APPRAISAL OF CLINICAL TRIALS

Su May Liew
CEBM Oxford
‘Good study’

- Experimental - RCT
- Fit question - PICO
- Passes appraisal - RAMMbo
- Minimises random error - statistics
Why a trial?

It may be better to do nothing at all!
Clinical question

- Population
- Intervention
- Comparator
- Outcome/s
Clinical question

• Population: Adults with URTI and purulent nasal discharge
• Intervention: Antibiotics
• Comparator: No Antibiotics
• Outcome/s: Reduction of duration of symptoms, reduction of severity of symptoms, etc
Participants

Intervention Group (IG) & Comparison Group (CG)

Outcome

QUESTION:

DESIGN:

Selection?

Allocation?

Maintenance of allocation?

Measurement of outcomes?

Allocation?

Randomised?

comparable groups?

Measurements blind subjective?

OR objective?

Allocation?

Representative?

Allocation?

Selection?
Participants

Intervention Group (IG) & Comparison Group (CG)

Outcome

Allocation?

Maintenance of allocation?

Measurement of outcomes?

QUESTION:

VALIDITY

1. Fair start?

2. Few drop outs?

3. Fair finish?
Was it a fair race?

1. Fair start?

2. Few drop outs?

3. Fair finish?
Use RAMMbo to check validity

**Representative**
- Who did the subjects represent?

**Allocation**
- Was the assignment to treatments randomised?
- Were the groups similar at the trial’s start?

**Maintenance**
- Were the groups treated equally?
- Were outcomes ascertained & analysed for most patients?

**Measurements blinded OR objective**
- Were patients and clinicians “blinded” to treatment? OR
- Were measurements objective & standardised?
Does Amoxicillin Improve Outcomes in Patients with Purulent Rhinorrhea?

A Pragmatic Randomized Double-Blind Controlled Trial in Family Practice

An I. De Sutter, MD; Marc J. De Meyere, MD, PhD; Thierry C. Christiaens, MD; Mieke L. van Driel, MD, MSc; Wim Peersman; and Jan M. De Maeseneer, MD, PhD

Ghent, Belgium
Use RAMMbo to check validity

Representative
• Who did the subjects represent?

Allocation
• Was the assignment to treatments randomised?
• Were the groups similar at the trial’s start?

Maintenance
• Were the groups treated equally?
• Were outcomes ascertained & analysed for most patients?

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

User Guide. JAMA, 1993
scribe antibiotics.\textsuperscript{23-26} The trial was designed as a pragmatic effectiveness trial. Patient inclusion and evaluation were defined on a purely clinical basis to maximize relevance for routine daily practice.

**METHODS**

**Study Population**

Between October 1998 and December 1999, 69 FPs in Flanders, Belgium, agreed to enroll patients meeting the following inclusion criteria: age 12 years or older, presenting with a respiratory tract infection, and having purulent rhinorrhea. Exclusion criteria were allergy to penicillin or ampicillin; having received antibiotic therapy within the previous week; complaints lasting for more than 30 days; abnormality on clinical chest examination; complications of sinusitis (facial edema or cellulitis; orbital, visual, meningeal, or cerebral signs)\textsuperscript{27}; pregnancy or lactation; comorbidity that might impair immune competence; and inability to follow the protocol because of language or mental problems. The Ethics Committee of the Ghent University Hospital (GUH) approved the study. All patients (or their guardians, for those younger than 16 years of age) gave written informed consent.
Participants

Study Setting: FP, Flanders, Belgium, 1999

Eligible Participants: >12 years, RTI + purulent rhinorrhea

Participants: 416 randomised, 8 excluded, 34 withdrew leaving 374 with f/up data
Use **RAMMbo** to check validity

**Representative**
- Who did the subjects represent?

**Allocation**
- Was the assignment to treatments randomised?
- Were the groups similar at the trial’s start?

**Maintenance**
- Were the groups treated equally?
- Were outcomes ascertained & analysed for most patients?

**Measurements blinded OR objective**
- Were patients and clinicians “blinded” to treatment?
  OR
- Were measurements objective & standardised?
Comparable Groups: the only difference should be the treatment/s

Is the difference between I and C because of (i) the intervention or (ii) because the groups were not comparable in the first place?
Fair Allocation
How do we get comparable groups?

Was assignment to treatments randomised?
• Was the allocation process tamper proof?
  AND
• Were the groups similar at start of trial?
Treatment Assignment and Masking

In this double-blind trial, patients were assigned via a computer-generated random number list to receive 500 mg amoxicillin 3 times a day or placebo for 10 days. The trial medication was supplied in numbered uniform cardboard boxes, each containing 30 capsules of the same size, color, and shape for active and placebo treatment. The randomization list, kept at the pharmacy of GUH, was accessible to the participating FPs only in case of a serious adverse event.

To assess the effectiveness of masking, patients and their FPs guessed the treatment group at 10-day follow-up. Data were encoded and entered without knowledge of treatment allocation. Compliance was assessed by counting leftover medication. All patients were allowed to use xylometazoline 1% nose drops and paracetamol or ibuprofen to alleviate symptoms; these data were registered.
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
Allocation Concealment

**BEST** – most valid technique
Central computer randomization

**DOUBTFUL**
Envelopes, etc

**NOT RANDOMISED**
Date of birth, alternate days, etc
<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>BASELINE CHARACTERISTICS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>General</strong> (placebo = 205, amoxicillin = 204)</td>
<td></td>
</tr>
<tr>
<td>Mean age (SD)</td>
<td>39 (15)</td>
</tr>
<tr>
<td>Mean days of complaint before contact (SD)</td>
<td>7.2 (5.5)</td>
</tr>
<tr>
<td>Women (%)</td>
<td>54</td>
</tr>
<tr>
<td><strong>Mean Score on SNOT-20</strong> (placebo = 196, amoxicillin = 192)</td>
<td>40.8 (SD 15.9)</td>
</tr>
<tr>
<td><strong>History</strong> (placebo = 196, amoxicillin = 192)</td>
<td></td>
</tr>
<tr>
<td>Generally ill to very ill (%)</td>
<td>46</td>
</tr>
<tr>
<td>Unilateral facial pain (%)</td>
<td>56</td>
</tr>
<tr>
<td>Pain on bending forward (%)</td>
<td>70</td>
</tr>
<tr>
<td>Pain in upper teeth or when chewing (%)</td>
<td>44</td>
</tr>
<tr>
<td><strong>Examination</strong> (placebo = 209, amoxicillin = 207)</td>
<td></td>
</tr>
<tr>
<td>Sinus tenderness (%)</td>
<td>61</td>
</tr>
<tr>
<td>Pain on bending forward (%)</td>
<td>60</td>
</tr>
<tr>
<td>Postnasal discharge on throat inspection (%)</td>
<td>55</td>
</tr>
<tr>
<td>Purulent rhinorrhea on rhinoscopy (%)</td>
<td>47</td>
</tr>
<tr>
<td>Body temperature &gt; 37°C (%)</td>
<td>38</td>
</tr>
</tbody>
</table>

SD denotes standard deviation; SNOT, Sino-Nasal Outcome Test.
Use **RAMMbo** to check validity

**Representative**
- Who did the subjects represent?

**Allocation**
- Was the assignment to treatments randomised?
- Were the groups similar at the trial’s start?

**Maintenance**
- Were the groups treated equally?
- Were outcomes ascertained & analysed for most patients?

**Measurements blinded OR objective**
- Were patients and clinicians “blinded” to treatment? **OR**
- Were measurements objective & standardised?
Effects of non-equal treatment

Apart from actual intervention - groups should receive identical care!

*Trial of Vitamin E in pre-term infants (1948)*
Vit E "prevented" retrolental fibroplasia
(By removal from 100% Oxygen to give the frequent doses of Vit E!)

Rx: Give placebo in an identical regime, and a standard protocol
Crime against Randomisation
Crossovers and Drop-outs
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 this study?

- **416 randomised** (207 amoxycillin; 209 placebo)
- **8 excluded; leaving 408**
- **34 withdrawn; leaving 374 (187 each arm)**
  - 8 clinical exacerbations
  - 2 complete recovery
  - 2 concurrent pathology
  - 5 allergic reactions
  - 1 side effect
  - 16 lost to follow-up
Use **RAMMbo** to check validity

**Representative**
- Who did the subjects represent?

**Allocation**
- Was the assignment to treatments randomised?
- Were the groups similar at the trial’s start?

**Maintenance**
- Were the groups treated equally?
- Were outcomes ascertained & analysed for most patients?

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

User Guide. JAMA, 1993
Measurement Bias - minimizing differential error

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

Schulz and Grimes. Lancet, 2002
<table>
<thead>
<tr>
<th>Symptom</th>
<th>Mean Score Reduction Amoxicillin n = 170</th>
<th>Mean Score Reduction Placebo n = 164</th>
<th>P *</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unilateral facial pain</td>
<td>1</td>
<td>1.1</td>
<td>.56</td>
</tr>
<tr>
<td>Pain on bending forward</td>
<td>1.21</td>
<td>1.32</td>
<td>.55</td>
</tr>
<tr>
<td>Pain in upper teeth or when chewing</td>
<td>0.7</td>
<td>0.93</td>
<td>.17</td>
</tr>
<tr>
<td>Need to blow nose</td>
<td>1.73</td>
<td>1.70</td>
<td>.85</td>
</tr>
<tr>
<td>Sneezing</td>
<td>1.13</td>
<td>1.05</td>
<td>.63</td>
</tr>
<tr>
<td>Runny nose</td>
<td>1.47</td>
<td>1.55</td>
<td>.33</td>
</tr>
<tr>
<td>Cough</td>
<td>1.0</td>
<td>1.11</td>
<td>.46</td>
</tr>
<tr>
<td>Thick nasal discharge</td>
<td>2.2</td>
<td>1.5</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Postnasal discharge</td>
<td>1.29</td>
<td>1.09</td>
<td>.26</td>
</tr>
<tr>
<td>Ear fullness</td>
<td>1.13</td>
<td>1.31</td>
<td>.32</td>
</tr>
<tr>
<td>Dizziness</td>
<td>0.95</td>
<td>0.87</td>
<td>.63</td>
</tr>
<tr>
<td>Ear pain</td>
<td>0.64</td>
<td>0.77</td>
<td>.36</td>
</tr>
<tr>
<td>Facial pain or pressure</td>
<td>1.54</td>
<td>1.61</td>
<td>.69</td>
</tr>
<tr>
<td>Difficulty falling asleep</td>
<td>1.14</td>
<td>1.26</td>
<td>.54</td>
</tr>
<tr>
<td>Wake up at night</td>
<td>1.39</td>
<td>1.44</td>
<td>.79</td>
</tr>
<tr>
<td>Lack of a good night’s sleep</td>
<td>1.24</td>
<td>1.44</td>
<td>.28</td>
</tr>
<tr>
<td>Wake up tired</td>
<td>1.34</td>
<td>1.65</td>
<td>.09</td>
</tr>
<tr>
<td>Fatigue</td>
<td>1.46</td>
<td>1.61</td>
<td>.38</td>
</tr>
<tr>
<td>Reduced productivity</td>
<td>1.45</td>
<td>1.63</td>
<td>.29</td>
</tr>
<tr>
<td>Reduced concentration</td>
<td>1.24</td>
<td>1.46</td>
<td>.19</td>
</tr>
<tr>
<td>Frustrated, restless, irritable</td>
<td>0.87</td>
<td>1.41</td>
<td>.91</td>
</tr>
<tr>
<td>Sad</td>
<td>0.38</td>
<td>0.52</td>
<td>.18</td>
</tr>
<tr>
<td>Embarrassed</td>
<td>0.36</td>
<td>0.76</td>
<td>.36</td>
</tr>
</tbody>
</table>

* Student’s t test.
Treatment Assignment and Masking
In this double-blind trial, patients were assigned via a computer-generated random number list to receive 500 mg amoxicillin 3 times a day or placebo for 10 days. The trial medication was supplied in numbered uniform cardboard boxes, each containing 30 capsules of the same size, color, and shape for active and placebo treatment. The randomization list, kept at the pharmacy of GUH, was accessible to the participating FPs only in case of a serious adverse event.

To assess the effectiveness of masking, patients and their FPs guessed the treatment group at 10-day follow-up. Data were encoded and entered without knowledge of treatment allocation. Compliance was assessed by counting leftover medication. All patients were allowed to use xylometazoline 1% nose drops and paracetamol or ibuprofen to alleviate symptoms; these data were registered.
Fundamental Equation of Error

- Measure = Truth + Bias + Random Error

Use good study design

Use large numbers

Critically Appraise Design

Confidence Intervals and P-values
So do you still want antibiotics?

Wellllll, antibiotics get rid of the purulent discharge
Yes, antibiotics are more effective in changing the colour of the snot BUT it does not change the duration or severity of your symptoms AND you are also more likely to have diarrhoea,