Evidence-Informed Practice Workshop Series

Level 2: Critical Appraisal of Research Evidence

Outline

- Introduction to Critical Appraisal
- Break
- Critical Appraisal - Intervention Study
  - Independent review and group discussion
- Critical Appraisal - Systematic Review
  - Independent review and group discussion

Objectives

- Introduce some basic statistical terminology.
- Outline a process to use when reviewing research studies – only the best evidence warrants your time!
- Discuss central elements of critical appraisal:
  - Validity – controls for bias
  - Results – how strong and how precise
  - Applicability – to your practice

Getting Started

- Instructors in the EIP Workshop have discovered an alarming trend in participants apparently falling asleep during the presentations

P: Among healthcare workers who attend full day workshops,

I: does ingestion of a caffeinated beverage at the beginning of the day

C: versus ingestion of non-caffeinated beverage

O: improve knowledge retention, prevent falling asleep during the workshop, lead to better application in the clinical setting?

Free Tim's!

50 people sign up for this prestigious workshop

- Randomized to 2 groups:
  - Full caffeinated n=25
  - Decaffeinated n=25

Outcome measures:

- Fall asleep during workshop yes/no
- Wakefulness score 1-5 Likert scale self report
- Knowledge test out of 100
- Excessive urination yes/no

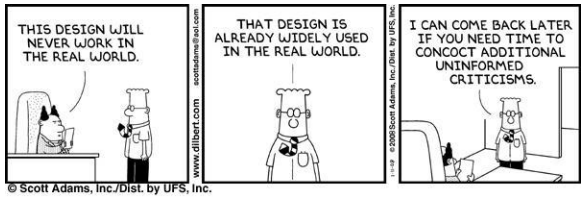
Measuring Outcomes

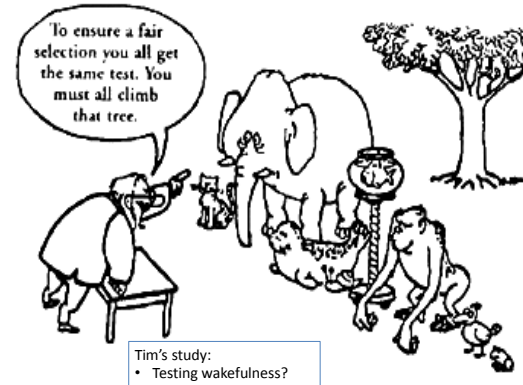
- Self report at the end of the day (did you fall asleep?)
- 5 point Likert scale?
- Hidden camera?
- Snore meter?

Clinical Relevance

- Method fit the question?
  - What are the goals and will the method illicit the information they want?
- Stretching it?
  - Logical?
  - Contribute meaningful info?

Tim's study:
- What is the goal?
- What do we know about caffeine?
- Just because people might be more awake, can we assume that they learn more?





Tim's study:
- Testing wakefulness?
- Testing knowledge?
- Testing application?

Research Design

Good match?

Tim's study:
- Are there other relevant outcomes that the design can't measure? (how caffeine makes you feel?)

Right vehicle?

Tim's study:
- Is a Tim's the right way to administer caffeine?
- Is a multiple choice test the best way to determine knowledge retention?

Determining Flaws

- Clarity
- Relevant lit review
- Conflict of interest
- Drop out rates
- Any obvious systematic erros

Tim's study:
- Who pays for the Tim's
- What if someone from the decaf group sneaks a Red Bull?
- What if 10 of the caffeine group get so hyper they don't come back after lunch? Or spend too much time in WC?

Validity - Bias

- Figures don't lie but liars figure!
- Guilty?

- "Intention to treat"

Tim's study:
- Some don't like coffee
- Someone took sleeping pills the night before that haven't worn off
- Same dose of caffeine for all sizes of people?
- Did they have a cup before they came to class?

Types of Bias

- Selection: criteria used to recruit and the actual enrolling of subject inherently different
- Recall: outcome of treatment may color subject's recollection
- Transfer: lost to follow-up

Randomization

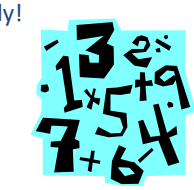
- A method that allows for the equal and independent change of being selected
- Groups should all have the same characteristics at baseline
- Controls for selection bias:
  - Blinding
  - Allocation concealment

Blinding / Concealing

- "Double blind" vs "multiple blinding"
  - Patient, researchers
  - Monitoring committee and evaluators
  - The individual responsible for randomizing does not know what the next treatment allocation will be

Sorting Through the Stats

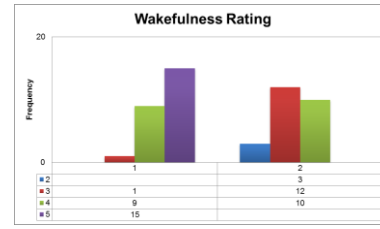
- Statistics don't tell you facts, they tell you what, not why!
- Not all data is created equally!



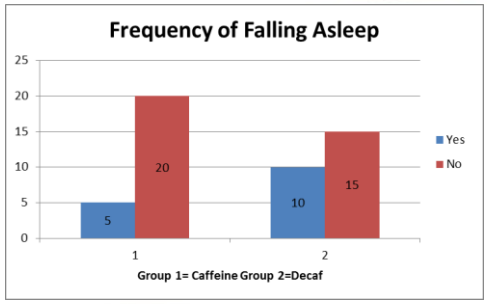
Sample Size

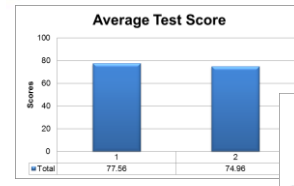
- Did they pre-determine how many subjects they needed (power analysis)
- Did they meet their goal?
- Does the reported number in each group add up correctly?

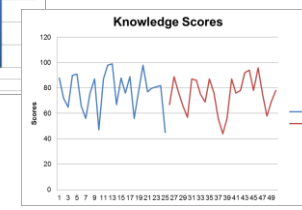
What are the results?



Tim's Study: Group 1 (caffeine) seems to be more awake than Group 2 (decaf)





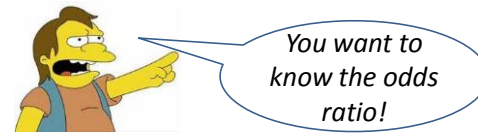


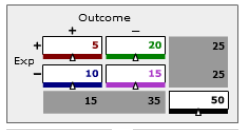
Tim's study:
- Average test scores look similar

Dichotomous vs Continuous

- Dichotomous / Discreet variables – yes/no questions, specific choices (ie. scale of 1-5)
- Significance is expressed in confidence intervals
- Continuous variables – any value is possible (ie. Weights and measures, points)
- Significance is expressed as p-values

What is the chance that I will fall asleep during this workshop if I drink a fully caffeinated Tim's at the beginning of the day vs if I drink decaf?





Measure	Value	95% CI
CER:	40%	20% - 59%
EER:	20%	4% - 35%
OR:	0.38	
RR:	0.50	
ARR:	20%	
RRR:	50%	
NNT:	5	

Tim's study:
Fall asleep (Outcome +)
Caffeine (the treatment or Exp +)

Caffeine Group 5/25 yes
Decaf Group 10/25 yes

Significance

It's relative!

Margin of Error

- Size matters!
- How much error are we willing to accept?
- 95% is the accepted limit

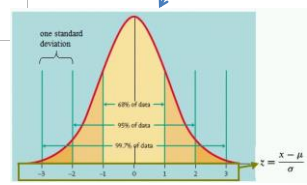
P-value

- A P value <0.05 means that there is a less than 1 in 20 probability that, on repeated performance of the experiment, the results as extreme as or more extreme than those observed would occur if the null hypothesis (that caffeine makes no difference to keeping you awake) were true

There will be one curve for each group. Now the two groups can be compared, and a p-value will tell you if they are statistically different. With only a 3% difference in the average test scores p>.05

Statistical manipulation (legally) turns This into

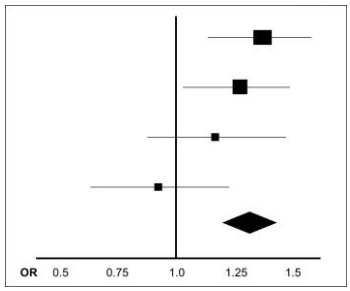
This



Confidence Intervals (CI)

- 95% Confidence Interval - you are 95% sure that the real answer lies between these two numbers
- Treatment studies - if the CI includes or encompasses 1, it is not significant

  CI -5:7 would not be significant, even if the study showed that people who got the intervention were 2% more likely to have a positive effect



Forrest Plot:
- Multiple studies on the same chart
- Odds ratio & confidence interval
- Compares results of studies of the same hypothesis

Strength vs Confidence

- The strength or significance of the result is balanced by the confidence that the result is correct
- Strength = how small is the p-value, how big is the difference between groups
- Confidence = how sure are you that there is no bias (precision or margin of error)

Reducing Risk or Events (Illness)

Compares the rate of the problem in the
- Control group vs treatment group
- Exposed group vs protected group

Tim's study:
- Is Caffeine the treatment for sleepiness, or the protection from boring presenter?

Tim's Coffee Decreases Falling Asleep In Class by 50%!

Relative Risk Reduction

50% RRR:

Look closely at the real numbers:
- 40% risk reduced to 20%
- 10% risk reduced to 5%
- 1% risk reduced to .5%

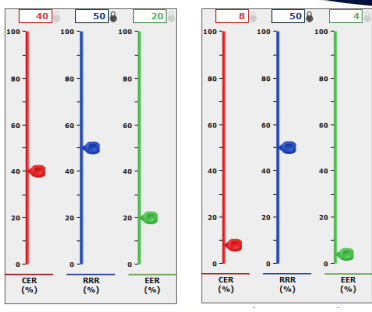
Tim's study results:
Fall asleep (the disease!)
yes/no

Decaf Group 2/25 yes
Caffeine Group 1/25 yes

OR

Decaf Group 10/25 yes
Caffeine Group 5/25 yes

Same RRR!



RRR is the same, but the impact of the treatment is different.

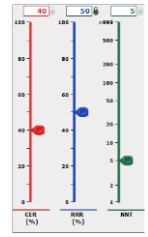
Absolute Risk Reduction

- 40% risk reduced to 20% - a 20% reduction
- 10% risk reduced to 5% - a 5% reduction
- 1% risk reduced to .5% - a .5% reduction

Tim's study results: Fall asleep yes/no
Decaf (Control) Group 2/25 (8%) Caffeine (Treatment)Group 1/25 (4%) = 4% ARR
OR
Decaf (Control) Group 10/25 (40%) Caffeine (Treatment) Group 5/25 (20%) = 20% ARR

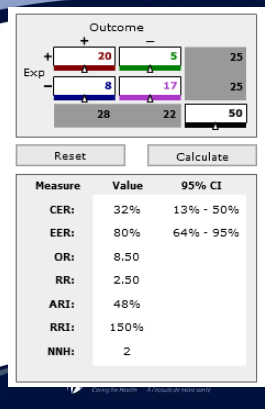
Number Needed to Treat

- How many people need to get caffeine to save one from falling asleep? 5



Number Needed to Harm

- The intervention causes a complication
- How many people need to get the intervention in order to have one harmed?
- What's worse, having the disease, or the harm resulting from the treatment?



Tim's study:
Excessive Urination (Outcome +)
Caffeine (the treatment or Exp +)

Caffeine Group 20/25 yes
Decaf Group 8/25 yes

- How many people need to get the intervention in order to have one harmed? 2

Evaluating Charts

- Look for
  - original raw numbers in the data
  - Scale
  - Starting points
  - Units of measure

"What" does not explain "Why"

Control: Many Variables, Many Answers

- A lot of things contribute to illness
- You can't control all of them
- You can "tease them out" statistically
- Called a "regression analysis"
- The stats program can determine the relative contribution of each of the variables entered ie. Socioeconomic, age, co-illnesses etc.

Application to Practice

- Were the study patients similar to the people in my clinical setting?
- Were all important outcomes considered?
- Are the likely intervention benefits worth the potential harms and costs?

Application to Practice

Next Steps

- Critical Appraisal
- Worksheets
  - Health Care Intervention
  - Qualitative Research
  - Systematic Reviews
  - Available for a large variety of types of research – on the cdrom, on the internet

Break?

Critical Appraisal
Health Care Intervention

- Take 20-30 minutes to review the article provided using the worksheet in the handout
- Review the worksheet as a group

Critical Appraisal
Systematic Review

- Take 20-30 minutes to review the article provided using the worksheet in the handout
- Review the worksheet as a group

Conclusion of Level 2

Next level: What to do with all that literature

Making Recommendations