



Rapid Critical Appraisal of Controlled Trials

Carl Heneghan
Dept of Primary Health Care
University of Oxford

November 23rd 2009





Five steps in EBM

1. Formulate an answerable question
2. Track down the best evidence
3. Critically appraise the evidence for:
 - Relevance
 - Validity
 - Impact (size of the benefit)
 - Applicability
4. Integrate with clinical expertise and patient values
5. Evaluate our effectiveness and efficiency
 - keep a record; improve the process



Searching for critical appraisal checklists randomized controlled trials .

11,100 articles (0.40 seconds)

A CHECKLIST FOR APPRAISING RANDOMIZED CONTROLLED TRIALS

1. Was the objective of the trial sufficiently described?
2. Was a satisfactory statement given of the diagnostic criteria for entry to the trial?
3. Were concurrent controls used (as opposed to historical controls)?
4. Were the treatments well defined?
5. Was random allocation to treatments used?
6. Was the potential degree of blindness used?
7. Was there a satisfactory statement of criteria for outcome measures? Was a primary outcome measure identified?
8. Were the outcome measures appropriate?
9. Was a pre-study calculation of required sample size reported?
10. Was the duration of post-treatment follow-up stated?
11. Were the treatment and control groups comparable in relevant measures?
12. Were a high proportion of the subjects followed up?
13. Were the drop-outs described by treatment and control groups?
14. Were the side-effects of treatment reported?
15. How were the ethical issues dealt with?
16. Was there a statement adequately describing or referencing all statistical procedures used?
17. What tests were used to compare the outcome in test and control patients?
18. Were 95% confidence intervals given for the main results?
19. Were any additional analyses done to see whether baseline characteristics (prognostic factors) influenced the outcomes observed?
20. Were the conclusions drawn from the statistical analyses justified?



Victims of DVT are told that they can't sue



By Andrew Law
and David Smith

WINDMILLERS who have suffered the horror of a blood clot during a flight might worry whether they will not claim as regards compensation.

More than 100 airlines and millions of passengers who die, are left devastated by the airline. The High Court in London is now hearing whether a group of passengers, including the family of a man who died, can sue the airline for negligence.

The British group would sue the airline for negligence by failing to warn passengers of the risk of a blood clot during a flight.

With Chris Jackson, who is a partner in the law firm of Jackson & Clegg, the group is suing the airline for negligence by failing to warn passengers of the risk of a blood clot during a flight.

The claimants want to sue the airline for negligence by failing to warn passengers of the risk of a blood clot during a flight.

(Except if it's in Australia)

When the Boeing 777-300ER was launched in 2009, it was the world's largest passenger jet. It was designed to fly for 12 hours non-stop. The airline industry was confident that it was safe for long-haul flights.

But the airline industry was wrong. In 2010, a British Airways flight from London to Los Angeles was delayed for several hours because of a problem with the engine. The airline was forced to divert the plane to a nearby airport.

DVT is a blood clot that can form in the legs during a long flight. It is a serious condition that can lead to a heart attack or stroke.

The airline industry has been accused of negligence for not warning passengers of the risk of DVT during a long flight.

The airline industry has been accused of negligence for not warning passengers of the risk of DVT during a long flight.

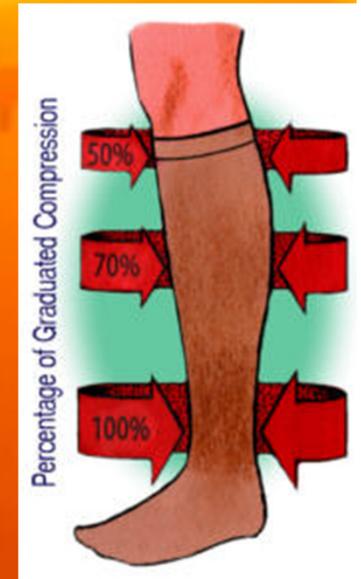
The airline industry has been accused of negligence for not warning passengers of the risk of DVT during a long flight.



Clinical Question

In people who take long-haul flights does wearing graduated compression stockings prevent DVT?

Page 71 and 95 in your books





Causes of an “Effect” in a controlled trial

- Who would consider wearing stockings on a long haul flight?



QUESTION:

Participants

Intervention Group (IG)
& Comparison Group
(CG)

Outcome

VALIDITY

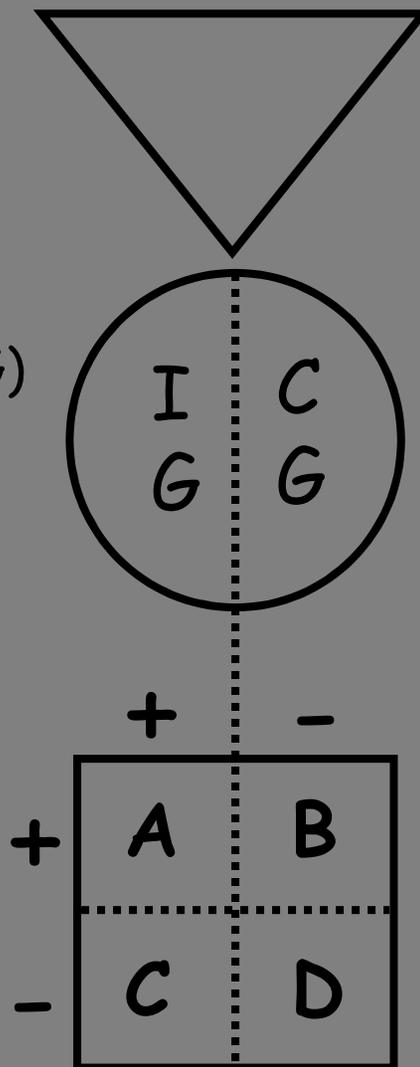


QUESTION:

Participants

Intervention Group (IG)
& Comparison Group (CG)

Outcome



VALIDITY

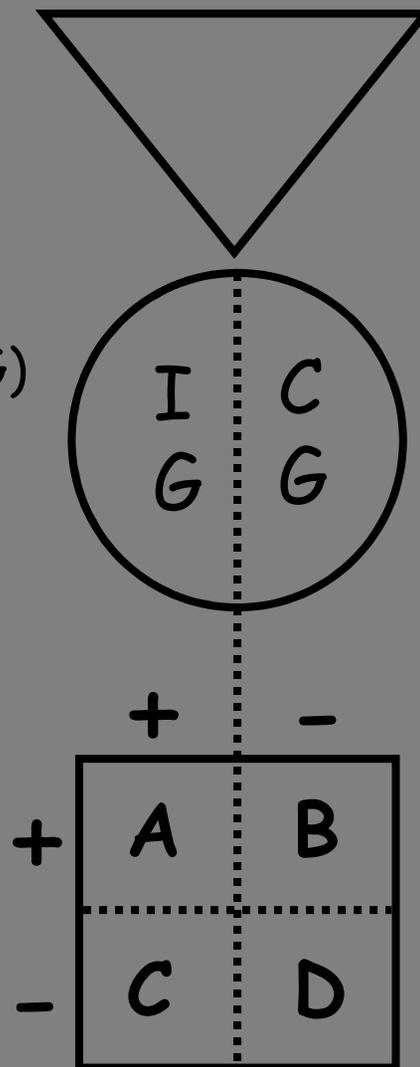


QUESTION:

Participants

Intervention Group (IG)
& Comparison Group (CG)

Outcome



VALIDITY

Recruitment

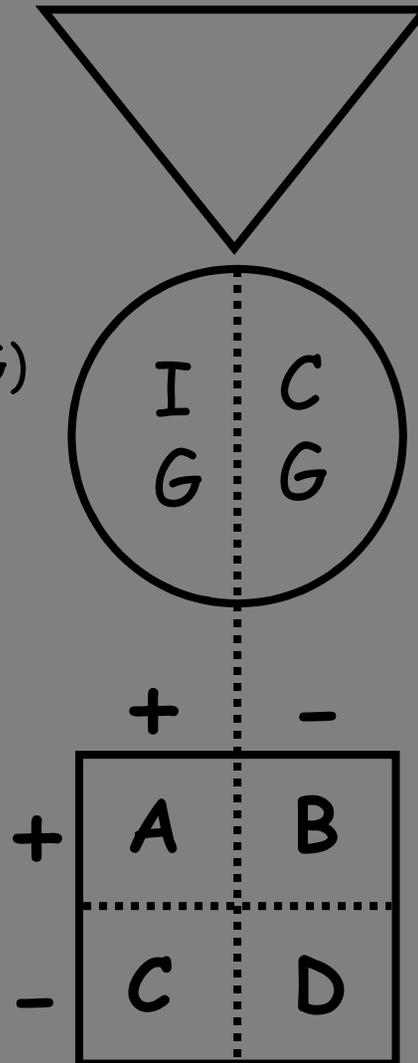


QUESTION:

Participants

Intervention Group (IG)
& Comparison Group (CG)

Outcome



VALIDITY

Recruitment

Allocation
concealment
comparable groups

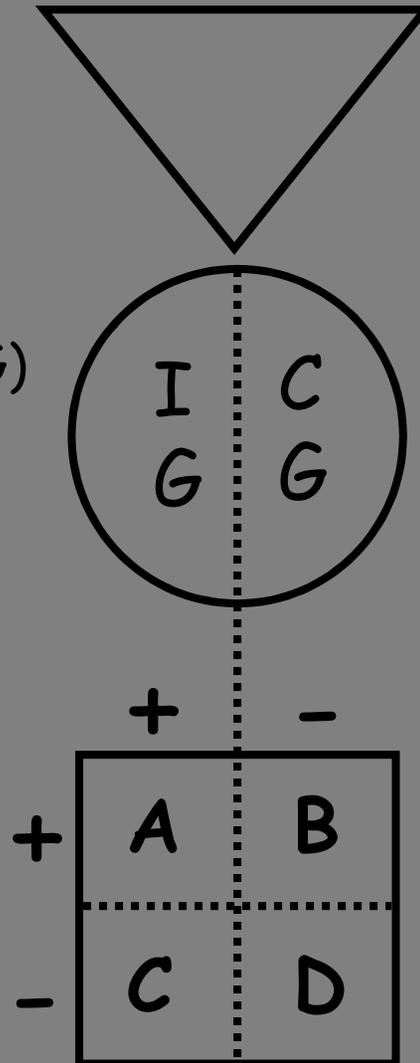


QUESTION:

Participants

Intervention Group (IG)
& Comparison Group (CG)

Outcome



VALIDITY

Recruitment

- Allocation
- concealment
 - comparable groups

- Maintenance
- treated equally
 - compliant

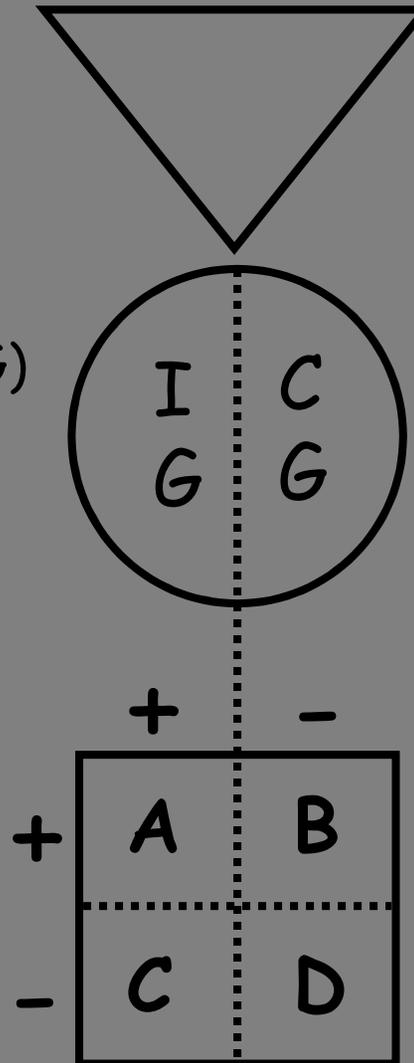


QUESTION:

Participants

Intervention Group (IG)
& Comparison Group (CG)

Outcome



VALIDITY

Recruitment

Allocation concealment
comparable groups

Maintenance
• Treated equally
• compliant

Measurements
blind? OR
objective?



Appraisal checklist - **RAMMbo**



Study biases

1. **Recruitment**
 - Who did the subjects represent?
2. **Allocation**
 - Was the assignment to treatments randomised?
 - Were the groups similar at the trial's start?
3. **Maintenance**
 - Were the groups treated equally?
 - Were outcomes ascertained & analysed for most patients?
4. **Measurements**
 - Were patients and clinicians "blinded" to treatment? OR
 - Were measurements objective & standardised?

Study statistics (p-values & confidence intervals)

Guyatt. JAMA, 1993



Frequency and prevention of symptomless deep-vein thrombosis in long-haul flights: a randomised trial

John H Scurr, Samuel J Machin, Sarah Bailey-King, Ian J Mackie, Sally McDonald, Philip D Coleridge Smith

How were the patients recruited?



Frequency and prevention of symptomless deep-vein thrombosis in long-haul flights: a randomised trial

John H Scurr, Samuel J Machin, Sarah Bailey-King, Ian J Mackie, Sally McDonald, Philip D Coleridge Smith

Randomization
Volunteers were
randomized by
sealed envelope
to one of two
groups.

Passengers were randomly allocated to one of two groups: one group wore class-I below-knee graduated elastic compression stockings, the other group did not.



- Take out the envelopes
- Sign the back



- You have now consented to the trial
- Please open your envelopes now



| | Blue Bunnies | Pink Bunnies |
|-----------------------|--------------|--------------|
| Argued with your boss | | |
| Been to New York | | |



Ensuring Allocation Concealment

BEST – most valid technique

- Central computer randomization



DOUBTFUL

- Envelopes, etc



NOT RANDOMIZED

- Date of birth, alternate days, etc



Were the groups similar at the trial's start?

By chance a greater proportion of women were included in the stocking group

Results

Volunteers were excluded before randomisation if they did not fulfil the entry requirements or could not attend hospital for investigation both before and after travel (figure). Thus, 231 of 479 volunteers were randomised. 27 passengers were unable to attend for subsequent ultrasound investigation because of ill-health (three), change of travel plans, or inability to keep appointments (24). Two who

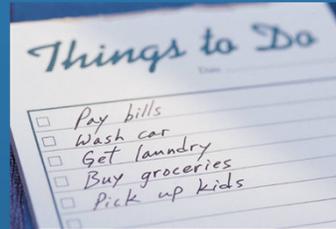
| | No stockings | Stockings |
|-------------------------------|------------------|------------------|
| Number | 116 | 115 |
| Age (years) | 62 (56-68) | 61 (56-66) |
| Number of women (%) | 61 (53%) | 81 (70%) |
| Number with varicose veins | 41 | 45 |
| Days of stay | 17 (13-32) | 16 (13-27) |
| Hours flying time | 22 (18-36) | 24 (19-35) |
| Haemoglobin (g/L) | 142 (133-149) | 140 (133-147) |
| WBC ($\times 10^9/L$) | 5.9 (5.0-7.3) | 6.0 (5.0-6.9) |
| Packed cell volume | 0.44 (0.42-0.47) | 0.44 (0.41-0.46) |
| Platelets ($\times 10^9/L$) | 240 (206-272) | 242 (219-290) |
| Number FVL positive | 7 | 4 |
| Number PGM positive | 1 | 3 |

Median (interquartile range) shown, unless otherwise indicated. WBC=white blood cells. FVL=factor V Leiden. PGM=prothrombin gene mutation.

Table 1: Characteristics of study groups



Appraisal checklist - **RAMMbo**



Study biases

1. **Recruitment**
 - Who did the subjects represent?
2. **Allocation**
 - Was the assignment to treatments randomised?
 - Were the groups similar at the trial's start?
3. **Maintenance**
 - **Were outcomes ascertained & analysed for most patients?**
 - **Were the groups treated equally?**
4. **Measurements**
 - Were patients and clinicians "blinded" to treatment? OR
 - Were measurements objective & standardised?

Study statistics (p-values & confidence intervals)



Effects of non-equal treatment

- Apart from actual intervention - groups should receive identical care!
 - Trial of Vitamin E in pre-term infants (1949)
 - Vit E "prevented" retrolental fibroplasia

Rx: Give placebo in an identical regime, and a standard protocol



Equal treatment in DVT study?

Table 3: All drugs taken by volunteers who attended for examination before and after air travel*

| | Number of participants | |
|----------------------------------------|------------------------|-----------|
| | No stockings | Stockings |
| Aspirin | 9 | 11 |
| Hormone replacement therapy | 8 | 16 |
| Thyroxine | 6 | 6 |
| Antihypertensives, including diuretics | 10 | 12 |
| Antipeptic ulcer drugs | 8 | 3 |

*Includes additions to usual drugs.

Table 3: All drugs taken by volunteers who attended for examination before and after air travel*

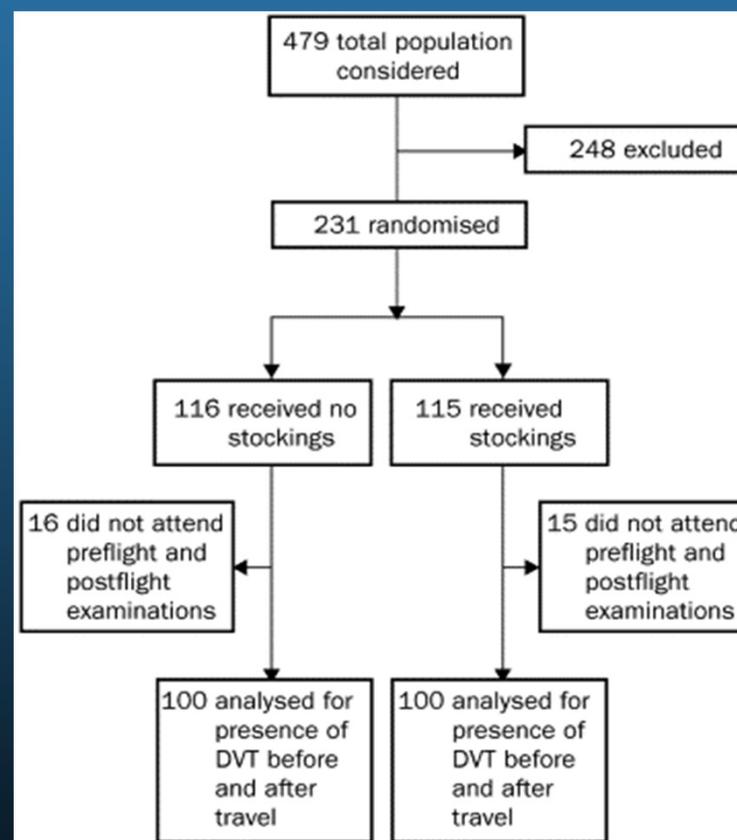
THE LANCET • Vol 357 • May 12, 2001



Follow-up in DVT study?

- 200 of 231 analyzed (87%)
- 27 were unable to attend for subsequent ultrasound
- 2 were excluded from analysis because they were upgraded to business class
- 2 were excluded from analysis because they were taking anticoagulants

See figure on page 96



Scurr et al, Lancet 2001; 357:1485-89



Losses-to-follow-up

How many is too many?

"5-and-20 rule of thumb"

- 5% probably leads to little bias
- >20% poses serious threats to validity

Depends on outcome event rate and comparative loss rates in the groups

Loss to follow-up rate should not exceed outcome event rate and should not be differential



How important are the losses?

- Equally distributed?
 - **Stocking group:** 6 men, 9 women - 15
 - **No stocking group:** 7 men, 9 women - 16
- Similar characteristics?
 - No information provided



Intention-to-Treat Principle

Maintaining the randomization

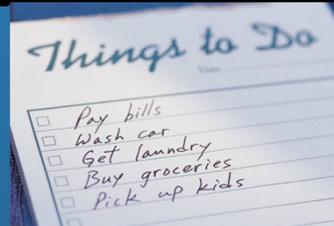
Principle:

Once a patient is randomized, s/he should be analyzed in the group randomized to - even if they discontinue, never receive treatment, or crossover.

Exception: If patient is found on BLIND reassessment to be ineligible based on pre-randomization criteria.



Appraisal checklist



Study biases

1. **R**ecruitment
 - Who did the subjects represent?
2. **A**llocation
 - Was the assignment to treatments randomised?
 - Were the groups similar at the trial's start?
3. **M**aintenance
 - Were outcomes ascertained & analysed for most patients?
 - Were the groups treated equally?
4. **M**easurements
 - Were patients and clinicians "blinded" to treatment? OR
 - Were measurements objective & standardised?

Study statistics (p-values & confidence intervals)



Measures in DVT study?

- Blood was taken from all participants before travel
- All participants had US once before travel (30 had US twice)
- All participants were seen within 48 hr of return flight, were interviewed and completed a questionnaire, had repeat US



Measurement Bias - minimizing differential error

- Blinding – Who?
 - Participants?
 - Investigators?
 - Outcome assessors?
 - Analysts?

- Most important to use "blinded" outcome assessors when outcome is **not objective!**

- Papers should report **WHO** was blinded and **HOW** it was done



Figure 1: The authors: double blinded versus single blinded



Figure 2: The authors blinded and masked



Frequency and prevention of symptomless deep-vein thrombosis in long-haul flights: a randomised trial

John H Scurr, Samuel J Machin, Sarah Bailey-King, Ian J Mackie, Sally McDonald, Philip D Coleridge Smith

Evaluation

Most passengers removed their stockings on completion of their journey. The nurse removed the stockings of those passengers who had continued to wear them. A further duplex examination was then undertaken with the technician unaware of the group to which the volunteer had been randomized.

Summary

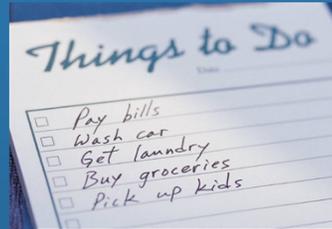
Background The true frequency of deep-vein thrombosis (DVT) during long-haul air travel is unknown. We sought to determine the frequency of DVT in the lower limb during long-haul economy-class air travel and the efficacy of graduated elastic compression stockings in its prevention.

Methods We recruited 89 male and 142 female passengers over 50 years of age with no history of thromboembolic problems. Passengers were randomly allocated to one of two groups: one group wore class-I below-knee graduated elastic compression stockings, the other group did not. All the passengers made journeys lasting more than 8 h per group (median total duration 24 h), returning to the UK within 72 h. Duplex ultrasonography was used to assess the deep veins before and after travel. Blood samples were analysed for two specific common gene mutations, factor V Leiden (FVL) and prothrombin G20210A (PGM), which predispose to venous thromboembolism. A sensitive D-dimer assay was used to screen for the development of recent thrombosis.

Findings 12/116 passengers (10%; 95% CI 4.8–16.0%) developed symptomless DVT in the calf (five men, seven women). None of these passengers wore elastic



Appraisal checklist



Study biases

1. **Recruitment**
 - Who did the subjects represent?
2. **Allocation**
 - Was the assignment to treatments randomised?
 - Were the groups similar at the trial's start?
3. **Maintainence**
 - Were the groups treated equally?
 - Were outcomes ascertained & analysed for most patients?
4. **Measurements**
 - Were patients and clinicians "blinded" to treatment? OR
 - Were measurements objective & standardised?
5. **Placebo Effect**
6. **Chance**
7. **Real Effect**

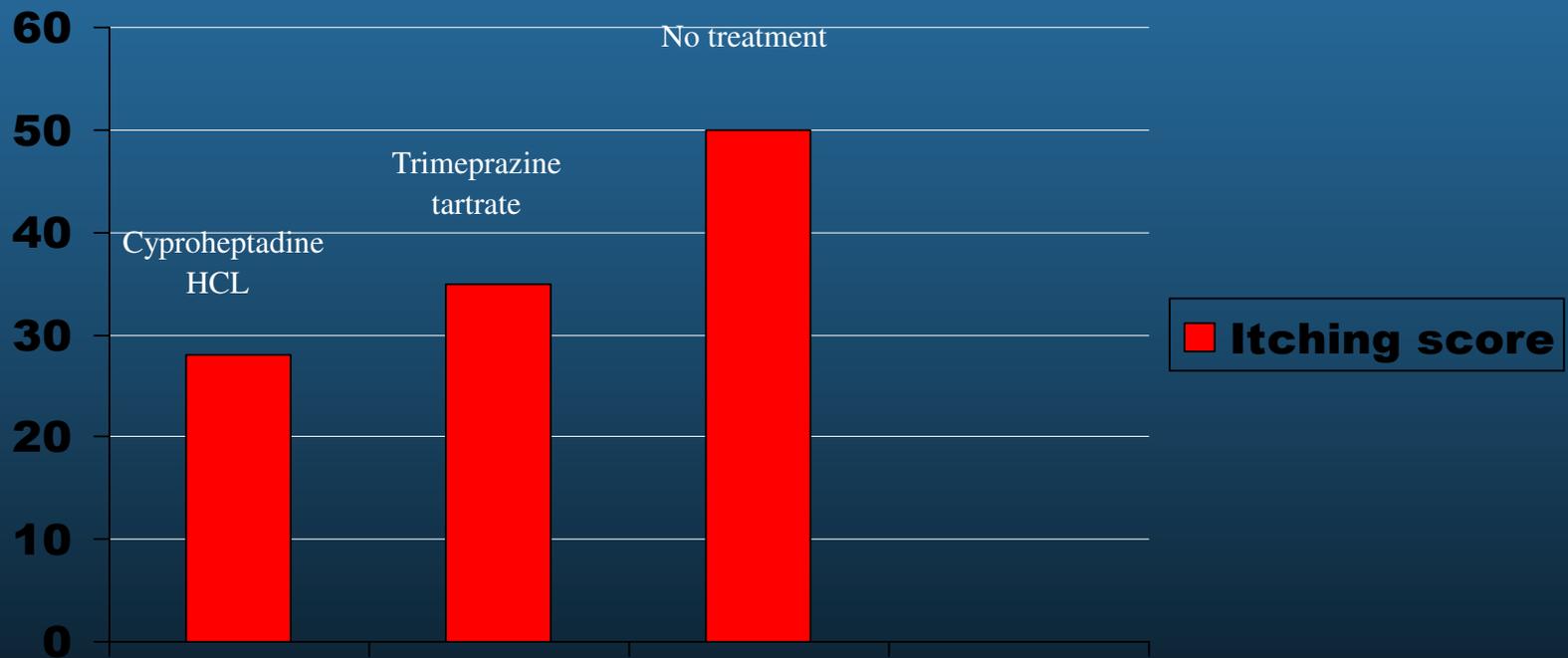
Study statistics (p-values & confidence intervals)

Guyatt. JAMA, 1993



Placebo effect

Trial in patients with chronic severe itching

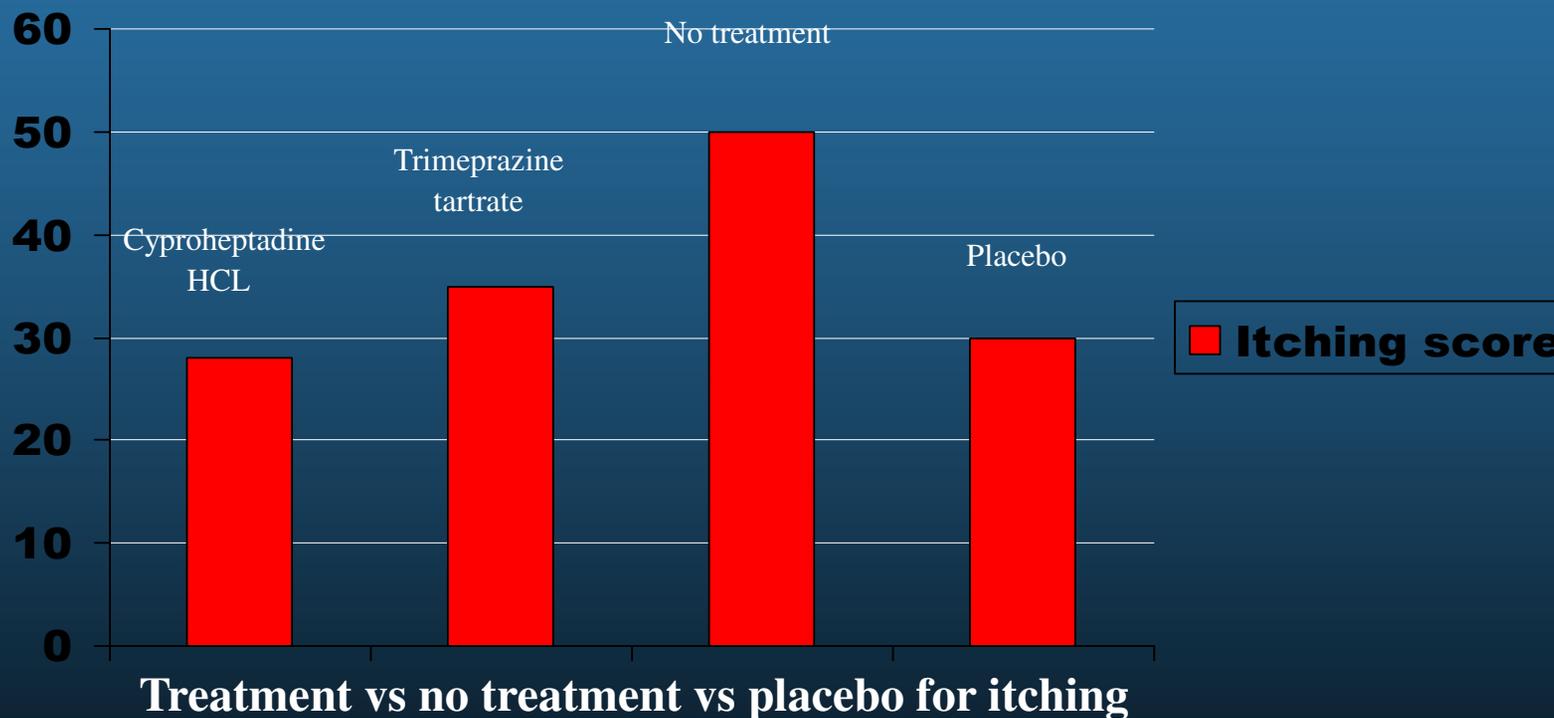


Treatment vs no treatment for itching



Placebo effect

Trial in patients with chronic severe itching



Placebo effect - attributable to the expectation that the treatment will have an effect



Appraisal checklist



Study biases

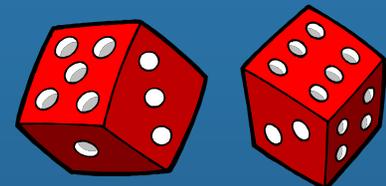
1. **Recruitment**
 - Who did the subjects represent?
2. **Allocation**
 - Was the assignment to treatments randomised?
 - Were the groups similar at the trial's start?
3. **Maintainence**
 - Were the groups treated equally?
 - Were outcomes ascertained & analysed for most patients?
4. **Measurements**
 - Were patients and clinicians "blinded" to treatment? OR
 - Were measurements objective & standardised?
5. **Placebo Effect**
6. **Chance**
7. **Real Effect**

Study statistics (p-values & confidence intervals)

Guyatt. JAMA, 1993



Two methods of assessing the role of chance



- **P-values (Hypothesis Testing)**
 - use statistical test to examine the 'null' hypothesis
 - associated with "p values" - if $p < 0.05$ then result is statistically significant
- **Confidence Intervals (Estimation)**
 - estimates the range of values that is likely to include the true value



P-values (Hypothesis Testing) - in DVT study

- **Incidence of DVT**
 - Stocking group - 0
 - No Stocking group - 0.12



Risk difference = $0.12 - 0 = 0.12$ (P=0.001)

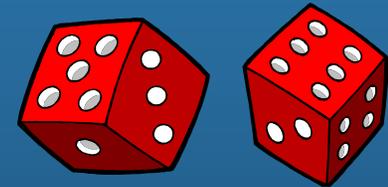
The probability that this result would only occur by chance is

1 in 1000 → **statistically significant**



Confidence Intervals (Estimation)

- Incidence of DVT
 - Stocking group = 0
 - No Stocking group = 0.12



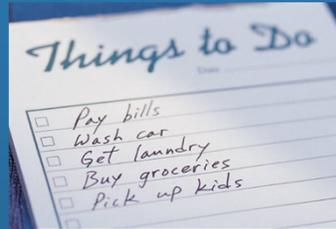
Risk difference = $(0.12 - 0) = 0.12$ (95% CI, 0.058 - 0.20)

The true value could be as low as 0.058 or as high as 0.20 - *but is probably closer to 0.12*

Since the CI does not include the 'no effect' value of '0' → the result is **statistically significant**



Appraisal checklist



Study biases

1. **Recruitment**
 - Who did the subjects represent?
2. **Allocation**
 - Was the assignment to treatments randomised?
 - Were the groups similar at the trial's start?
3. **Maintainence**
 - Were the groups treated equally?
 - Were outcomes ascertained & analysed for most patients?
4. **Measurements**
 - Were patients and clinicians "blinded" to treatment? OR
 - Were measurements objective & standardised?
5. **Placebo Effect**
6. **Chance**
7. **Real Effect**

Study statistics (p-values & confidence intervals)

Guyatt. JAMA, 1993



Causes of an “Effect” in a controlled trial

- Who would now consider wearing stockings on a long haul flight?



M Clarke, S Hopewell, E Juszczak, A Eisinga, M Kjeldstrøm

Compression stockings for preventing deep vein thrombosis in airline passengers

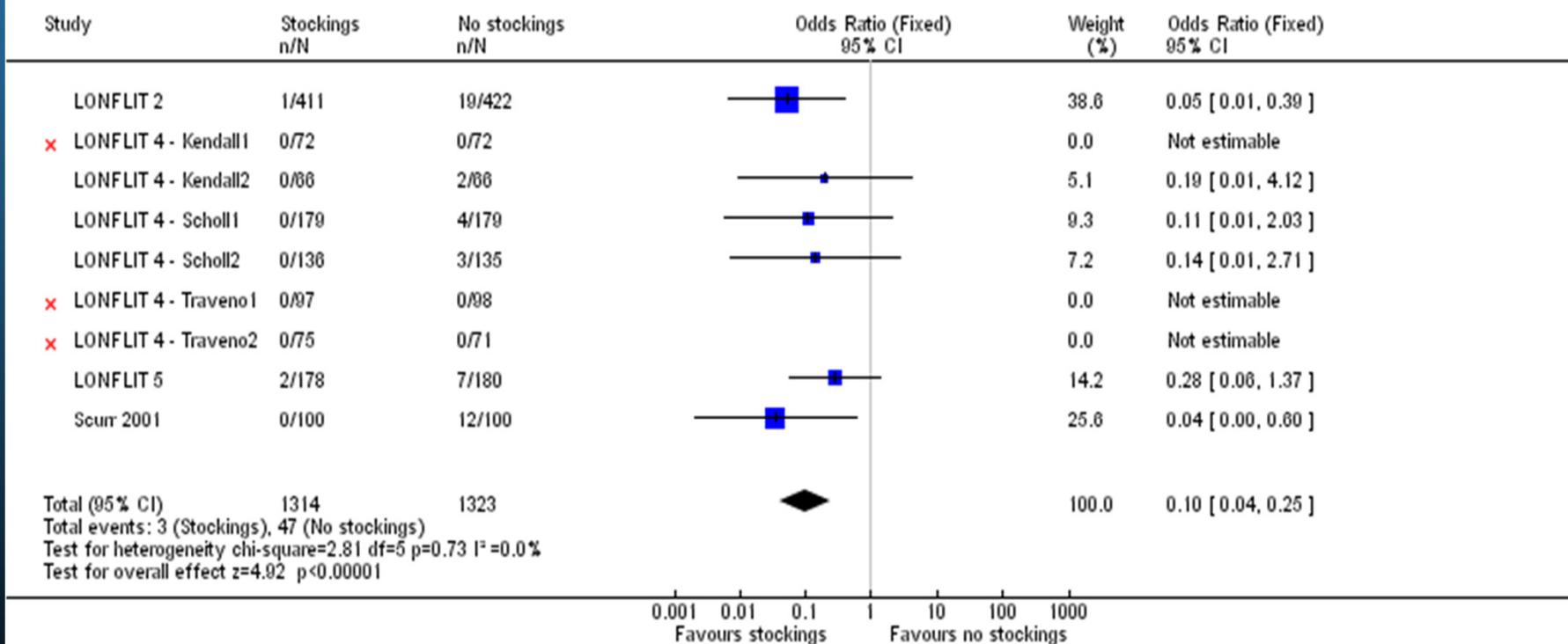
Cochrane Database of Systematic Reviews 2006 Issue 4

- 10 RCTs (n = 2856)
- Seven included low or medium risk (n = 1548) and two included high risk participants (n = 1273).
- All flights > seven hours.
- Fifty of 2,637 participants in the trials of wearing stockings on both legs had a symptomless DVT; three wore stockings, 47 did not
(OR 0.10, 95% CI 0.04 to 0.25, P < 0.00001).
- No deaths, pulmonary emboli or symptomatic DVTs were reported.
- Wearing stockings had a significant impact in reducing oedema (based on six trials).
- No significant adverse effects were reported.



M Clarke, S Hopewell, E Juszczak, A Eisinga, M Kjeldstrøm
Compression stockings for preventing deep vein thrombosis in airline passengers
Cochrane Database of Systematic Reviews 2006 Issue 4

Review: Compression stockings for preventing deep vein thrombosis in airline passengers
 Comparison: 01 Wearing stockings versus not wearing stockings
 Outcome: 01 Symptomless deep vein thrombosis





Thank you

Small groups