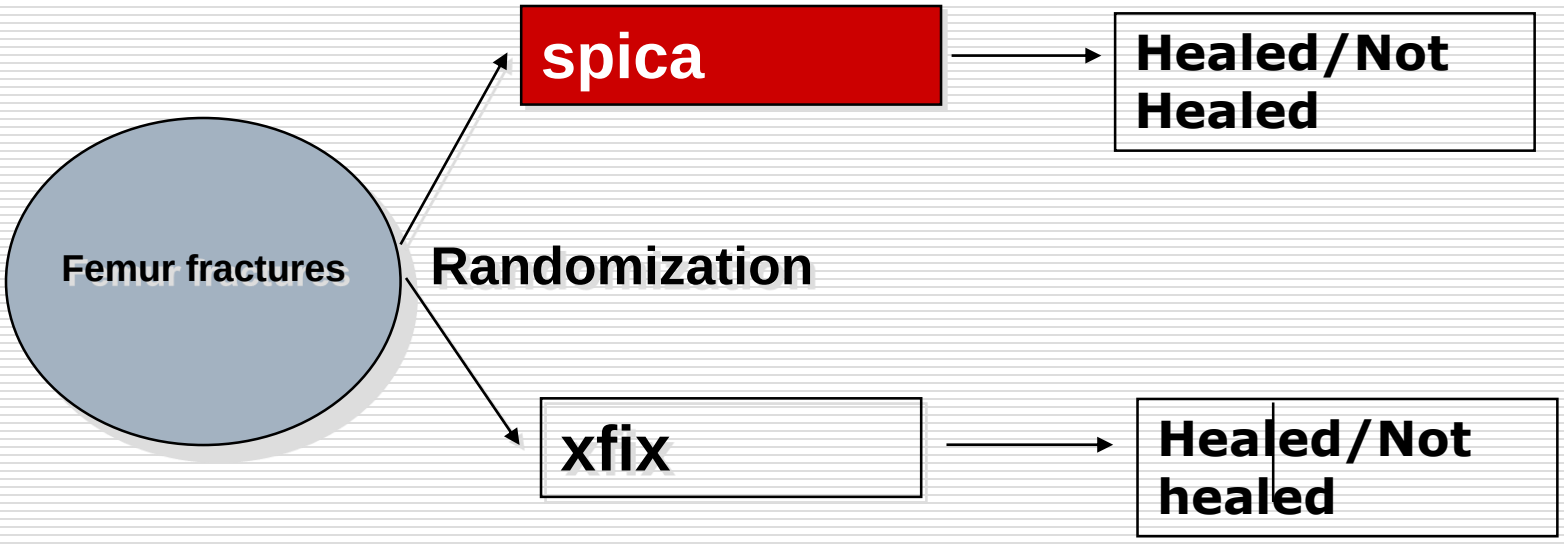
- Investigator controls the predictor variable (intervention or treatment)
- Major advantage over observational studies is ability to demonstrate causality
- Randomization used to control unmeasured or unknown confounding variables.
- Only for mature research questions

Basic Trial Design



Back to surgical Femur Fracture trial

- Can best be answered using this design.



- Randomization more feasible in this example

Steps in a randomized controlled trial

1. Select participants(inclusion criteria)
   - high-risk for outcome (high incidence)
   - Likely to benefit and not be harmed
   - Likely to be compliant
     - (to avoid problems of missing data)

2. Measure baseline variables

3. Randomize
   - Eliminates baseline confounding
   - Randomization could be simple, stratified, block

Steps in a randomized controlled trial

4. Blinding the intervention
   - As important as randomization
   - Eliminates
     - biased outcome ascertainment
     - biased measurement of outcome
5. Follow subjects
   - Adherence to protocol
   - Lost to follow up
6. Measure outcome
   - Clinically important measures
   - Adverse events

What is Blinding?

- Single blind - participants are not aware of treatment group
- Double blind - both participants and investigators unaware
- Triple blind - various meanings
  - persons who perform tests
  - outcome adjudicators
  - safety monitoring group

Why blind?:
Biased Outcome Ascertainment

- If group assignment is known
  - *participants* may report symptoms or outcomes differently
  - *physicians or investigators* may elicit symptoms or outcomes differently
  - *Study staff or adjudicators* may classify similar events differently in treatment groups

- More Problematic with "soft" outcomes
  - investigator judgment
  - participant reported symptoms, scales

Analysis of RCT

- Analyzed like cohort study with RR or GLM/Mixed analysis(for continuous type outcomes)
- Don’t need to adjust for baseline scores.
- Intention to treat analysis
  - Most conservative interpretation
  - Include all persons assigned to intervention group (including those who did not get treatment or dropped out)

- RCT

- Exposure/treatment by randomization, and hence confounders also randomized.

- Chance imbal. can be ↓ by ↑
- sample size. (i.e. inverse relation between α and N)

- Subject to bias and chance imbalance

- Causality likely

- **Observational Studies**

- exposure "chosen" →
- confounding variables associated with exposure remain

- effects of confounders do not ↓ with ↑ N.

- Subject to bias and chance imbalance

- Cannot make causal statements.

Features of High Quality Randomized Trials

- Tamper-proof randomization
- Blinding of participants, study staff, lab staff, outcome ascertainment and adjudication
- Adherence to study intervention and protocol
- Complete follow-up

Equivalence Trials (optional slides)

- Some studies seek to show that treatments are NOT different.

- These are called equivalence trials
- In a comparative(difference) trials the standard analysis uses statistical significance tests to determine whether the null hypothesis of "no difference" may be rejected, together with confidence limits to place bounds on the possible size of the difference between the treatments.
  - Null: Mean1 = Mean2
    Alternative: Mean1 ≠ Mean2

Equivalence Trials

- In an equivalence trial the conventional significance test has little relevance: failure to detect a difference does not imply equivalence
- A difference which is detected may not have any clinical relevance and may correspond to practical equivalence.
- *The relevance of the confidence interval, however, is easier to see.*

Equivalence Trials

- The CI defines a range for the possible true differences between the treatments, any point of which is reasonably compatible with the Null Hypothesis. If every point within this range does not correspond to a clinically important difference then the treatments may be considered to be equivalent.

- A margin of equivalence must be specified before the study.

Equivalence Trials

- If we have predefined a range of equivalence as an interval from - to + we can then simply check whether the confidence interval centered on the observed difference lies entirely between - and + .
- If it does, equivalence is demonstrated; if it does not, there is still room for doubt.
- It is important to emphasize that absolute equivalence can never be demonstrated
- It is possible only to assert that the true difference is unlikely to be outside a range which depends on the size of the trial(N), the results of the trial(δ,σ) and the specified probabilities of error(α, β).

Equivalence Trials