Contributors: Johnathan Cooper and Saaid Siddiqui

Document Outline

Epidemiology Section

Statistics Section

Epidemiology

I. Clinical Study Design

Clinical studies are either **observational** or **experimental**. Observational studies may be **descriptive**, which generate hypotheses for further studies or **analytic**, which test hypotheses and can report associations. Experimental studies are used to test hypotheses and report cause and effect relationships.

Longer descriptions of these types of tests can be found online. The following chart includes key words to help you remember what is important about each type of test and the level of evidence associated with each test is included for most.

Study Type	Key Words	Level of Evidence
Descriptive Observational Studies	**Generate Hypotheses**	
Case Report	- Signs, Symptoms, Diagnosis, Treatment, Follow up - Novel/Rare - Single case	- Level 5
Case Series	- Signs, Symptoms, Diagnosis, Treatment, Follow up - Novel/Rare - Multiple cases - Similar subjects/treatment	- Level 4
Correlation Study	- Association (OR,RR,R) - Large Samples - Disease—Risk factor	
Cross-sectional	- Moment in time - Prevalance	- Level 4

Study Type	Key Words	Level of Evidence
Systematic Review	- Literature summary - Basic statistics - Objective/Predetermined Inclusion Criteria	- Can be Level I, II, or III depending on the types of studies it is based on.
Analytic Observational Studies	**Test hypotheses Report Associations**	
Case Control	- Two Groups - Control/Diseased group - Reports OR for risk factor - Retrospective - High risk of recall bias	- Level 3
Interventional Study	- Clinical Interventions - Effect of intervention on outcome	
Meta-analysis	- Literature Analysis - Advanced Statistics - Objective/Predetermined inclusion criteria - High risk of publication bias	- Level 1 if Meta-analysis of randomized trials with homogenous results - Level 2 if Meta- analysis of level 2 studies - Level 3 if meta- analysis of level 3 studies
Outcomes Research	- Patient Centered - Benefits/Risks of treatment method - Patient individual situation	
Prospective Cohort	- Two Groups - Control/Exposed group - Routine documentation - Prospective	- Level 2
Experimental Study		

Study Type	Key Words	Level of Evidence
Randomized Controlled Trial (RCT)	- Random Treatment Assignment - Control Group - Prospective - Unmeasured covariates randomly distributed	- Level 1 if follow up > 80% - Level 2 if follow up < 80%

II. Levels of Evidence

The levels of evidence is/are a ranking system used in evidence based practices which determines the clinical value of a study. Keep in mind both the study design and the endpoints measured determine the level of evidence. Definitions to make the chart easier to understand are included below.

Level of Evidence	Therapeutic Studies	Prognostic Studies	Diagnostic Studies
Level I	- High quality RCT - Systematic review of level 1 RCT w/ homogenous results	- High quality prospective study - Systematic review of level 1 studies	- Testing of previously developed diagnostic criteria on consecutive patients (w/universally applied gold standard) - Systematic review of level I studies
Level II	- Lesser quality RCT - Prospective Cohort Study - Systematic review of level II studies w/ heterogenous results	- Retrospective study - Untreated controls from RCT - Lesser quality prospective study - Systematic review of level II studies	- Development of previously developed diagnostic criteria on consecutive patients (w/universally applied gold standard) - Systematic review of level II studies
Level III	- Case control study - Retrospective cohort study - Systematic review of level III studies	- Case control study	- Study of non-consecutive patients without universally applied reference gold standard - Systematic review of level III studies

Level of Evidence	Therapeutic Studies	Prognostic Studies	Diagnostic Studies
Level IV	- Case Series	- Case series	- Case control - Poor reference standard
Level V	- Case Study - Expert Opinion	- Case study - Expert opinion	- Expert opinion

Term	Definition
Therapeutic Studies	Treatment under investigation is believed to be beneficial to participants in some way.
Prognostic Studies	Examine selected predictive variables or risk factors and assess their influence on the outcome of a disease.
Diagnostic Studies	Procedure performed to confirm or determine the presence of disease in an individual suspected of having it, following the report of symptoms or based on other tests.
High quality RCT	Require at least 80% follow up, proper blinding, and properly random treatment assignment.
Prospective Study	Study began prior to initial patient enrollment.
Retrospective Study	Study began after initial patient enrollment.
High quality prospective study	Require all patients to be enrolled at the sme point in their disease with at least 80% follow up of enrolled patients.
Universally applied gold standard	Currently accepted diagnostic critera (ex. X-ray for bone fracture)

III. Sources of Bias

Bias and Study Errors

There are more types of bias than those described in this booklet, but the ones included are quite common and are important to be familiar with.

Types of Bias When Recruiting Participants

Selection Bias

Four Characteristics of Selection Bias:

1. Nonrandom sampling or assignment to treatment.
2. The sample does not effectively represent the population of interest.
3. Patients are lost in follow up.
4. The study produces a different results than expected if the study included the entire target population.

How to Reduce the Chances of Selection Bias:

Ensure randomization during:

1. Sampling: Randomly sample from population or smaller defined groups (strata) of the population (ex. When sampling average length of sleep for high school students, sample from jocks, cool kids, mathletes, cheerleaders, etc.).

2. Assignment to Treatment: Assignment of subjects or smaller defined groups of subjects to treatment groups must be random.

Types of Bias When Performing the Study

Recall Bias

Characteristics of Recall Bias:

1. Sample population self-reports data. (ex. As a child, did your house have lead paint?)
2. One group either intentienally or unintentially misremembers a piece of information about exposure to a risk factor due to their being in the treatment group. (ex. Response: My house did not have lead paint — but it actually did.)
3. Because the subjects misremembered their exposure or non-exposure, they are incorrectly sorted into the control group or treatment group. (ex. Control group did not have lead paint in their houses as children, Treatment group did have lead paint in their houses — This particular subject is incorrectly sorted into the control group.)

Risk Factors for Recall Bias:

1. The disease or event in question is significant or critical (ex. cancer)
2. A particular exposure is thought of by the patient as a risk factor for a high burden disease.
3. A scientifically ill-established association is made public by the media.
4. The exposure under investigation is socially undesirable (ex. AIDS).
5. The event in question took place a long time ago.

How to Reduce Recall Bias:

1. Use a well constructed, standardized questionnaire.
2. Use a double-blind study: blind subjects and data collectors to the hypothesis of the study.

3. Use any available proxy sources of reported data to confirm results. (ex. tree ring width to measure historic rainfall, you can ask family members about a patient's QOL vs the patient's own description of their QOL)
4. If studying a disease or condition, choose participants with a new diagnosis when possible.

Measurement Bias

How to Reduce the Chances of Measurement Bias:

1. It results from a systematic error (an error that is consistently repeated, and that results in a specific favored outcome.) ex. a scale isn't properly tared and consistently overestimates the weight of objects by 2 lbs

2. Information is measured so that the true value of the information is obscured. (ex. a person is colorblind and is asked the color of 100 blue balls, and they say the balls are red.)

How to Reduce the Chances of Measurement Bias:

1. Use a predetermined standardized method of data collection compared against clinical assessment.
2. A placebo group is useful because the result is expected to be null, finding a measurement that is consistently higher than expected may indicate measurement bias.
3. A measurement may be validated by its ability to predict future illness
4. Use a reliable standard to confirm validity of findings

Procedure Bias

Characteristics of Procedure Bias:

1. Non-random treatment assignment — Patients and/or physicians responsible for treatment assignment.
2. Subjects in different groups are treated differently.

How to Reduce the Chances of Procedure Bias:

1. Random treatment assignment.
2. Double blind- blinding of patients and physicians to treatment.

Observer-expectancy Bias

Characteristics of Observer-expectancy bias:

1. Researcher shares their expectations for the outcome of the study with the subjects.
2. We see what we expect to see — We feel how we expect to feel, subjects act in accordance with the researchers expectations

How to Reduce the Chances of Observer-expectancy Bias:

1. Double blind- blinding of patients and physicians to treatment: Without expectations of what might occur, this form of bias is very unlikely

Examples of observer-expectancy bias:

1. Two groups observe a painting and are asked to rank it from 1-10, 1 being the ugly and 10 being the most beautiful. One group is told the painting is beautiful. The other group is told the painting is ugly. The group that is told the painting is beautiful has higher average scores than the group that was told the painting is ugly.
2. Researcher expects group who takes experimental sleep inducing drug to be more lethargic than placebo group — more likely to document fewer body movements in treatment group.

Interpreting Results

Confounding Bias

Confounding Variable: A variable other than the independent variable(s) that has an impact on the dependent variable.

Characteristics of Confounding Bias:

1. An relationship exists between the confounding variable and the outcome that is independent of the exposure.
2. The confounding variable is not a proxy for the exposure, but is associated with the exposure.
3. A confoundint variable is not an intermeditate between the exposure and the outcome.

How to Reduce the Chances of Confounding Bias:

1. Measure and report all potential confounding variables including diagnostic features, comorbidities, and any factor that may impact patient outcome
2. Routinely assess the role of confounding factors and adjust for them in analyses

- Restriction: Incusion criteria prevents confounders- if age is a confounder, set age boundaries for subjects (ex. 28-34 yrs old) OR stratification
- Multivariate analysis: Allows for adjustment of multiple variables simultaneously via mathematical modeling, mathematical controls

3. Report adjusted and crude estimates of association and discuss limitations

- If adjusted estimate is greater than or equal to 10% of the crude estimate, the variable can be considered a confounder

Lead-time Bias

Lead-time: The length of time between the detection of a disease and its diagnosis.

Characteristics of lead-time bias:

1. Disease is diagnosed earlier than usual, typically due to a novel screening method.
2. Early treatment often allows for earlier and more treatment than usual .
3. Disease runs its regular course but due to the early diagnosis, it is believed that the survival time has increased due to the extra treatment.

How to Reduce the Chances of Lead-time Bias:

1. Evalate severity of disease at time of diagnosis.
2. Compare survival times from different stages of the disease rather than survival times from diagnosis.
3. Measure "back-end" survival (adjust survival according to the severity of disease at the time of diagnosis).

IV. Measures of Disease

Term	Definition
Inflow	Proportion of people developing a condition over a period of time.
Pool	Total number of cases at a period of time.
Incidence	A measure of the number of new cases of a characteristic (such as illness or risk factor) that arise in a population over a given period.
Prevalance	The proportion of a population who have (or had) a specific characteristic in a given time period.

V. Odds Ratio and Relative Risk

Odds: The ratio of the probability of one event to that of an alternative event.

Odds Ratio (OR): The odds that an outcome will occur given a particular exposure, compared to the odds of the outcome occurring in the absence of that exposure.

Why is the Odds Ratio Important? The odds ratio is used to quantify how strongly the presence or absence of property A is associated with the presence or absence of property B in a given population.

$$\text{Odds Ratio} = \frac{\text{Presence of A/Absence of A}}{\text{Presence of B/ Absence of B}}$$

How to Interpret the Odds Ratio

OR > 1: A is "associated" with B, having B **increases** the chances of having B

OR = 1: B **does not affect** the chances of having A

OR < 1: A is "associated" with B, having B **reduces** the chances of having B

The OR is commonly reported in case control studies which determine the association between risk factors and developing a disease or sustaining an injury. Relative risk (RR)

and absolute risk reduction (ARR) are given in prospective studies (prospective cohort, clinical trials) as measures of association.

Risk: The probability that an event will occur within a stated period of time. Because risk is a probability, it lies between 0% and 100%.

$$\text{Risk} = \frac{\text{\# with outcome}}{\text{\# at risk of outcome}}$$

Absolute Risk: The total amount of risk of a given 'thing' occurring after all risk factors and confounding variables are summed up.

Relative Risk (RR): Also known as "Risk Ratio", the relative risk is a measure of the risk of a specified event occuring in one group compared to the risk of its occuring in another group.

$$\text{Relative Risk} = \frac{\text{Risk of event in treatment group}}{\text{Risk of event in control group}}$$

How to Interpret Relative Risk

RR > 1: exposure variable **increases** the risk of the outcome developing

RR < 1: exposure variable **decreases** the risk of the outcome developing

Absolute Risk Reduction (ARR): Also known as risk difference (RD), absolute risk reduction is the total percent reduction in risk that results from a given treatment.

Number Needed to Treat (NNT): The number of patients who need to be subjected to a treatment for that treatment to successfully treat one patient.

I. Basic Statistics

Term	Definition
Types of Variables	
Continuous Numerical	The variable include all real numbers. (1.12343, 2.4324, etc.)
Discrete Numerical	The variable include only whole numbers. (1, 2, 3, 4)
Ordinal Categorical	Natural order to the levels of the variables (Shortest- Average Height- Tallest)
Nominal Categorical	No natural order to the levels of the variables (US States)
Measures of Central Tendency	
Mean	Given as the (Sum of values/ Number of cases), the mean is a better estimate of the center of the data for normal data.
Median	Defined as the value at the midpoint of a distribution, the median is a more reliable estimate of the center of the data for skewed data.
Mode	The value which occurs most frequently in a distribution.
Measures of Dispersion	
Variance	The average of the squared departures from the mean.
Standard Deviation	Better estimate of variability for normal data

Term	Definition
Standard Error	More reliable estimate of the variabilty for non-normal/ skewed data
Statistical Hypotheses	
Null Hypothesis H_0	Hypothesis of no difference: population values are not significantly different.
Alternative Hypothesis H_a	Hypothesis of some difference: population values are significantly different.
Hypothesis Testing	
Significance Level	The probability of rejecting the null hypothesis when it is true. A .05, 0.01, or 0.001 significance level is common. The null hypothesis is rejected if the p-value is less than the significance level.
Confidence Interval	A range of values so defined that there is a specified probability that the value of a parameter lies within it.
95% confidence interval	If a population was sampled many times, a confidence interval drawn from 95% of those samples will contain the true population parameter (mean).
p-value	The probability, under the null hypothesis , of obtaining a result at least as extreme as the result obtained.
p-value < 0.05	There is less than a 5% chance of obtaining a result at least this extreme, given the null hypothesis is true.
Correlation coefficient, r	Measure of association between the independent and dependent variables.
Coefficient of determination, r^2	Percentage of variability in the dependent variable accounted for by independent variable.

II. Hypothesis Testing

Test	Description	Conditions
Linear Regression	- Used to study the linear relationship between a continuous dependent variable and one or more continuous independent variables - Can be used to predict a value of the dependent variable given some value of the independent variable	- The errors are normally distributed and are independent
One-sample t-test	- Compares the sample mean to a known mean. (ex. is the rainfall observed this year in Toledo significantly different than normal?)	- Sample size at least 30 if abnormally distributed - Random sample - Sample is less than 10% of the population
Two-sample unpaired t-test	- Used to compare the mean responses between two independent groups - H_0: responses are the same for both groups - H_a: responses are significantly different	- Random sample or random treatment assignment - n_{1+} n_2 at least 30 if sample is abnormally distributed - Sample is less than 10% of the population

Test	Description	Conditions
Two-sample paired t-test	- Used to compare the mean responses of groups of individuals who experiences both conditions of the variable of interest.	- Paired differences are normally distributed or there are at least 30 differences - Sample of paired differences is random
ANOVA	- Used to compare the means of 2 or more groups. - Categorical predictor(s), numerical response - H0: mean is the same between all groups - Ha: mean is different between 2 or more groups	- The errors are normally distributed and the errors are independent.
Chi-square	- Tests the association between 2 categorical variables - H0: No association - Ha: Some association - Compares the observed frequencies with the frequencies that would be expected if the null hypothesis of no association was true. - By assuming the variables are independent, we can predict an expected frequency for each cell in the contingency table	- Independence - Sample size/distribution: Each particular scenario (Cell count) must have at least 5 expected cases OR no more than 20% of the cells has an expected frequency less than 5 and no empty cells
Fisher's exact test	- Used the same way as chi-square, but functions even when chi-square conditions are not met. - 2 categorical independent variables tested against 2 categorical dependent variables	- Observations are independent - The row and column totals are fixed

Test	Description	Conditions
Kruskall Wallace Test	- If ANOVA conditions are not met, this test compares the median(s) of 2 or more groups of categorical variable(s)	- Observations in each group come from populations with the same shape of distribution
Logistic Regression	- Binary dependent variable - Continuous and/or categorical independent variable(s) - Used to determine OR - The empirical logit plot can be used to determine the OR for any level of the independent variable(s).	- Errors are independent - Independent variables are linearly related to the log odds - Sample size >10 per independent variable

III. Power

Power: The power of a study is the probability that one will correctly reject the null hypothesis if the alternative hypothesis is actually true.

Why is Power Important?

- Power is used to decide before initiation of a clinical study to descide whether it is worthwhile given the cost, effort, and patient involvement. Often times to have enough patients to run a high power study, you need a lot of money and many patients.
- A hypothesis test with little power will likely yield large p values and large confidence intervals. Thus when the power of a proposed study is low, even when there are real differences between treatments under investigation, the most likely result of the study will be that there is not enough evidence to reject the H0 and meaningful clinical differences will remain in question.

The following terms are commonly used when discussing power.

Term	Definition
Type 1 error	- Rejection of a true null hypothesis - False positive
Type 2 error	- Failing to reject a false null hypothesis - False negative
Sensitivity	- the ability of a test to correctly identify those with the disease (true positive rate) - true positives/(true positive + false negative)
Specificity	- the ability of the test to correctly identify those without the disease (true negative rate). - true negatives/(true negative + false positives)
Positive Predictive Value	- If the test result is positive what is the probability that the patient actually has the disease?" - PV+= true positive/(true positive + false positive)
Negative Predictive Value	- If the test result is negative what is the probability that the patient does not have disease?" - PV-= true negatives/(true negatives +false negatives)

IV. Miscellaneous

Likelihood Ratio

- Defined as the probability of a person who has the disease testing positive divided by the probability of a person who does not have the disease testing positive.
- P(T|D) / P(T|D') which is = Sensitivity/ (1 - Specificity), since Specificity = P(T'|D') and Sensitivity = P(T|D)
- Likelihood ratio positive = sensitivity / (1 − specificity)

Accuracy vs. Precision

Accuracy

- The difference between the measurement and the actual value.

Precision

- The variation between repeated measurements when using the same device.

Validity vs. Reliability

Validity:

- Measuring what is intended to be measured.
- There is low nonrandom (systematic) errors.
- It is assessed by one of three methods: content validation, criterion-related validation, and construct validation.

Reliability

- Consistency in measurement over repeated measures.
- There is a low random (chance) error.
- It is assessed by one of four methods: retest, alternative-form test, split-halves test, or internal consistency test.