CEBM CENTRE FOR EVIDENCE BASED MEDICINE

Rapid Critical Appraisal of Randomised Controlled Trials

Dr Dan Lasserson
Clinical Lecturer
Centre for Evidence Based Medicine
University of Oxford
www.cebm.net

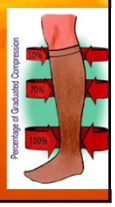
CEBM CENTRE FOR EVIDENCE BASED MEDICINE

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

Step 3 in EBM: appraisal

1. Formulate an answerable question
2. Track down the best evidence
3. Critically appraise the evidence for:
   - 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

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



CEBM CENTRE FOR EVIDENCE BASED MEDICINE

Google

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?

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

QUESTION:

Participants

Intervention Group (IG)
& Comparison Group
(CG)

Outcome

VALIDITY

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

QUESTION:

Participants

Intervention Group (IG) & Comparison Group (CG)

Outcome

	+	–
+	A	B
–	C	D

VALIDITY

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

QUESTION:

Participants

Intervention Group (IG) & Comparison Group (CG)

Outcome

	+	–
+	A	B
–	C	D

VALIDITY

Recruitment

Allocation concealment? comparable groups?

Maintenance

treated equally? compliant?

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

QUESTION:

Participants

Intervention Group (IG) & Comparison Group (CG)

Outcome

	+	–
+	A	B
–	C	D

VALIDITY

Recruitment

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

QUESTION:

Participants

Intervention Group (IG) & Comparison Group (CG)

Outcome

	+	–
+	A	B
–	C	D

VALIDITY

Recruitment

Allocation concealment? comparable groups?

Maintenance

treated equally? compliant?

Measurements blind? OR objective?

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

QUESTION:

Participants

Intervention Group (IG) & Comparison Group (CG)

Outcome

	+	–
+	A	B
–	C	D

VALIDITY

Recruitment

Allocation concealment? comparable groups?

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

QUESTION: DESIGN: VALIDITY

Participants — Recruitment?

Allocation?

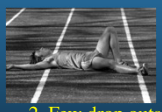
Intervention Group (IG) & Comparison Group (CG) — I G / C G

Maintenance of allocation?

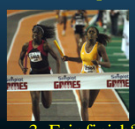
Outcome — A B C D

Measurement of outcomes?

1. Fair start?



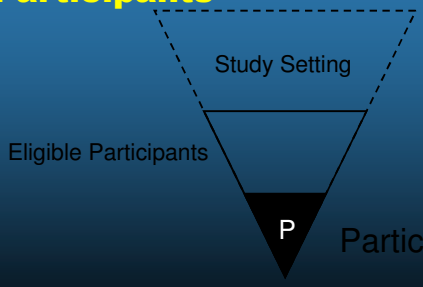
2. Few drop outs?



3. Fair finish?

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

Participants



CEBM CENTRE FOR EVIDENCE BASED MEDICINE

Using the PICO to orient us

Clinical Question

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

ARTICLES

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

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

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

Participants

Study Setting: volunteers, UK, ? 1990s

Eligible Participants: no previous DVT, > 50 yrs, planned economy air travel 2 sectors > 8 hours

P

Participants: 200, mean age 61-62 years

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

Use the **RAMMbo** to check validity

Was the Study valid?

1. **Recruitment**
   - 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**aintainence
   - Were the groups treated equally?
   - Were outcomes ascertained & analysed for most patients?
4. **M**easurements blinded OR objective
   - Were patients and clinicians "**b**linded" to treatment? OR
   - Were measurements **o**bjective & standardised?

User Guide. JAMA, 1993

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

Appraisal checklist - **RAMMbo**

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 the groups treated equally?
   - Were outcomes ascertained & analysed for most patients?
4. **M**easurements
   - Were patients and clinicians "blinded" to treatment? OR
   - Were measurements objective & standardised?

Study statistics (p-values & confidence intervals)

Guyatt. JAMA, 1993

ARTICLES

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

...ily McDonald, Philip D Coleridge Smith

Randomisation
Volunteers were randomised by sealed envelope to one of two groups.

...(DVT) during long-haul air travel is unknown. ...limb during long-haul economy-class air ...ockings 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 flight (median total duration 24 h), returning to the UK within 6 weeks. 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 compression stockings, and two were heterozygous for FVL. Four further patients who wore elastic compression stockings, had varicose veins and developed superficial thrombophlebitis. One of these passengers was heterozygous for both FVL and PGM. None of the passengers who wore class-I compression stockings developed DVT (95% CI 0–3·2%).

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

Lancet 2001; 357: 1485–89 See Commentary page 1461

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

Intervention & Comparison Groups

Intervention Group
Below knee compression stockings

115 116
100 100

Comparison or Control Group (CG): no stockings

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

Benefits of Randomisation (and Allocation Concealment)

- Minimises confounding - known and unknown potential confounders are evenly distributed between study groups
  - reduces bias in those selected for treatment
  - guarantees treatment assignment will not be based on patients' prognosis

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

Fair Allocation – balance achieved?

Were the groups similar at the start?

- ✓ Usually Table 1 in Results section
- ✓ Do imbalances favour one treatment?

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 (×10⁹/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 (×10⁹/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.

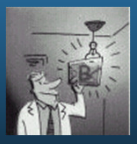
Table 1: **Characteristics of study groups**

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

Allocation Concealment

BEST – most valid technique

- Central computer randomization



DOUBTFUL

- Envelopes, etc

NOT RANDOMISED

- Date of birth, alternate days, etc – WHY?

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

Appraisal checklist - **RAMMbo**

Study biases

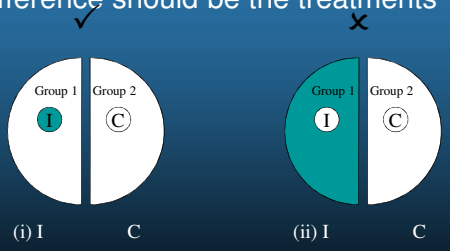
1. **R**ecruitment
   - Who did the subjects represent?
2. Allocation
   - Was the assignment to treatments randomised?
   - Were the groups similar at the trial's start?
3. **Maintenance**
   - **Were the groups treated equally?**
   - **Were outcomes ascertained & analysed for most patients?**
4. **Measurements**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

Study statistics (p-values & confidence intervals)

Guyatt. JAMA, 1993

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

Comparable Groups: the only difference should be the treatments



(i) I C (ii) I C

Is the difference between I and C because of (i) the intervention or (ii) because the groups were not comparable in the first place?

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

Maintaining the Randomisation

- Principle 1 (Intention to treat)
  - Once a patient is randomised, s/he should be analysed in the group randomised to - even if they discontinue, never receive treatment, or crossover.
- Principle 2 (adequate follow up)
  - "5-and-20 rule of thumb"
  - 5% probably leads to little bias
  - >20% poses serious threats to validity

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

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

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

Follow-up in DVT study?

- 231 randomised (115 to stockings; 116 none)
- 200 analysed
  - 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 1486

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

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

Equal treatment in DVT study?

	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***

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

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

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

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

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. 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

ARTICLES

Frequency and prevention of symptomless deep-vein thrombosis in long-haul flights

John H Scurr, Samuel J Machin, Sa

Summary

Background The true freque
We sought to determine the fr
travel and the efficacy of grad

Methods We recruited 89 ma
thromboembolic problems. Pa
class-I below-knee graduated
made journeys lasting more than 8 h
weeks. 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 compression stockings, and two were heterozygous for FVL. Four further patients who wore elastic compression stockings, had varicose veins and developed superficial thrombophlebitis. One of these passengers was heterozygous for both FVL and PGM. None of the passengers who wore class-I compression stockings developed DVT (95% CI 0–3·2%).

Lancet 2001; 357: 1485–89 See Commentary page 1461

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 randomised

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

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

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

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

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 "**b**linded" to treatment? OR
   - Were measurements **o**bjective & standardised?
5. **Placebo Effect**
6. Chance
7. Real Effect

Study statistics (p-values & confidence intervals)

Guyatt. JAMA, 1993

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

Measurement Bias

- Objective
- Blinded?
  - Participants?
  - Investigators?
  - Outcome assessors?
  - Analysts?
- Papers should report WHO was blinded and HOW it was done


Figure 1: The authors: double blinded versus single blinded

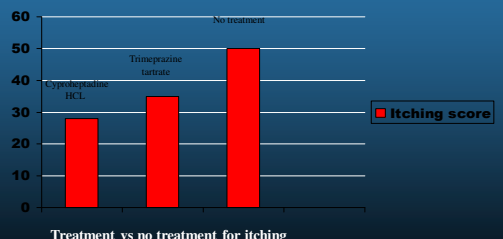

Figure 2: The authors blinded and masked

Schulz and Grimes. Lancet, 2002

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

Placebo effect

Trial in patients with chronic severe itching

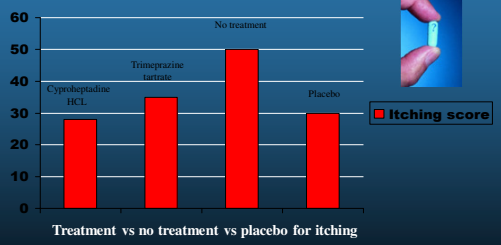


Treatment vs no treatment for itching

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

Placebo effect

Trial in patients with chronic severe itching



Treatment vs no treatment vs placebo for itching

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

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

Wonder Drug Trial

- New drug for stroke
- 100 patients randomised to get Wonder drug or standard care
- 50 patients get Wonder drug
- 50 patients get standard care

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

How do we measure differences?

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

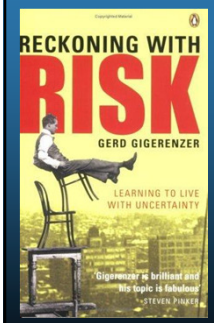
Wonder Drug Trial

- RESULTS

WD group – 5 patients die

SC group – 25 patients die

CEBM CENTRE FOR EVIDENCE BASED MEDICINE



RISK

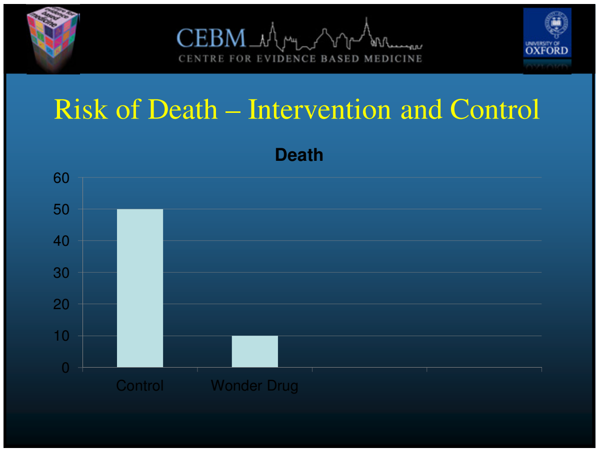
- Chance that something bad will happed
- Different ways of measuring it
- Choice depends on how you want to spin your results

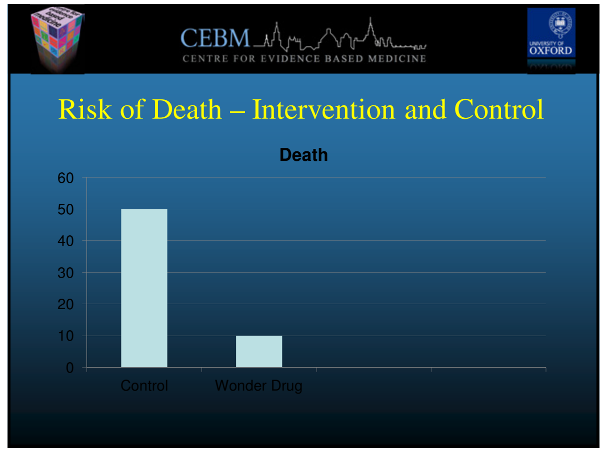
CEBM CENTRE FOR EVIDENCE BASED MEDICINE

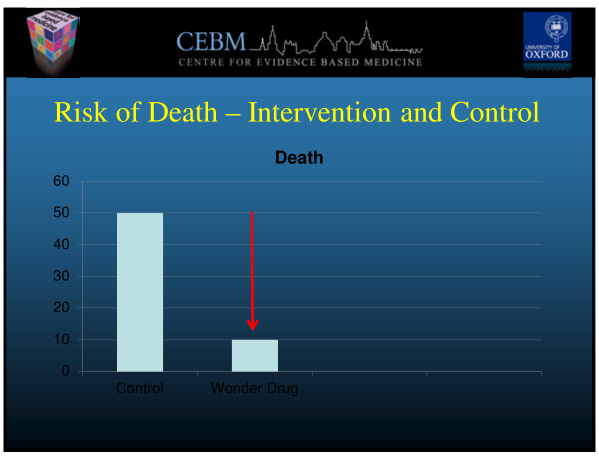
RISKS

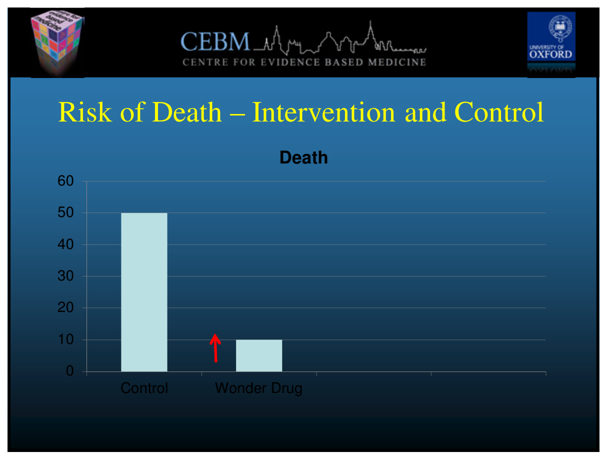
Risk of death in WD group = 5/50
= 1/10
= 10%

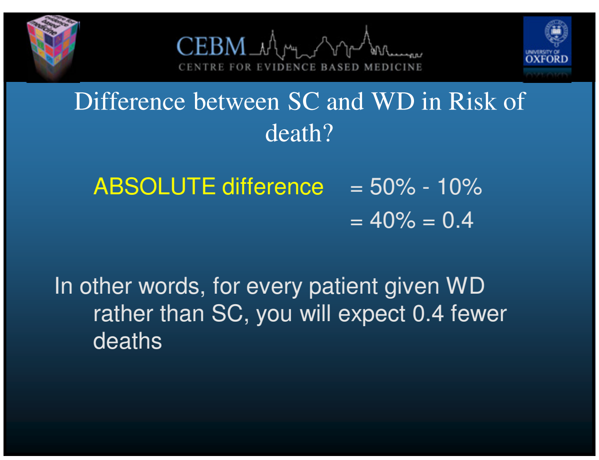
Risk of death in SC group = 25/50
= ½
= 50%

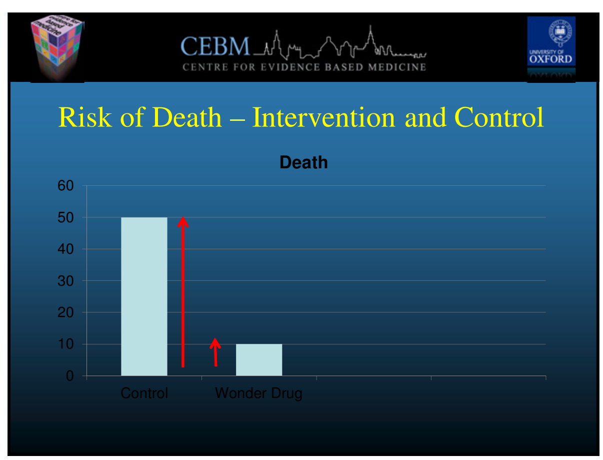

CEBM
CENTRE FOR EVIDENCE BASED MEDICINE
OXFORD
Risk of Death – Intervention and Control
Death
60
50
40
30
20
10
0
Control
Wonder Drug


CEBM
CENTRE FOR EVIDENCE BASED MEDICINE
OXFORD
Risk of Death – Intervention and Control
Death
60
50
40
30
20
10
0
Control
Wonder Drug


CEBM
CENTRE FOR EVIDENCE BASED MEDICINE
OXFORD
Risk of Death – Intervention and Control
Death
60
50
40
30
20
10
0
Control
Wonder Drug


CEBM
CENTRE FOR EVIDENCE BASED MEDICINE
OXFORD
Risk of Death – Intervention and Control
Death
60
50
40
30
20
10
0
Control
Wonder Drug


CEBM
CENTRE FOR EVIDENCE BASED MEDICINE
OXFORD
Difference between SC and WD in Risk of death?
ABSOLUTE difference = 50% - 10%
= 40% = 0.4
In other words, for every patient given WD rather than SC, you will expect 0.4 fewer deaths


CEBM
CENTRE FOR EVIDENCE BASED MEDICINE
OXFORD
Risk of Death – Intervention and Control
Death
60
50
40
30
20
10
0
Control
Wonder Drug

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

Difference between SC and WD in Risk of death?

How much risk is there in the treatment group as a percentage of original (control) risk

RELATIVE RISK = risk in WD (10%) / risk in SC (50%)

= 1/5

= 20%

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

Difference between SC and WD in Risk of death?

RELATIVE RISK REDUCTION = (50% - 10%) / 50%

= 40/50 = 80%

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

Risk of Death – Intervention and Control

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

Difference between SC and WD in Risk of death?

- Number needed to treat NNT
- How many people would you expect to have to treat with WD rather than SC in order to prevent 1 death?
- The Absolute risk reduction gives us the number of events prevented per patient given WD rather than SC

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

Difference between SC and WD in Risk of death?

How much of control risk has been taken away, as a percentage of that original control risk?

RELATIVE RISK REDUCTION = (50% - 10%) / 50%

= 40/50 = 80%

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

Difference between SC and WD in Risk of death?

- So to get the number of patients treated to prevent 1 event

= 1 / Absolute Risk Reduction

= 1/(40/100)

= 2.5 (round up to 3)

(ARR is events prevented/patient, for the NNT, we need patients/event prevented)

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

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**aintainence
   - Were the groups treated equally?
   - Were outcomes ascertained & analysed for most patients?
4. **M**easurements
   - Were patients and clinicians "**b**linded" to treatment? OR
   - Were measurements **o**bjective & standardised?
5. **Placebo Effect**
6. **Chance**
7. Real Effect

Study statistics (p-values & confidence intervals)

Guyatt. JAMA, 1993

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

Confidence Intervals (Estimation) - in DVT study

- 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

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

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

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

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

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

P-values (Hypothesis Testing) - in DVT study

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

Absolute risk reduction = 0.12 - 0 = 0.12
(P=0.001)

The probability that this result would only occur by chance is
1 in 1000 → statistically significant

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

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**