# Experimental Design and Statistical Analysis Seminar

Shah Golshan
sgolshan@Health.ucsd.edu
858-642-1264



1

## EXPERIMENTAL DESIGN AND STATISTICAL ANALYSIS SEMINAR:
### STUDY DESIGNS

1. Classification of Study Designs
2. Essential Design Features of a Controlled Clinical Trial
3. Quality Assurance
4. Computer Facility
5. Data Security Precautions
6. Preparation of Analysis Files
7. Broader classifications of study designs
8. Randomization
9. Activities By Stage Of Clinical Trial



64

## 3: Quality Assurance

- Quality assurance is any method or procedure for collecting, processing, or analyzing study data that is aimed at maintaining or enhancing their reliability or validity.
- Most of the quality assurance procedures require a continuous and timely flow of data from the clinic to the data center to be useful.
- The data edits and analyses carried out during the trial to assess data quality and clinic performance will lose much of their value if there is a large time gap between data generation and conversion into computer-readable formats.



65

## 3: Quality Assurance Procedures

1) Ongoing data cleaning process
2) Visual check for illegible responses and for unanswered or incorrectly answered items
3) Replication of the coding and data entry process as a means of error detection
4) Computer edit of keyed data for inadmissible codes or missing values
5) Data edit queries concerning completed data forms (Locking System)



66

## 3: Quality Assurance Procedures

6) Generation of periodic status reports concerning the data collection process
7) Multiple independent readings of ECGs, X-rays, etc. (inter-rater reliability).
8) Independent review of patient records
9) Submission of blinded duplicate specimens or records
10) Interim analyses for inadequacies or inconsistencies in the data collected



67

## EXPERIMENTAL DESIGN AND STATISTICAL ANALYSIS SEMINAR:
### STUDY DESIGNS

1. Classification of Study Designs
2. Essential Design Features of a Controlled Clinical Trial
3. Quality Assurance
4. **Computer Facility**
5. Data Security Precautions
6. Preparation of Analysis Files
7. Activities By Stage Of Clinical Trial



68

## 4: Computer Facility: Multiple levels of General-Use Facility

**A: Pros and Cons**

1. More computing power, but limited access to the facility
2. Investigators are freed of responsibility for the operation of the facility; however, less insensitive to specific needs of the trial
3. Larger number of programming options
4. Wider array of hardware
5. More difficult to protect the data files

**B: Factors in Their Favor**

1. Existence of good general-use facility
2. Total duration of the trial (e.g., ≤ 3 years)
3. Small programming and data processing staff need
4. Lack of appropriate staff locally



69

## 4: Computer Facility: Dedicated Facility

**A: Pros and Cons**

1. Limited computing power and number of hardware and software
2. Access to computer can be limited to study personnel (more resources are available)
3. Easier to protect data files against unauthorized entry **(Not True)**
4. Responsibility for operation of the facility rests with study personnel (Difficult task to do).

**B: Factors in Their Favor**

1. No general-use facility in the institution
2. Sizable data processing needs ( > 3 years)
3. Large programming and data processing staff need ( ≥ 4)
4. The existence of qualify staff



70

## 4: Computer Facility:

**Considerations in choosing among different general-use facilities:**

1. Type and number of staffing available
2. Hours of operation and modes of access
3. Primary use of the facility (research versus administration)
4. Compatibility of hardware/software with other facilities
5. Array of available hardware and software packages, particularly for data management (e.g., ACCESS, Oracle) and data analysis (e.g., SPSS, SAS)
6. Past history of operation and upgrades
7. Users level of satisfaction
8. Charging Fee: computer time, storage, printer, etc.



71

## 4: Computer Facility:

Considerations in choosing among different **dedicated facilities** are:

1. Available hardware and software features (Computing power, response time, database maintenance, and data analysis)
2. Compatibility of programming languages with other operating systems
3. Standard Operating Manuals
4. Expertise of service personnel
5. Levels of access to systems personnel
6. Cost and maintenance charges



72

## EXPERIMENTAL DESIGN AND STATISTICAL ANALYSIS SEMINAR:
### STUDY DESIGNS

1. Classification of Study Designs
2. Essential Design Features of a Controlled Clinical Trial
3. Quality Assurance
4. Computer Facility
5. **Data Security Precautions**
6. Preparation of Analysis Files
7. Activities By Stage Of Clinical Trial



73

## 5: Data Security Precautions

The study database must be safe-guarded against loss or unauthorized use.

**A: General Precautions & Safeguards**

**B: Patient Confidentiality Safeguards**

**C: Safeguards Against Misuse**

**D: Loss Safeguards**



74

## 5: Data Security Precautions:

### A: General Precautions & Safeguards

1. Staff training/experienced in the operation of a database and in protecting it against loss or misuse
2. Periodic staff meetings to remind database personnel of required operating procedures and safeguards



75

## 5: Data Security Precautions

The study database must be safe-guarded against loss or unauthorized use.

**A:** General Precautions & Safeguards

**B: Patient Confidentiality Safeguards**

**C: Safeguards Against Misuse**

**D: Loss Safeguards**



76

## 5: Data Security Precautions:

### B: Patient Confidentiality Safeguards

- HIPAA: In 1996, Congress passed the Health Insurance Portability and Accountability Act.
- Full implementation of the rule became effective April 14, 2003.
- Staff with access to sensitive information are required to complete annual training.
- VA sensitive information: Personally Identifiable Information (PII), Protected Health Information (PHI).



77

## 5: Data Security Precautions:

18 HIPAA Elements

1. Names
2. All geographical subdivisions smaller than a State
3. All elements of dates
4. Phone numbers
5. Fax numbers
6. Electronic mail addresses
7. Social Security numbers
8. Medical record numbers
9. Health plan beneficiary numbers
10. Account numbers
11. Certificate/license numbers
12. Vehicle identifiers, license and serial numbers
13. Device identifiers and serial numbers
14. Web Universal Resource Locators (URLs)
15. Internet Protocol (IP) address numbers
16. Biometric identifiers, including finger and voice prints
17. Full face photographic images and any comparable images
18. Any other unique identifying number, characteristic, or code (note this does not mean the unique code assigned by the investigator to code the data)



78

## 5: Data Security Precautions:

### B: Patient Confidentiality Safeguards

- All employees shall comply with all Federal laws, regulations, UCSD, VA, and VHA policies.
- De-identification = Removing all 18 HIPAA elements.
- De-identified information is not considered to be individually identifiable; therefore, the provisions of the Privacy Act, HIPAA, and UCSD/VA confidentiality statutes do not apply.
- PHI is compiled into a limited data set and only disclosed by using a data use agreement (DUA).



79

## 5: Data Security Precautions:

### B: Patient Confidentiality Safeguards

1. Data flow procedures should exclude patient identifying information
2. Physical separation of pages containing personal identifying information from other pages of the data forms
3. Electronic storage of patient identifying information in enciphered form or in a separate file
4. Separation of the file containing patient identifying information from other files



80

**5: Data Security Precautions:**

**B: Patient Confidentiality Safeguards**

5. Proscription against the distribution of data listings that contain Subject identifiers
6. Proscription against the use of any Subject identifiers in any published data listing
7. Secure procedure for disposing of computer output and backups that contain patient-identifying information
8. Denial of access to any Subject record to persons outside the research group without the express written consent of the patient



81

**5: Data Security Precautions**

The study database must be safe-guarded against loss or unauthorized use.

A: General Precautions & Safeguards

B: Patient Confidentiality Safeguards

**C: Safeguards Against Misuse**

D: Loss Safeguards



82

**5: Data Security Precautions:**

**C: Safeguards Against Misuse:**

1. **Data misuse** is the inappropriate use of data as defined when the data was initially collected.
2. Limit the number of persons who have access to the original study forms or any patient identifying information
3. Restrict access to the analysis computer files containing study results
4. Proscribe release of any data listing, without approval of the study leadership committee
5. Store completed study forms, data tapes/disks, in an attended, locked area



83

**5: Data Security Precautions**

The study database must be safe-guarded against loss or unauthorized use.

A: General Precautions & Safeguards

B: Patient Confidentiality Safeguards

C: Safeguards Against Misuse

**D: Loss Safeguards**



84

**5: Data Security Precautions:**

**D: Loss Safeguards:**

1. Scan original data forms for storage
2. Maintain duplicates of the electronic forms
3. AutoSave: Roll back option
4. Establish and maintain a series of data backups
5. Store copies of backup in an off-site location
6. Establish strict rules to safeguard access to backup systems
7. Make a backup of all essential programs



85

**EXPERIMENTAL DESIGN AND STATISTICAL ANALYSIS SEMINAR:**
**STUDY DESIGNS**

1. Classification of Study Designs
2. Essential Design Features of a Controlled Clinical Trial
3. Quality Assurance
4. Computer Facility
5. Data Security Precautions
6. Preparation of Analysis Files
7. **Broader classifications of study designs**
8. Randomization
9. Activities By Stage Of Clinical Trial



86