Rapid Critical Appraisal of Randomised Controlled Trials

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

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?

QUESTION:

Participants
- Intervention Group (IG)
- Comparison Group (CG)

Outcome
APPRAISAL OF RCTs

QUESTION: Participants
- Intervention Group (IG) & Comparison Group (CG)

OUTCOME
- A B
- C D

VALIDITY

QUESTION: Recruitment
- Allocation concealment? comparable groups?
- Maintenance treated equally? compliant?

QUESTION: Allocation concealment?
- comparable groups?

QUESTION: Maintenance
- treated equally? compliant?

QUESTION: Measurements
- blind? OR objective?
Participants

Intervention Group (IG) & Comparison Group (CG)

Outcome

Allocation?

Maintenance of allocation?

Measurement of outcomes?

QUESTION:

Study Setting: volunteers, UK, ? 1990s

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

Participants: 200, mean age 61-62 years

Using the PICO to orient us

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

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

Appraisal checklist - RAMMbo

Was the Study valid?

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 blinded OR objective
   - Were patients and clinicians “blinded” to treatment? OR
   - Were measurements objective & standardised?

User Guide. JAMA, 1993

Study biases
1. Recruitment
   - Who did the subjects represent?
2. Allocation
   - Was the assignment to treatments randomised?
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
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 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%).

Lancet 2001; 357: 1485–89

See Commentary page 1461

Randomisation

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

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%).

Lancet 2001; 357: 1485–89

See Commentary page 1461

Randomisation

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

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%).

Lancet 2001; 357: 1485–89

See Commentary page 1461

Randomisation

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

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%).

Lancet 2001; 357: 1485–89

See Commentary page 1461

Randomisation

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

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%).

Lancet 2001; 357: 1485–89

See Commentary page 1461

Randomisation

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

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%).

Lancet 2001; 357: 1485–89

See Commentary page 1461

Randomisation

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

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%).

Lancet 2001; 357: 1485–89

See Commentary page 1461
Comparable Groups: the only difference should be the treatments

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

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

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

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

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

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

<table>
<thead>
<tr>
<th></th>
<th>No Stockings</th>
<th>Stockings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspirin</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>Hormone replacement therapy</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>Thyroxine</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Antihypertensives, including diuretics</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>Antipptic ulcer drugs</td>
<td>8</td>
<td>3</td>
</tr>
</tbody>
</table>

*Includes additions to usual drugs

Scurr et al, Lancet 2001; 357:1485-89
**Measures in DVT study?**
- Blood was taken from all participants before travel
- All participants had US 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


**Appraisal checklist**

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?

Guyatt. JAMA, 1993

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

**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. Made journeys lasting more than 4 h in duration before return to the UK within 6 weeks.

**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 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 20210A (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%).


**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 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 20210A (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%).

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

Schulz and Grimes. Lancet, 2002
Placebo effect
Trial in patients with chronic severe itching

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

How do we measure differences?

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

Wonder Drug Trial
• RESULTS
WD group – 5 patients die
SC group – 25 patients die

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

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

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.
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?

\[
\text{RELATIVE RISK} = \frac{\text{risk in WD (10\%)}}{\text{risk in SC (50\%)}}
\]

\[
= \frac{1}{5}
\]

\[
= 20\%
\]

Difference between SC and WD in Risk of death?

\[
\text{RELATIVE RISK REDUCTION} = \frac{50\% - 10\%}{50\%}
\]

\[
= \frac{40}{50} = 80\%
\]

Risk of Death – Intervention and Control

- 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

Difference between SC and WD in Risk of death?

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

\[
\text{Absolute Risk Reduction} = \frac{1}{40/100}
\]

\[
= \frac{1}{0.4} = 2.5 \text{ (round up to 3)}
\]

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

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?

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

**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

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

**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