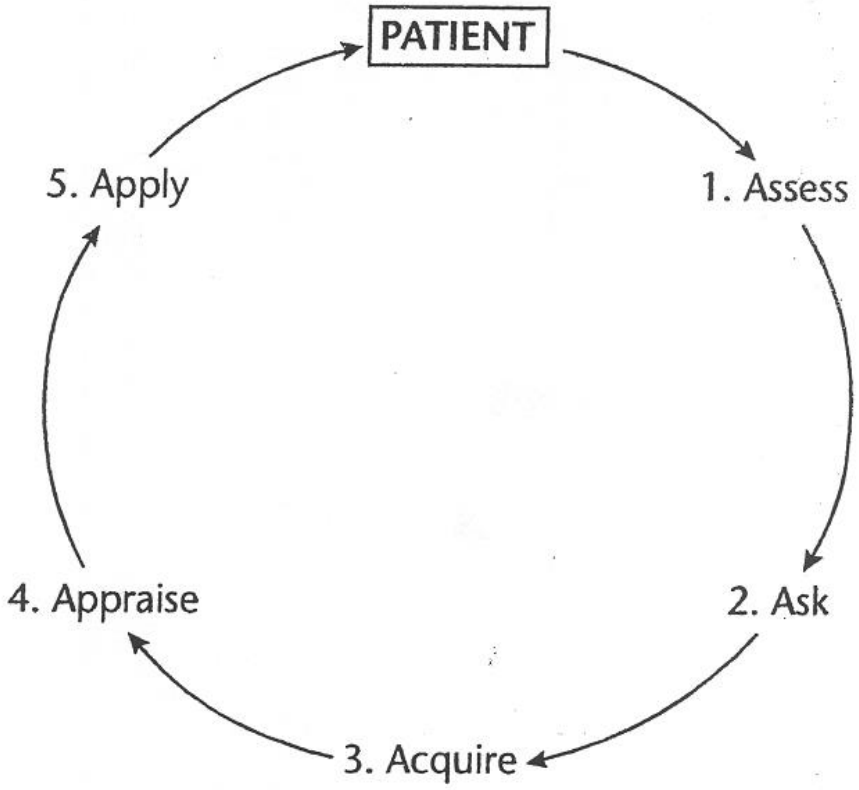
Evidence-Based Medicine Course: Therapy Article Review

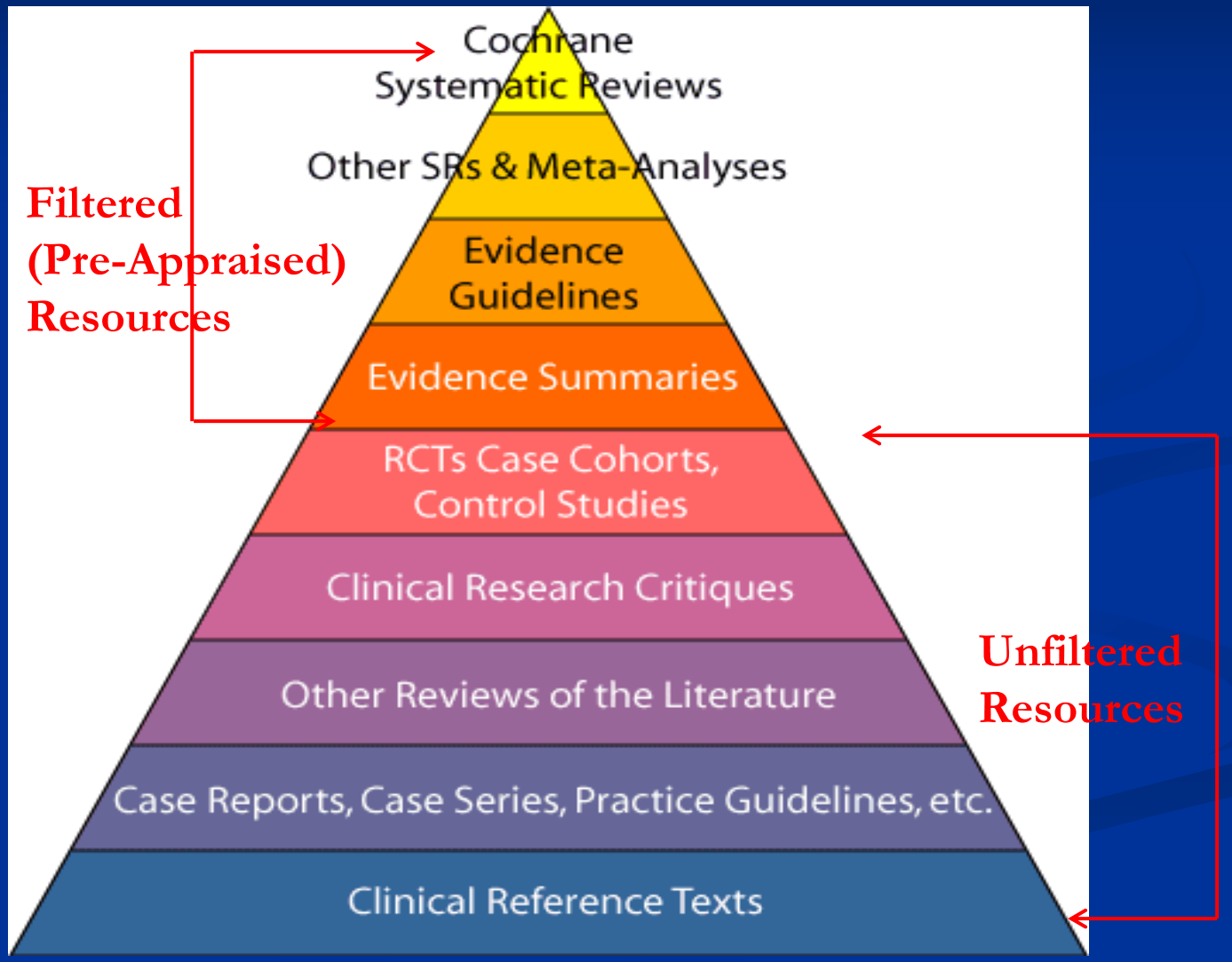
Des Moines Area

Medical Education Consortium

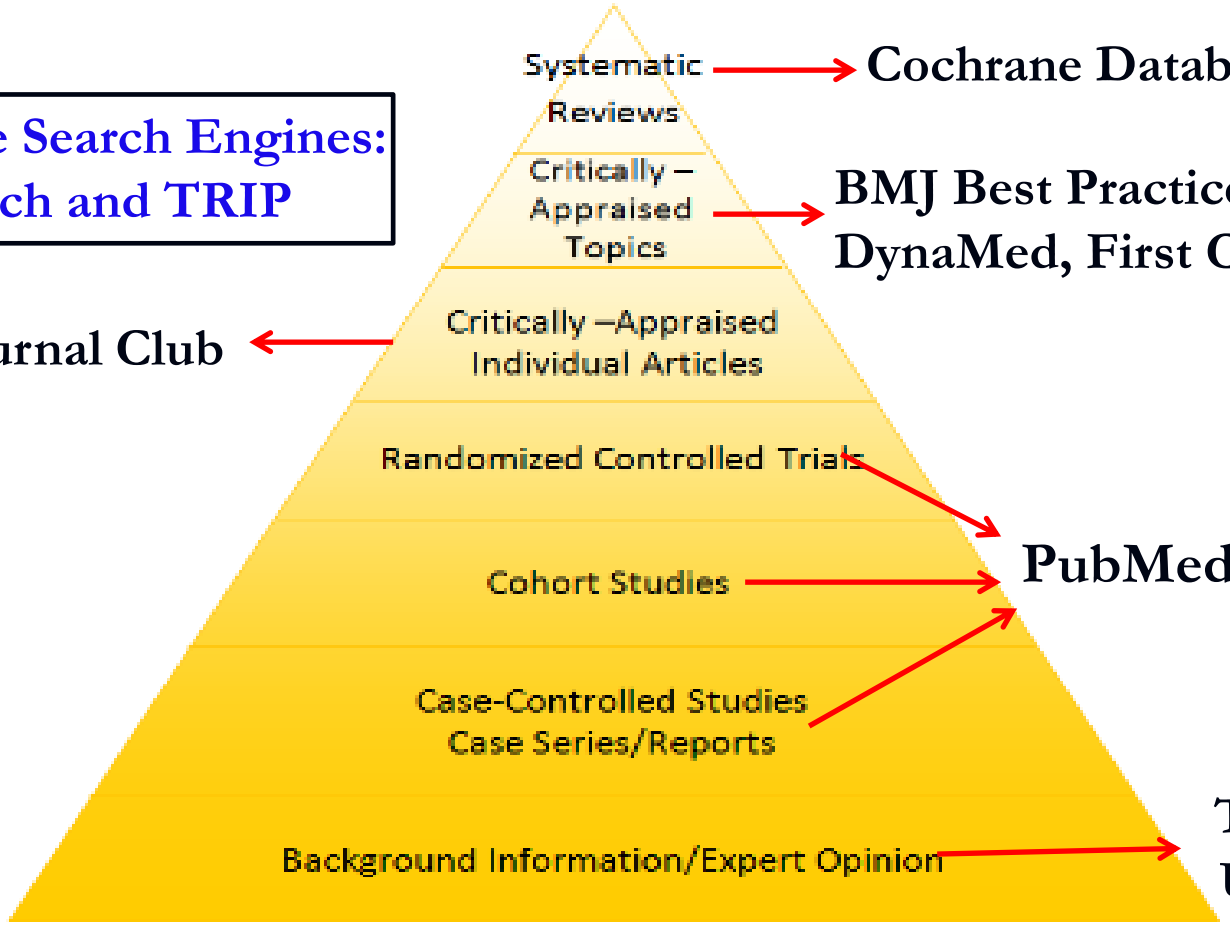
Evidence-Based Medicine



Results of Searching the Literature: Levels of Evidence



Evidence Pyramid



Multi-Source Search Engines: SUMSearch and TRIP

Cochrane Database

BMJ Best Practice, DynaMed, First Consult

ACP Journal Club

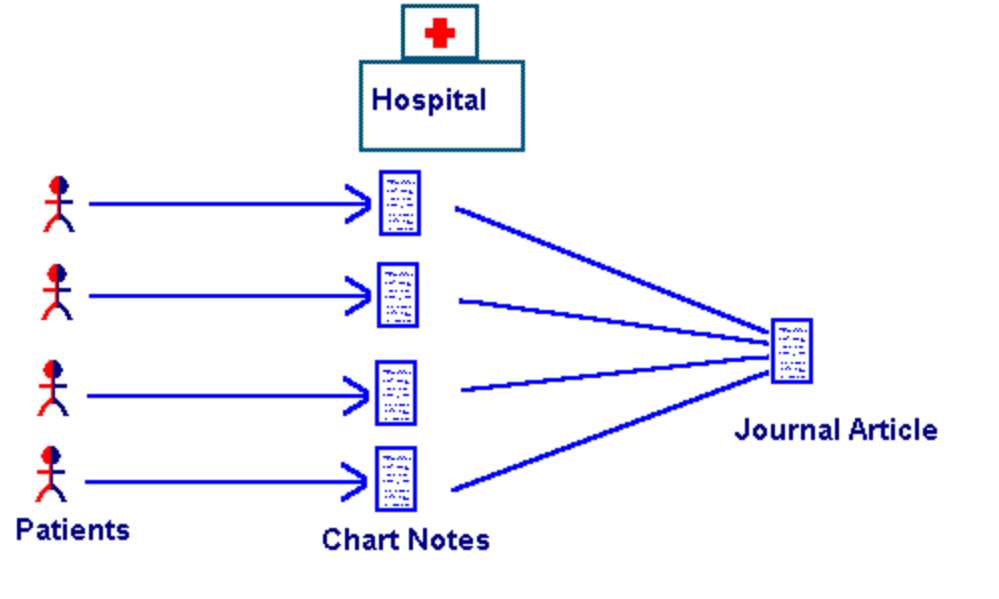
PubMed

Textbooks, UpToDate

The Hierarchy of Research Study Designs

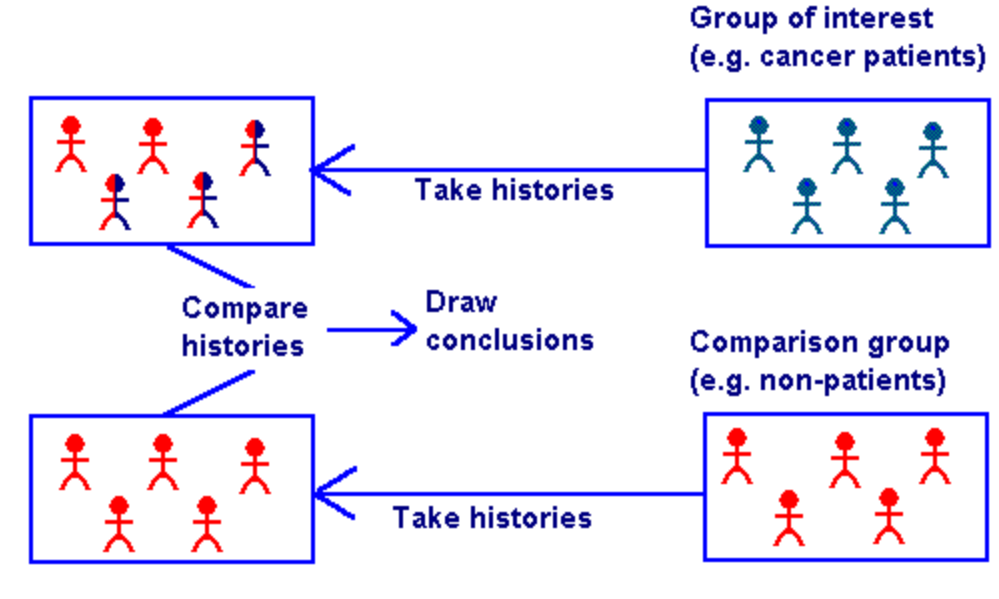
- Systematic Reviews
- Meta-Analyses
- Evidence Summaries & Evidence Guidelines
- Randomized Controlled Trials
- Cohort Studies
- Case Control Studies
- Case Reports & Case Series

Case Reports and Case Series



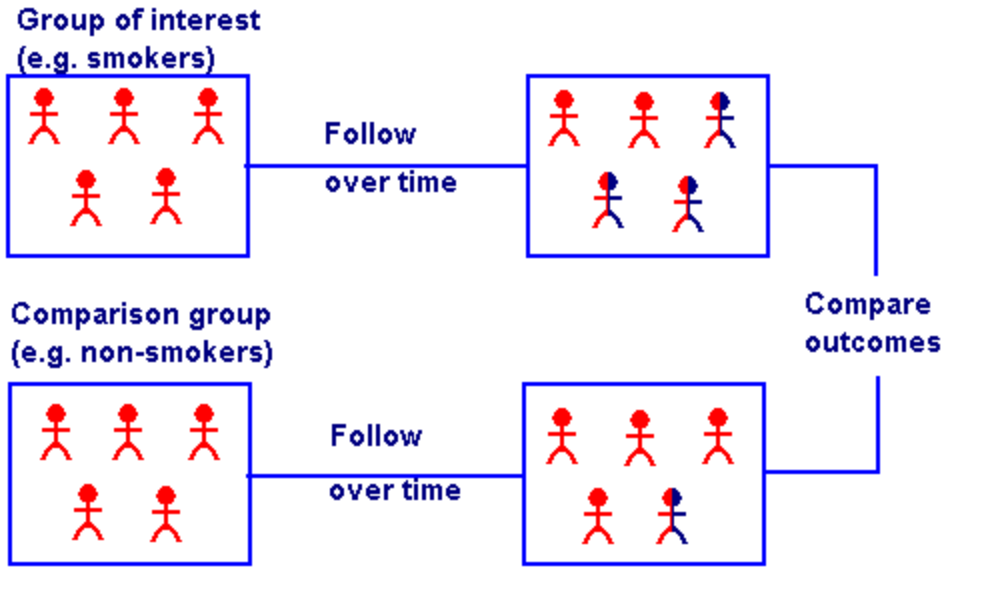
- Report on a single patient or several patients with the same condition
- Used to clarify characteristics of the condition, treatment effects, adverse effects of treatment, etc.
- Most helpful with uncommon conditions
- No control group & no statistical validity
- Can be written up in short period of time

Case Control Studies



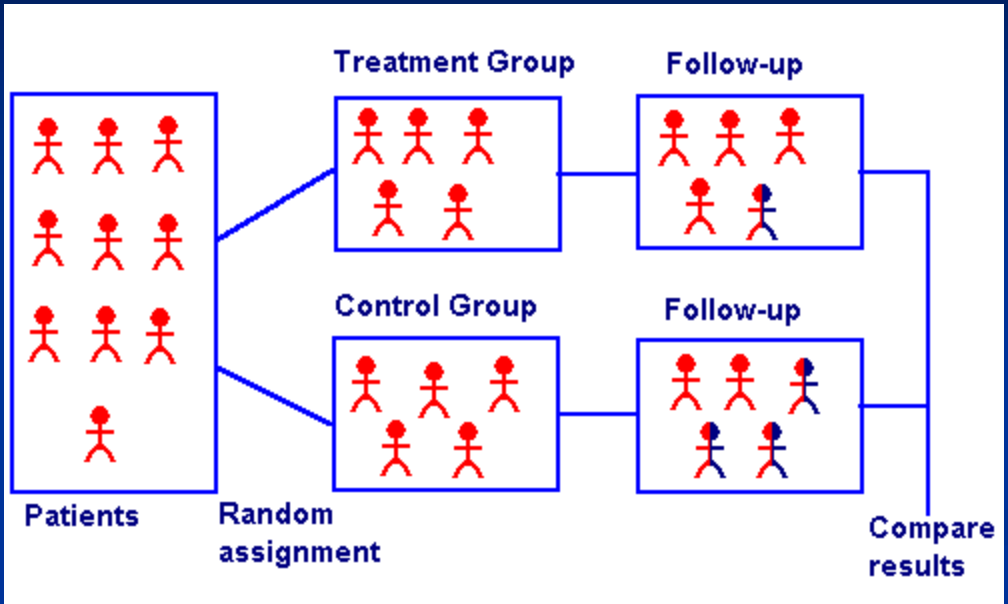
- Patients who already have a certain condition or treatment are compared with people who do not
- Try to draw conclusions from observations over time
- Often used to estimate odds of developing the condition being studied
- Can help determine if there is an association between a risk factor and the condition but can't establish absolute risk

Cohort Studies



- Longitudinal study following patients with a certain exposure or treatment over time
- Can compare to another group of patients not effected by the exposure or treatment under study
- May be either prospective or historical/retrospective
- Used to establish causation of a disease or evaluate the impact of a treatment when RCTs not possible
- Generally require large sample size and long follow-up period

Randomized Controlled Studies



- Gold standard in research
- Best at answering treatment questions
- Randomization avoids bias in the choice of patients receiving a given treatment
- Double blinding further reduces bias (minimizes the placebo effect)

Evidence Guidelines & Evidence Summaries

- Guidelines/Summaries generated by expert panel who together critically review available literature
- Must consider source & potential for bias of panel
- Must review methods used to search out available literature
- Best when controversies in literature re how best to diagnose/treat a condition

Meta-Analysis Studies

- Systematic, objective way of combining data from many studies
- Allows a pooled estimate of treatment effectiveness & stronger statistical significance of results
- Problems include publication bias & varying quality of studies from which data is extracted

Systematic Review Studies

- Comprehensive survey of a topic to include all relevant high level studies
- Assess all studies, synthesize the findings and present a balanced summary of the findings
- Especially good for evaluation of new technologies & new treatments
- Can include both published and unpublished studies
- More rigorous & less bias than a literature review

Practice Exercise

To determine if fasting is associated with dengue fever, data from 40 patients with dengue fever were collected. These patients were matched for age, sex, and race to 40 patients without dengue fever. The hospital charts of these patients were then reviewed to determine whether they also fasted prior to their illness. This study type is known as:

- Case series
- Concurrent cohort study
- Case-control study
- Retrospective cohort study
- Randomized clinical trial

Treatment Decisions

Always look for the highest level evidence to determine the best treatment for your patient.

So… when there is no Systematic Review or Meta-Analysis, or reliable Evidence Guideline or Summary to guide you, do a literature search looking for recent RCTs or Cohort Studies to help you plan your treatment.

Then… Do a critical appraisal of these therapy studies.

So… How do you critically appraise a therapy study?

Critically Appraising a Therapy Study

Steve Craig, M.D.

Three Basic Questions for Evaluating a Published Study

1. Are the results of the study valid?
2. What are the study results?
3. Will the study results help me in caring for my patients?

Are the results of the study valid?

- An unbiased estimate of the treatment effect

vs.

- Influenced (biased) in some systematic fashion

What are the study results?

- Must first establish significant benefit of treatment
- Then consider the size and precision of the treatment benefit
- Remember: Precision is superior in larger studies

Will the study results help me in caring for my patients?

- Are the results applicable to my patients? (inclusion / exclusion criteria)
- What is the net impact of the treatment? (risk-benefit ratio)

Therapy Study

1. Are the study results valid?

- **Primary Guides**: Can be easily applied by readers with limited time

- **Secondary Guides**: Reserved for articles that meet the 1° guides + when reader has time and/or need for more detailed review

Therapy Study

1. Are the study results valid?

PRIMARY GUIDES:

1. Was allocation of patients properly concealed?
2. Was assignment of patients randomized?
3. Were all patients accounted for and attributed at conclusion of study?
    - * Study drop outs
    - * Patients lost to follow-up
    - * Follow Intention-to-Treat

Therapy Study

1. Are the study results valid?

SECONDARY GUIDES:

1. Study blinded?
2. Control & treatment groups same at entry?
3. Control & treatment groups treated equally?
4. Study funding / potential for bias?
5. Statistical analysis: Power Analysis done?
   * **Sample size adequate?** (Larger sample needed if higher power or aim to detect smaller differences)

* ***Power Analysis needed when trial results negative***

Therapy Study

TYPES OF DATA REPORTED IN THERAPY STUDIES

- Parametric data = measured data (normally distributed quantitative data) reported as Mean ± SEM
- Non-Parametric **Nominal** data = categorical data reported as Risk Ratios, Relative Risks, Odds Ratios, Likelihood Ratios with 95% Confidence Intervals
- Non-Parametric **Ordinal** data = rating, ranking, scoring data reported as Median + Range

Types of Data Reported in Therapy Studies

1. Parametric Data

e.g. Compare 2 Cholesterol-Lowering Medications After One Year Treatment

Drug A: Cholesterol 190 ± 12 mg% (Mean ± SEM)

Drug B: Cholesterol 165 ± 10 mg% (Mean ± SEM)

---

2. Nominal (Categorical) Data: Dichotomous Data

e.g. Study Effect of Drug A vs. Drug B in Reducing Risk for Recurrent MI

Risk Ratio for Recurrent MI with Drug B vs. Drug A

= 0.66 (95% C.I., 0.60 – 0.75)

---

3. Ordinal (Rating, Ranking, Scoring) Data

e.g. Study Effect of 2 Different Medications to Treat Alzheimer's Dz on MMSE Scores

Enroll Patients with Mild Dementia and baseline MMSE scores 24-26 range

Drug A Group: MMSE Score at 1 Year = 22 (25^{th}-75^{th}% range, 20-24)

Drug B Group: MMSE Score at 1 Year = 17 (25^{th}-75^{th}% range, 12-19)

Practice Exercise

- Review the Parametric vs. Non-Parametric Data Worksheet

- For the two studies described, determine if the type of data being collected is:
    * Parametric Data
    * Non-Parametric Nominal Data
    * Non-Parametric Ordinal Data

Therapy Study

2. What are the study results?

a. Was the treatment benefit proven to a $p < 0.05$ level?

b. Was the treatment benefit large?

c. Was the treatment benefit shown to be precise?

2a. Was the treatment proven beneficial to a p <0.05 level?

Comparing Treatment Groups = Hypothesis Testing. Involves use of p values.

Hypothesis Testing

- Null Hypothesis = There is no difference between groups
- p value = Measure of the strength of the evidence in favor of the null hypothesis
- $p < 0.05$ = enough evidence against the null hypothesis to conclude there is a statistically significant difference between groups

Determining Significance of the Treatment Effect

1. Parametric Data

e.g. Compare 2 Cholesterol-Lowering Medications After One Year Treatment

If SEM don't overlap, the results will be significant at the $p < 0.05$ level

Drug A: Cholesterol 190 $\pm$ 12 mg% (178-202)

Drug B: Cholesterol 165 $\pm$ 10 mg% (155-175)

} **No Overlap**

---

2. Nominal (Categorical) Data

e.g. Study Effect of Drug A vs. Drug B in Reducing Risk for Recurrent MI

If 95% Confidence Interval for Odds Ratio or Risk Ratio doesn't include 1, results will be significant at the $p < 0.05$ level

Risk Ratio for Recurrent MI with Drug B vs. Drug A

= 0.66 (95% C.I., 0.60 – 0.75) } **0.60 - 0.75 does not cross 1**

---

3. Ordinal (Rating, Ranking, Scoring) Data

e.g. Study Effect of 2 Different Medications to Treat Alzheimer's Dz on MMSE Scores

Scores presented as Medians + Ranges: No easy way to predict if results are significant

Drug A: MMSE Score at 1 Year = 22 (25-75% range, 20-24)

Drug B: MMSE Score at 1 Year = 17 (25-75% range, 12-19)

Parametric Data: Significance Testing

If for 2 means, the SEM do not overlap, the 2 means will be significantly different ($p < 0.05$)

Example: 12 month study of 2 drugs used to lower cholesterol

Drug A: 190 ± 12 (178-202)

Drug B: 165 ± 10 (155-175)

SEMs don't overlap so p value will be < 0.05

Non-Parametric Nominal Data: Significance Testing

When 95% CI for odds or risk ratios don't cross one, results will be significant ($p < 0.05$)

Example: 5-year study comparing 2 drugs used to prevent future heart attacks

Drug B vs. Drug A: RR 0.66 (0.60-0.75)

95% CI doesn't cross 1 so p value will be < 0.05

Non-Parametric Ordinal Data: Significance Testing

- Rating / Ranking / Scoring Data
- Data reported as Median Scores ± Range
- Data less exact and significance harder to estimate

Example: 1-year study comparing 2 drugs used to treat Alzheimer's (baseline MMSE scores 24-26)

Drug A: Median MMSE Score 22 (20-24)*
Drug B: Median MMSE Score 17 (14-20)*

***Results expressed as Median (25th-75th% range)**

2b. Was the treatment benefit large?

- Parametric data: Absolute (quantitative) size of benefit
- Nominal (categorical) data: Look at ARR / **NNT**
- Ordinal data: Degree of improvement (qualitative)

Determining the Size of the Benefit in Therapy Studies

1. Parametric Data

e.g. Compare 2 Cholesterol-Lowering Medications After One Year Treatment
Mean Cholesterol on Entry for Patients Enrolled in Study = 250 mg%

Drug A: Cholesterol 190 $\pm$ 12 mg% = 24% reduction

Drug B: Cholesterol 165 $\pm$ 10 mg% = 34% reduction

At One Year: Drug B provides 10% better reduction in Cholesterol than Drug A

---

2. Nominal (Categorical) Data: Dichotomous Data

e.g. Study Effect of Drug A vs. Drug B in Reducing Risk for Recurrent MI after 5 years

Recurrent MI after 5 years: Drug A Group = 8.9% Drug B Group = 5.9%

Risk Ratio for Recurrent MI with Drug B vs. Drug A = 0.66 (95% C.I., 0.60 – 0.75)

To determine the size of the benefit, must calculate the Number Needed to Treat (NNT)

NNT = 1 / Absolute Risk Reduction so NNT = 1 / .03 = 33.3

At 5 years, for every 34 patients treated with Drug B vs. Drug A, you can prevent 1 MI

---

3. Ordinal (Rating, Ranking, Scoring) Data

e.g. Study Effect of 2 Different Medications to Treat Alzheimer's Dz on MMSE Scores

Enroll Patients with Mild Dementia and baseline MMSE scores 24-26 range

Drug A Group: MMSE Score at 1 Year = 22 (25^{th}-75^{th}% range, 20-24)

Drug B Group: MMSE Score at 1 Year = 17 (25^{th}-75^{th}% range, 12-19)

You cannot assess the size of the benefit because rating / ranking / scoring data is more qualitative than quantitative (data more subjective, less exact)

Parametric Data

- Quantitative Data / Measured Variables
- Data reported as Average ± SD or Mean ± SEM

Example: 12-month study of 2 drugs used to lower cholesterol in pts with high Cholesterol

Baseline Cholesterol Mean = 250 ± 15 (SEM)
Drug A: Mean 190 ± 12 (24% lowering)
Drug B: Mean 165 ± 10 (34% lowering)

Non-Parametric Nominal Data

- Categorical Data
- Most common = dichotomous data (2 categories)
- Data reported as: Relative Risks / Risk Ratios / Odds Ratios / Likelihood Ratios (with 95% C.I.)

Example: 5-year study comparing 2 drugs used to prevent future heart attacks

Drug A: 8.9% MIs Drug B: 5.9% MIs

Drug B vs. Drug A: RR 0.66 (0.60-0.75)

Non-Parametric Nominal Data (2)

- **Example: 5-year study comparing 2 drugs used to prevent future heart attacks**
  Drug A: 8.9% MIs Drug B: 5.9% MIs
  Drug B vs. Drug A: RR 0.66 (0.60-0.75)

  RRR: 34%

  ARR: **3%**

  NNT = 1/ARR = 1/.03 = 33.3
  Therefore, 34 patients would need to be treated with Drug B instead of Drug A to prevent one MI

Non-Parametric Ordinal Data

- **Rating / Ranking / Scoring Data**
- **Data reported as Median Scores ± Range**
- **Data less exact and only note degree of improvement with treatment**

Example: 2-year study comparing 2 drugs used to treat Alzheimer's Dementia (baseline MMSE scores 24-26)
Drug A: 22 (25-75% range, 20-24)
Drug B: 17 (25-75% range, 14-20)

2c. Was the treatment benefit shown to be precise?

- Parametric Data
- Non-Parametric Nominal (Categorical) Data
- Non-Parametric Ordinal Data: **Not precise**

How Precise Was the Estimate of the Treatment Effect?

1. Parametric Data

e.g. Compare 2 Cholesterol-Lowering Medications After One Year Treatment

If the SEM is ± 10% of the mean, the data are very precise

Drug A: Cholesterol 190 ± 12 mg% so SEM = 12/190 (6%)

Drug B: Cholesterol 165 ± 10 mg% so SEM = 10/165 (6%)

} **so the data are very precise**

---

2. Nominal (Categorical) Data

e.g. Study Effect of Drug A vs. Drug B in Reducing Risk for Recurrent MI

Risk Ratio for Recurrent MI with Drug B vs. Drug A = 0.66 (95% C.I., 0.60 – 0.75)

If the 95% C.I. is ≤ 30% of the reported value, the data are very precise

Size of 95% Confidence Interval (0.75 – 0.60 = 0.15)

is less than 30% of the RR (30% of 0.66)

} **so the data are very precise**

---

3. Ordinal (Rating, Ranking, Scoring) Data

e.g. Study Effect of 2 Different Medications to Treat Alzheimer's Dz on MMSE Scores

Drug A: MMSE Score at 1 Year = 22 (25-75% range, 20-24)

Drug B: MMSE Score at 1 Year = 17 (25-75% range, 12-19)

*No easy way to estimate precision of result for ordinal data: Data is more subjective, and by definition, is **not precise**.*

Assessing Precision in Studies with Parametric Data

RULE: If the SEM is ± 10% of the mean, the data are very precise

Example: 12 month study of 2 drugs used to lower cholesterol (expressed as Mean ± SEM)

Drug A: 190 **± 12**

Drug B: 165 **± 10**

SEM are less than 10% of the Mean so data are **precise**

Assessing Precision in Studies with Non-Parametric Nominal Data

<u>RULE</u>: If the 95% CI difference is less than 30% of the reported value, the data are precise

Example: 5-year study comparing 2 drugs used to prevent future heart attacks

Drug B vs. Drug A: RR 0.66 (**0.60-0.75**)

The size of the CI difference (0.15) is < 30% of the RR (0.66 x 30% = .198) so the data are **<u>precise</u>**

Assessing Precision in Studies with Non-Parametric Ordinal Data

Ordinal data

- More subjective data reporting
- Data reported as median with range
- This data is NOT precise!

Practice Exercise

- Review the ***Determining Significance, Size and Precision of Treatment Benefits*** worksheet

- For the two studies described, determine:
  * Is the treatment benefit significant?
  * If so, how large is the benefit?
  * If so, how precise is the measured benefit?
  * If not precise, how can precision be improved?

Therapy Study

3. Will the study results help me in caring for my patients?

- Are the patients studied similar to mine?
- Were clinically important outcomes/benefits demonstrated? **(next slide)**
- Were significant adverse effects considered?
- Is the treatment benefit worth the possible harms and costs? (cost-benefit analysis)

Worksheet for Assessing a Therapy Article

ARE RESULTS OF THE STUDY VALID?

1. Primary Guides — **Circle Yes / No**

a. Was allocation concealed from those enrolling patients in study? Yes / No
b. Was the study a randomized controlled trial? Yes / No
c. Were all study patients properly accounted for at conclusion of the study? Yes / No
d. Were the number of patients dropping out or lost to follow-up small (< 20%) & approximately equal between groups? Yes / No
e. Were patients analyzed in the group to which they were randomized? (intention-to-treat principle) Yes / No

2. Secondary Guides

a. Were patients and study personnel blind to treatment? Yes / No
b. Were patients similar / balanced at the start of the trial? Yes / No
c. Were the groups treated equally (aside from the experimental intervention)? Yes / No
d. Was the study sponsored / funded by a pharmaceutical/device company? Yes / No
If so, is there evidence of bias? Yes / No
e. Regarding statistical analysis:
→Was a power calculation done & was the proper sample size then recruited? Yes* / No*

**Power Analysis needed when trial results negative*

WHAT ARE THE RESULTS?

1. Was the treatment effect proven significant to a $p < 0.05$ level? Yes / No

2. Was the treatment effect large (can ARR / RRR / NNT be determined)? Yes / No

3. Are the results clinically as well as statistically significant? Yes / No

4. Was the treatment effect shown to be precise?

a. For quantitative data, were standard errors of the mean ≤ 10% of the mean? Yes / No
b. For categorical data, were 95% CI ≤ 30% of the reported value? Yes / No
c. For ordinal data, there is no good way to estimate precision.

WILL THE RESULTS HELP ME CARE FOR MY PATIENTS?

1. Are the patients in the study similar to mine (inclusion/exclusion criteria)? Yes / No

2. Were clinically important outcomes / benefits of treatment identified? Yes / No

3. Were significant adverse effects of treatment considered? Yes / No

4. Is the treatment benefit worth the possible harms and costs? Yes / No

Any Questions?