Lecture 2

Key Concepts in Clinical Research

Outline

- Key Statistical Concepts
  - Bias and Variability
  - Type I Error and Power
  - Confounding and Interaction
  - Statistical Difference vs Clinical Difference
  - One-sided Test vs Two-sided Test
- Basic Design Considerations
- Data Safety Monitoring Board

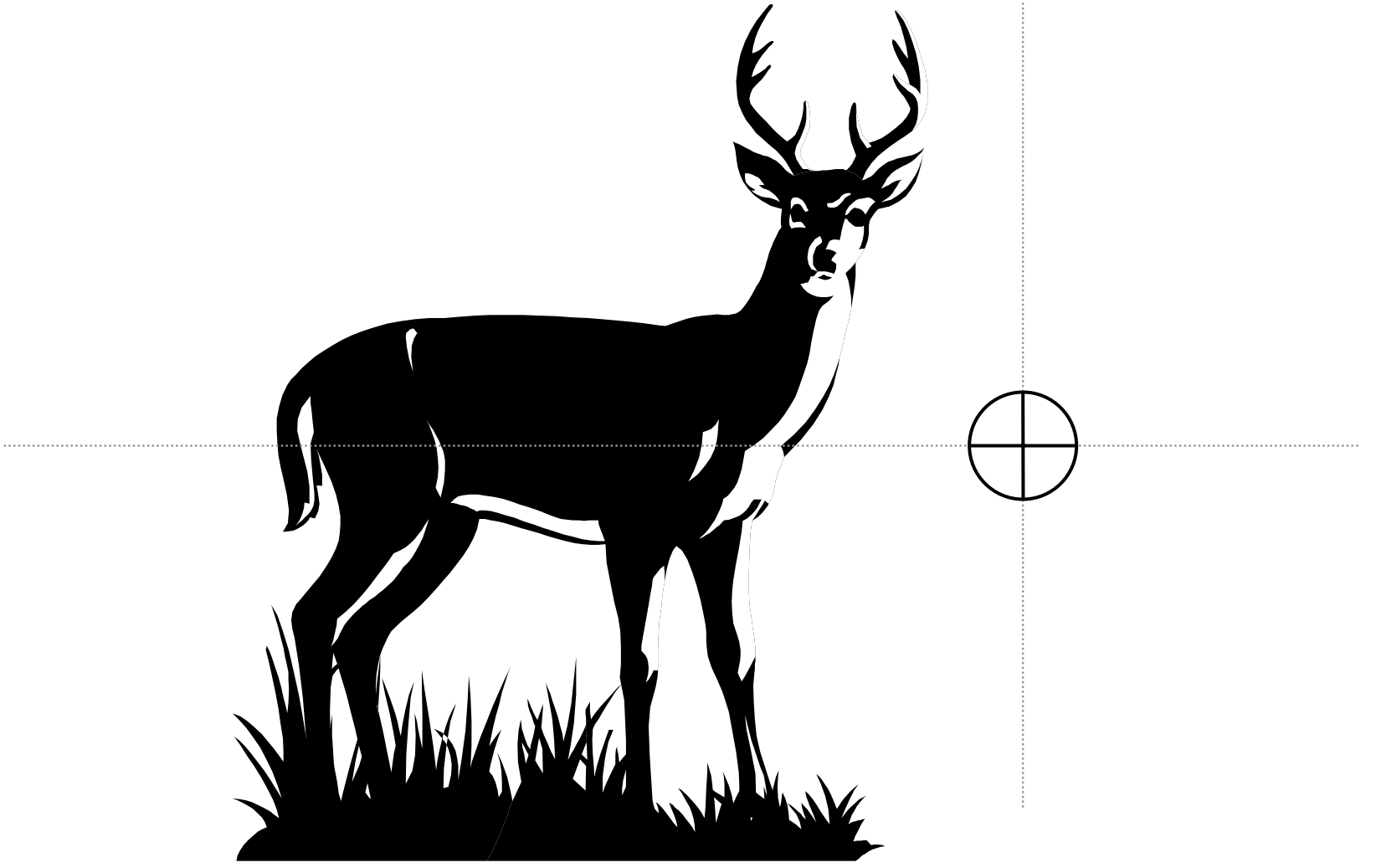
Bias and Variability

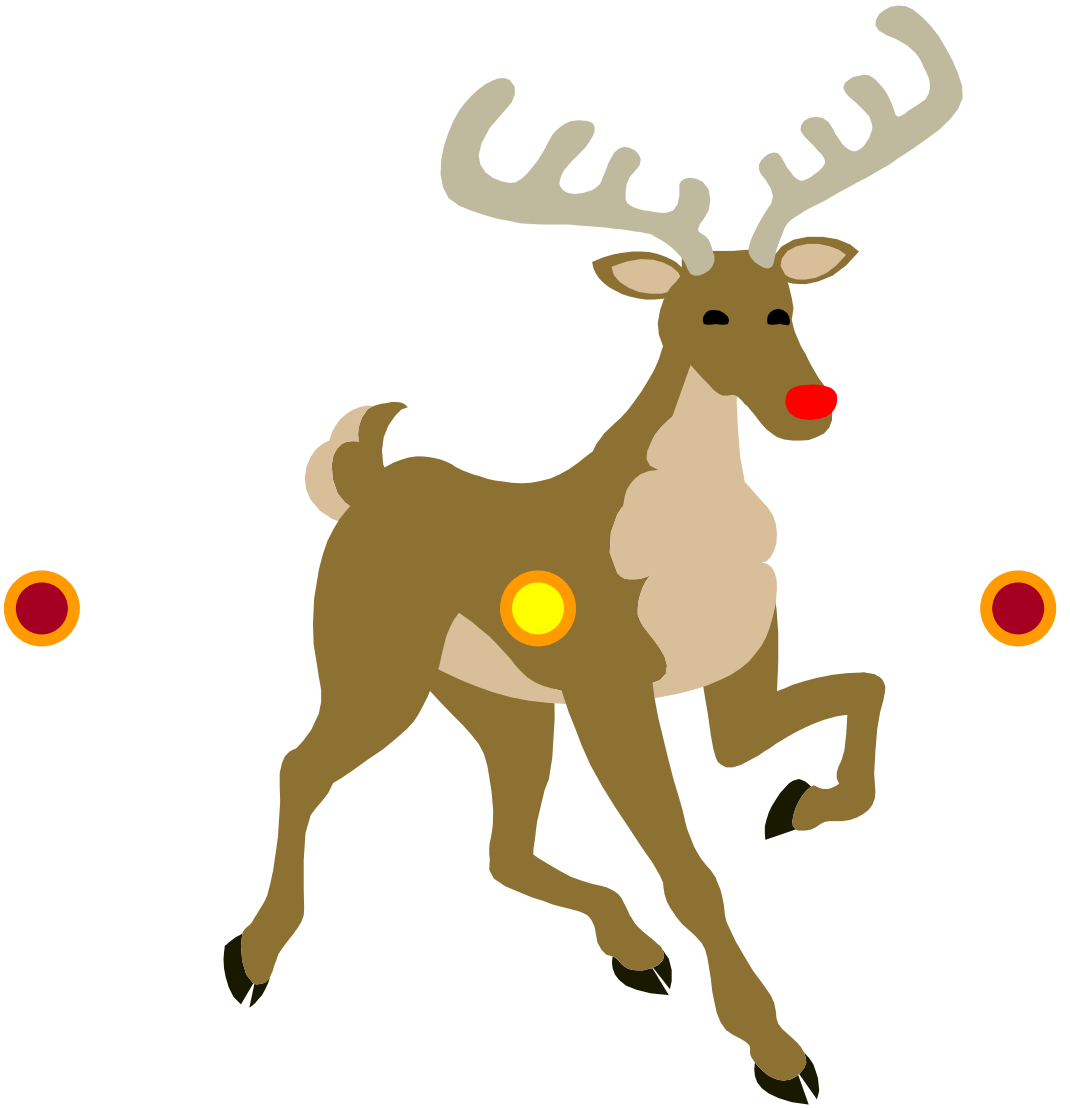
- Bias
  - It measures the **closeness** of the test result to the true value (e.g., population mean)
  - Accuracy
- Variability
  - It measures the **degree of the closeness** of the test result to the true value (e.g., population mean)
  - Precision

Bias and Variability

- Goal
  - Minimize the bias
  - Minimize the variability
- Why?
  - Deer hunting example

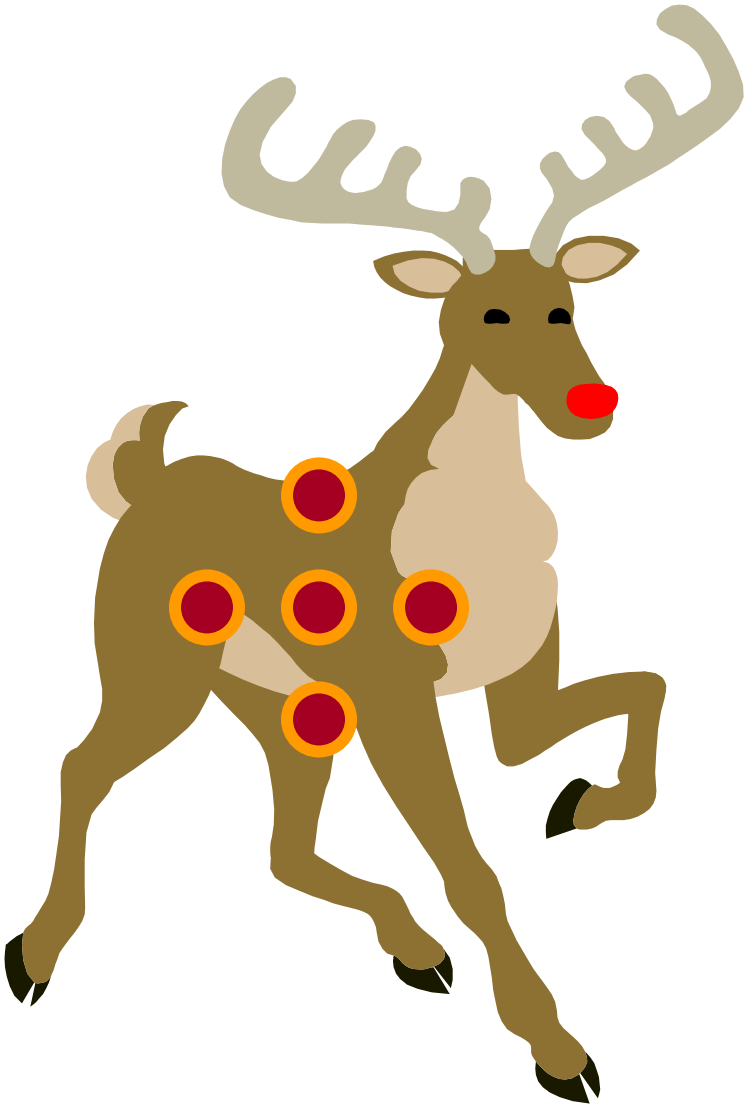
Bias and Variability

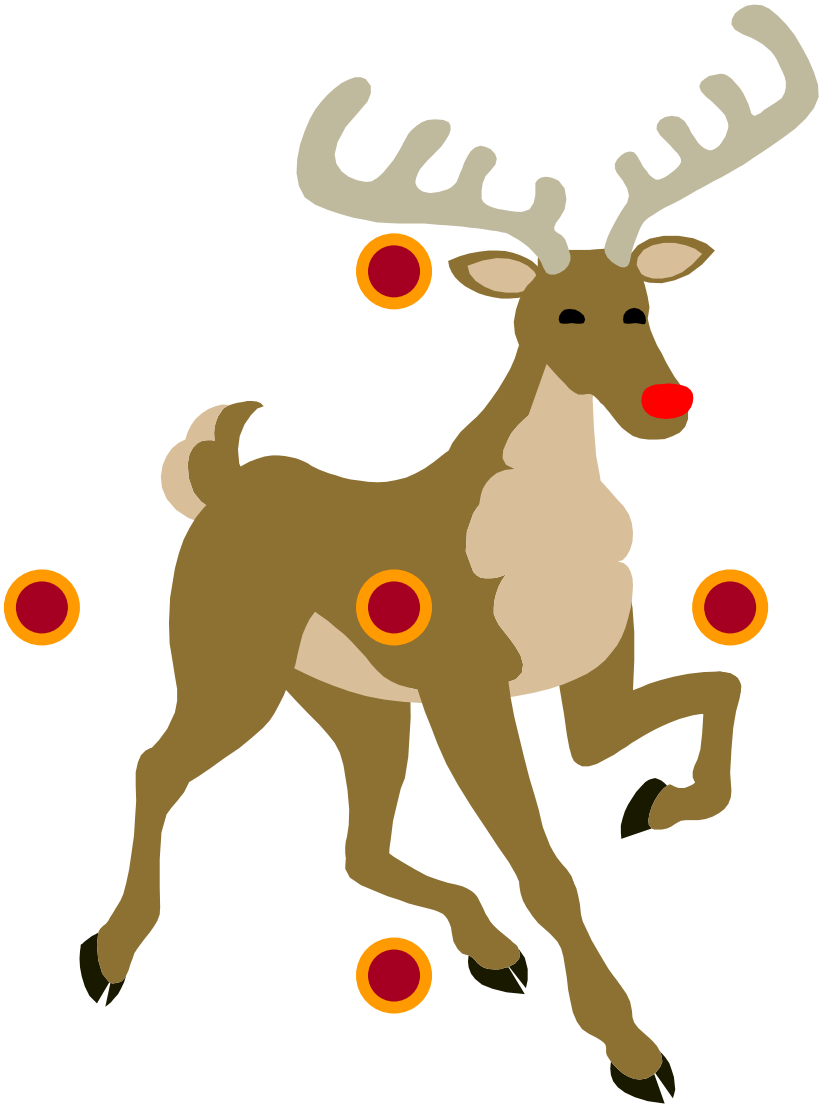




I am not afraid of statisticians !

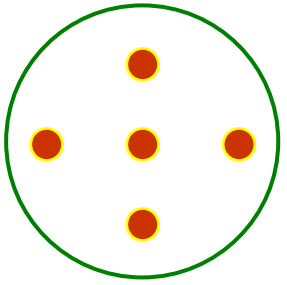
This is what I am afraid of !!



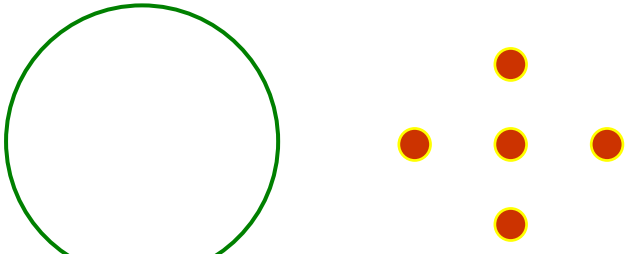


Bias and Variability

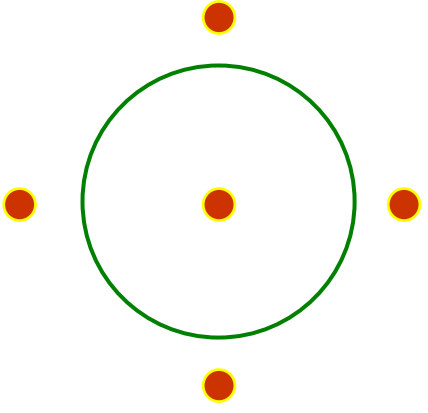
Less bias, small variability



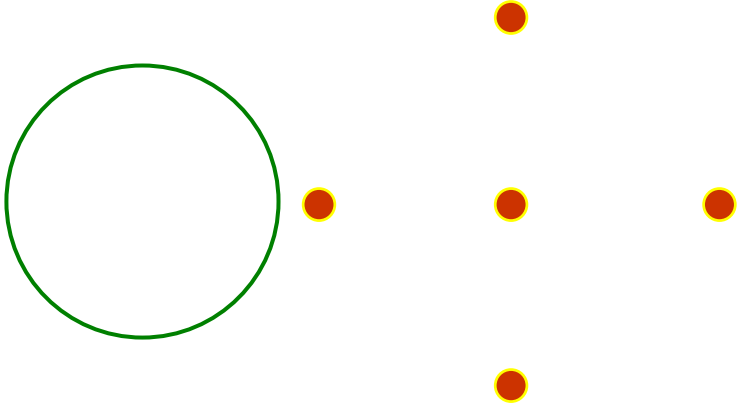
Large bias, small variability



Less bias, large variability



Large bias, large variability



Bias and Variability

- It is not possible to avoid bias and variability in real world.
- It is important to
  - identify,
  - eliminate, and
  - control

  the bias/variability to an acceptable limit.

Bias and Variability

Q: How many types of bias/variability that may incur in product research and development ?

(A) 1

(B) 5

(C) 57

(D) 63

(E) None of above

References

Sackett (1979) and Spilker (1991)

Type I Error and Power

Null hypothesis: The patient is alive

Alternative hypothesis: The patient is dead

Type I error : the patient is dead when in fact the patient is still alive.

Type II error : the patient is still alive when in fact the patient is dead.

P-value : the probability of observing a type I error.

Power : the probability of correctly concluding the death of the patient when the patient is dead.

Type I Error and Power

Null hypothesis: The drug is ineffective
Alternative hypothesis: The drug is effective

Type I error : the drug works when in fact it doesn't.

Type II error : the drug doesn't work when in fact it does.

P-value : the probability of observing a type I error.

Power : the probability of correctly concluding that the drug works when in fact it does.

Type I Error and Power

- Decrease type I error will result in increasing type II error, and consequently decreasing power.
- Increase sample size will decrease both type I and type II errors.
- Fixed type I error and select a sample size to achieve the desired power.

Confounding and Interaction

- Confounding
  - Confounding effects are defined as effects which are contributed by various factors that cannot be separated by the design under study.
- Interaction
  - The interaction effect between factors is defined as the joint effect contributed by more than one factor.

Confounding

- Heating Example
  - Mr. Smith's winter electricity bill for heating was very high last year
  - Mr. Smith decided to improve the insulation of his house
  - This year, Mr. Smith's winter electricity bill for heating is much lower than last year
- Question
  - This winter is much warmer than last year. We do not know the reduction in electricity bill is due to the warmer winter or due to the insulation of the house

Interaction

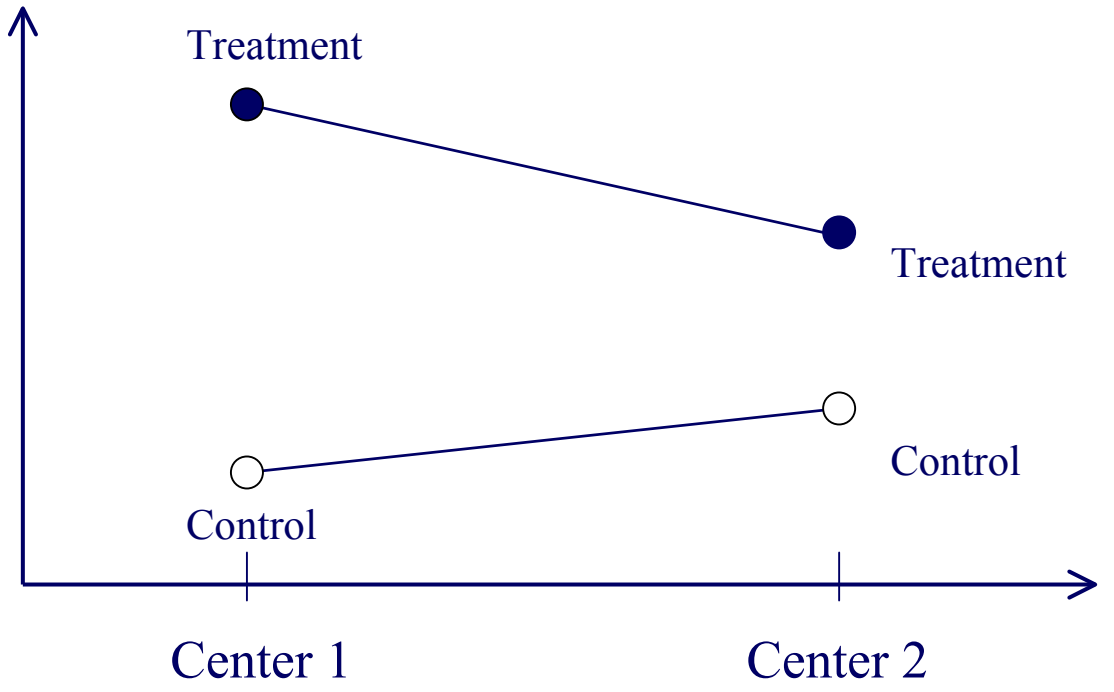
- Quantitative interaction

  Quantitative interaction between treatment and center (or study site) indicates that the treatment differences are in the same direction across centers but the magnitude differs from center to center.

- Qualitative interaction

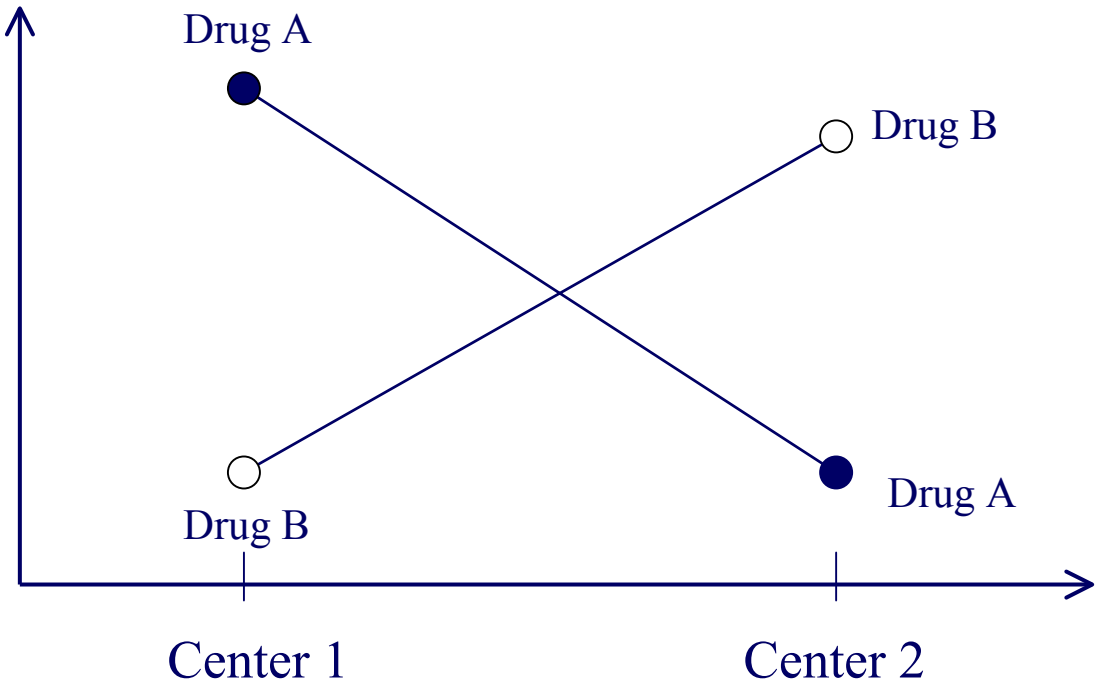
  Substantial treatment differences occur in different directions in different centers.

Study-by-Treatment Interaction



Quantitative interaction - centers could be combined.

Study-by-Treatment Interaction



Qualitative interaction - centers cannot be combined.

Interaction

- Drug Product Example
  - In San Diego, drug A is better than drug B
  - In New York, drug B is better than drug A
- Question
  - Should we sale drug A in San Diego and sale drug B in New York area?
  - No overall assessment of the product difference can be made.

Confounding and Interaction

- Study design should avoid or be able to account for potential
  - Confounding factors
  - Interaction factors
- Objectives
  - To provide a valid and fair assessment of the treatment effect
  - To assess the treatment difference efficiently

Statistical Difference vs. Scientific Difference

- Clinical Scientists & Researchers
  - The observed difference is of clinical meaning and yet not statistically significance.
  - The observed difference is of little clinical meaning but it is statistically significant.
- Statisticians
  - P-value must be less than 0.05.

Statistical Difference vs. Scientific Difference

- Clinical Scientists & Researchers
  - The observed difference is of clinical meaning and yet not statistically significance (You must be out of your mind!)
  - The observed difference is of little clinical meaning but it is statistically significant (Who cares?)
- Statisticians
  - P-value must be less than 0.05 in order to have statistical meaning

Statistical Difference vs. Scientific Difference

- Statistical difference
  - A difference which is unlikely to occur by chance alone.
- Clinical/Scientific difference
  - A difference which is considered important to the clinical scientists.

Statistical difference vs. Clinical Difference

Statistical difference

Clinical difference

	Significant	**Non-Significant**
Significant	No Confusion	*
Non-Significant	*	No Confusion

* May be due to **large variability** and/or **small sample size**

One-sided Test vs. Two-sided Test

- Pharmaceutical companies
  - Would not run a study if we thought the drug product would be worse.
  - When testing at the 5% level with 80% power, the sample size required decreases by about 27% when a one-sided test is used.
- Regulatory agency
  - One-sided test allows more bad drug products to be approved because of chance as compared to two-sided test.

One-sided Test vs. Two-sided Test

	One-sided test	Two-sided test
Hypotheses	**Superiority**	**Equality**
One trial	**1/20**	**1/40**
Two trials	**1/400**	**1/1600**
Academia	**Yes/No**	**Yes**
FDA		**Yes**
Industry	**Yes**	

One-sided Test vs. Two-sided Test

- Situation where one-sided test may be appropriate
  - There is truly only concern with outcomes in one tail
  - It is completely inconceivable that results could go in the opposite direction
- Situation where one-sided test may be justified
  - Toxicity studies
  - Safety evaluation
  - Analysis of occurrences of adverse drug reaction data
  - Risk evaluation
  - Laboratory data

Basic Design Considerations

- Patient Selection
- Randomization
- Blinding
- The Selection of Controls
- Sample Size Calculation

Patient Selection

- CFR 314.166
  - Provides assurance that the selected patients have the disease/condition being studied
- Define patient population
  - Eligibility criteria
- Select a representative sample from the patient population
  - Accuracy
  - Reliability
  - Generalization of the findings

Eligibility Criteria

- Inclusion Criteria
  - Patients must meet all of the criteria.
  - To roughly outline the intended patient population
- Exclusion Criteria
  - Patients must not have any of the criteria.
  - To fine-tune the intended patient population by removing the unexpected sources of variabilities

Eligibility Criteria

- Sources of variabilities
  - Expected and controllable
  - Expected but not controllable
  - Unexpected but controllable
  - Unexpected and not controllable
- Statistical inference
  - Validity
  - Reliability

Eligibility Criteria

- Medical History
- Patient Characteristics
- Diagnostic Criteria
- Treatment Duration
- Severity of the Disease
- Others

Randomization - Concept

- Assures that patients selected constitute a representative sample of the intended patient population

Statistical Inference

Probability Structure

Randomization

- No randomization, then no probability structure.
  No probability structure, then no statistical inference

Randomization - Purpose

- Comparable Groups with Similar Characteristics
  - Avoid bias
  - Control variability
- Valid Statistical Tests
  - Accuracy
  - Reliability
  - Integrity of the trial

Randomization - Regulatory Requirement

- CFR 314.166
  - Treatment must be randomly assigned to patients.
  - Random selection of a representative sample from a targeted patient population
  - Random assignment of patients

Randomization - Advantages

- Remove bias due to imbalance between groups.
  - Blocking
  - Stratification
- Avoid subjective assignment of treatments to patients.

Practical Issues

- How to determine block size?
  - Block size = 2?
  - Block size = 4?
  - Block size = 6?
- The impact of treatment imbalance on power?

Blinding

- Purpose

  To prevent bias caused by subjective judgment in
  - Reporting
  - Evaluation
  - Data Processing
  - Statistical Analysis

  due to the knowledge of the identity of the treatments

Types of Blinding

- Open-Label
- Single-Blind
- Double-Blind
- Triple-Blind

Open-Label

- Both the investigator and the patient have an idea about which treatment the patient receives
  - Psychological human bias
  - Does not affect analytical results
  - Not recommended for comparative clinical trials
  - Generally not accepted by the U.S. FDA as adequate well-controlled trials for providing substantial evidence
- Situations when open-label trials are appropriate
  - Ethical consideration
  - Pre-marketing/post-marketing surveillance
  - New surgical procedures

Single-Blind

- The patient is unaware of his/her treatment
- Offers a certain degree of assurance of the validity of the trial
- Investigator may bias the evaluation of the treatment
- Spilker (1991)
  - The results of single-blind trials are equivalent to those from open-label trials.

Double-Blind

- Neither the patients nor the investigator are aware of patient's treatment assignment
- Provide a fair and unbiased assessment of study medication
- Considered as adequate and well-controlled clinical trials for providing substantial evidence of the effectiveness and safety of the study medications

Triple-Blind

- In addition to the patients and the investigator, all members of clinical project team are also blinded.
- It ensures the integrity of the trial.
- It provides the most conclusive unbiased evidence for the evaluation of the effectiveness and safety of the study medications.

Practical Issue - Matching Placebo

A matching placebo should be identical to the active drug in all aspects of

- Size
- Color
- Coating
- Taste
- Texture
- Shape
- Order

except that it contains no active ingredient

Practical Issue - The Integrity of Blinding

- Patient's guess
- Investigator's guess
- Expected bias factor
- Example

	Actual Assignment	
Patient's Guess	A	P
A	40	11
P	12	39
Don't Know	49	39
Total	101	89

Selection of Controls

Comparison of the Results Between Uncontrolled and Controlled Trials

Therapeutic Areas	Percent of Positive Findings: Uncontrolled	Controlled
Psychiatric (Foulds, 1958)	83%	25%
Antidepressant (Wechsler et al., 1965)	57%	29%
Antidepressant (Smith, et al., 1969)	58%	33%
Respiratory distress syndrome (Sinclair, 1966)	89%	50%
Rheumatoid arthritis (O'Brien, 1968)	62%	25%

Source: Summarized and tabulated from Spilker (1991)

Selection of Controls

The U.S. FDA requires that adequate well-controlled studies use a design that permits a valid comparison with a control to provide a quantitative assessment of drug effect

- Section 314.126 in part 21 of CFR

Selection of Controls

- Concurrent control
  - Placebo concurrent control
  - Dose-comparison control
  - Active concurrent control
  - No treatment control
- Historical control

ICH E9 Expert Working Group

- ICH E9 Expert Working Group
  - U.S.
    - PhRMA: B. Louv, S. Ruberg
    - FDA: R. O'Neill, S. Ellenberg
  - Europe
    - EFPIA: B. Huitfeldt, T. Lewis
    - EU: J. Lewis, J. Röehmel
  - Japan
    - JPMA: T. Uwoi, H. Uesaka
    - MHW: I. Yoshimura, T. Sato

Intent of ICH E9

- Statement of fundamental principles relating to statistical methodology in assessing pharmaceutical products
- Primary focus on confirmatory studies
- Highlights role of statistician in development, implementation and interpretation of data from clinical trials

ICH E9 Expert Working Group

- Major Emphasis
  - Minimizing bias
  - Maximizing precision
- Robustness of Trial Results
  - Reproducibility
  - Generalizability
- Responsibility for Statistical Aspects of Trials
  - Appropriately qualified and experienced statistician
  - Implementation of good statistics practice in design, conduct, analysis, report, and review/approval of drug research and development

Statistical Issues - ICH E9

- Primary and Secondary Variables
- Multicenter Trials
- Equivalence Trials
- Sample Size
- Analysis Sets
- Missing and Outlying Values
- Data Transformations
- Use of P-Values

Primary and Secondary Variables

- Limit to a Single Primary Variable, if possible.
- When multiple measurements are relevant, consider developing composite variable.
- Multiple Primary Variables
  - Significance on at least one, some subset, all?
  - Effect on type I error rate
  - May consider correlation among variables in assessing type I and type II errors

Multicenter Studies

- Advantages
  - To expedite the patient enrollment process
  - To provide replication and generalizability of clinical results to the targeted patient population
- Concerns
  - What is center?
  - Similarity among centers?
  - How many centers should be used?
  - How to deal with treatment-by-center interaction?

Equivalence Trials

- Equally effective or equally ineffective
- Equivalence limits must be pre-specified and should be clinically justified
- Statistical approach focuses on use of confidence intervals

Sample Size

Controversial Sentence:

"The number of subjects in a clinical trial should always be large enough to provide a reliable answer to the questions addressed, but should also be the minimum necessary to achieve aim".

Intention-to-Treat

- Based on randomized patients
- An unbiased assessment of treatment effect
- Best reflect real clinical practice
- The method of last-observation-carried-forward (LOCF) is often applied.

Efficacy Subset

- All patients with any efficacy observations or with a certain minimum number of observations
- Only patients complete the study
- All patients with an observation during particular time window
- Only patients with a specified degree of compliance

…

Missing and Outlying Values

- A major concern with no universally accepted approach
- Sensitivity analyses may be worthwhile
- Influence of outliers should be explored

Data Transformations

- Analysis of transformed data is acceptable
- Should pre-specify the intent for analyzing transformed data
- Should consider clinical interpretation

Use of P-Values

- Accompany with confidence intervals
- Report precise values rather than simply "$p<0.05$"
- Clarify one-sided vs. two-sided and prospectively justify one-sided tests
- Adjustment of p-values for multiplicity should be considered when appropriate

Adjustment for Covariates (continued...)

- Concerns
  - Covariate imbalance may result in a biased estimate of treatment effect.
  - Treatment by covariate interaction
- Solutions?
  - Adjustment for covariates
  - Pre-study stratified randomization
  - Post-study subgroup analyses

Multiplicity

- Multiple Comparisons
  - Comparisons among more than two treatments
  - Comparisons of K doses of a test drug
- Multiple Endpoints
  - Primary
  - Secondary
- Subgroup Analyses
  - Demographic variables
  - Patient characteristics

Active Control Trials

- Ethical Consideration
  - Very ill patients
  - Patients with severe or life-threatening diseases
- Primary Objectives
  - To establish efficacy
  - To show superiority
  - To demonstrate equivalence

Active Control Trials

- Concerns
  - Does not provide direct evidence of efficacy
  - Equivalence may imply equally effective or equally ineffective.

Active Control Trials

- Case I: A>B
  - A>B>P - Both A&B are effective.
  - A>P>B - A effective but B is not effective.
  - P>A>B - Both A&B are not effective.
- Case II: A=B
  - A=B>P - Both A&B are effective.
  - P>A=B - Both A&B are not effective.

Active Control Trials

- The U.S. FDA prefers placebo-control trials.
- European Community resists to placebo-control trials.
- It is suggested that clinical trials including new drug, active control, and placebo be considered as an alternative.

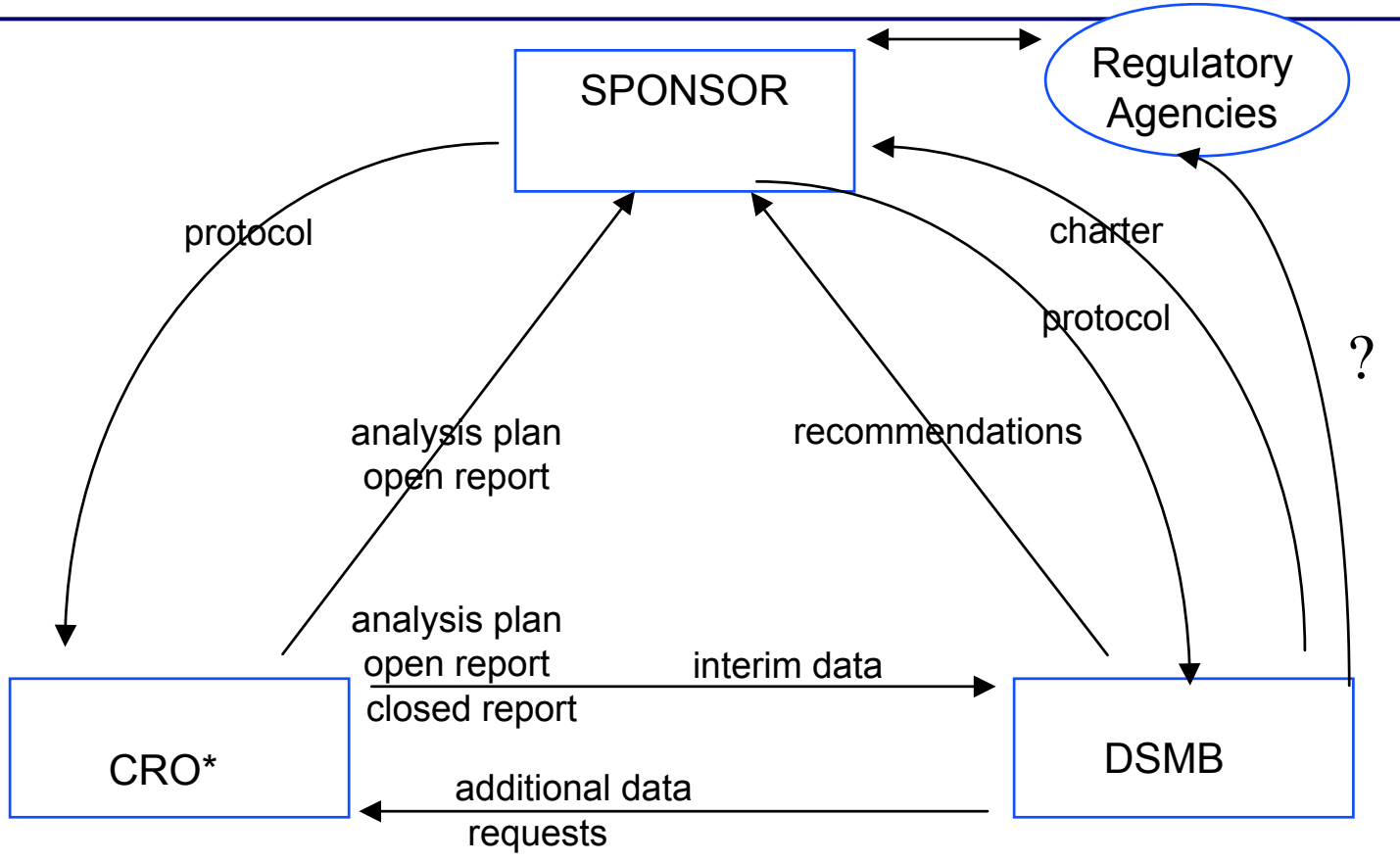
DSMB - Purpose

- Review of study endpoints
- Review of safety data
- Review of enrollment rate and placebo event rate
- Empowered to recommend expansion of sample size, early stopping of trial
- Opportunity to comment on data quality

DSMB - Characteristics

- Ongoing safety monitoring process
- Independent of any activities related to clinical operation of the study
- Comprised of experienced physicians and statisticians
- Separate DSMB staff support

DSMB - Organizational Flow



*Not otherwise involved with study

DSMB - Charter Outline

1. Introduction
2. Role of the Committee
3. Organizational Flow
4. Committee Membership
   - 4.1 Members
   - 4.2 Financial Disclosure
   - 4.3 Duration of DSMB Membership
5. Committee Meetings
   - 5.1 Organizational Meeting
   - 5.2 Scheduled Interim Analysis Meeting
   - 5.3 Unscheduled Meetings
6. Communication
   - 6.1 Open Reports
   - 6.2 Closed Reports
   - 6.3 Committee Minutes
   - 6.4 Committee Recommendations
   - 6.5 Sponsor Decision
   - 6.6 DSMB Additional Data Request
7. Timetable

DSMB - CRO's Role

- Assist with DSMB charter
- Arrange DSMB meetings
- Draft DSMB meeting agendas
- Prepare a procedure for unblinding
- Prepare a procedure for controlling the dissemination of the interim results
- Monitor validity, accuracy and reliability of database

DSMB - CRO's Role

- Perform interim analyses
- Prepare interim analysis reports
- Submit interim analysis reports to all DSMB members at least one week prior to the DSMB meeting
- Present interim analysis summaries at the DSMB meeting

DSMB - CRO's Role

- Perform additional analyses as requested by DSMB
- Draft minutes of the DSMB meetings
- Provide a chronology that documents all interim analyses, DSMB meeting minutes, DSMB recommendations, and any changes to the protocol.

DSMB - Project Team

- Project Team
  - Draft a report and analysis plan (RAP) including interim analysis section
  - Work with project programmer(s) to develop all outputs using dummy treatment codes
  - Coordinate a practice analysis (prior to the first DSMB meeting) using dummy codes
  - Submit draft outputs from the practice analysis to DSMB support statistician for review
  - Remain blinded throughout the study
  - Review and compile SAE reports.

DSMB - Support Staff

- Support Staff
  - Review the RAP and provide comments to the project team
  - Remain an independent function
  - Perform all unblinded and group analyses
  - Review all unblinded outputs
  - Alert DSMB to potential issues.