How to teach critical appraisal of RCTs

Associate Professor Dr Su May Liew
University of Malaya
CEBM Oxford

Teaching

Evidence Based Medicine

Su May's principles of teaching – the SENSES method

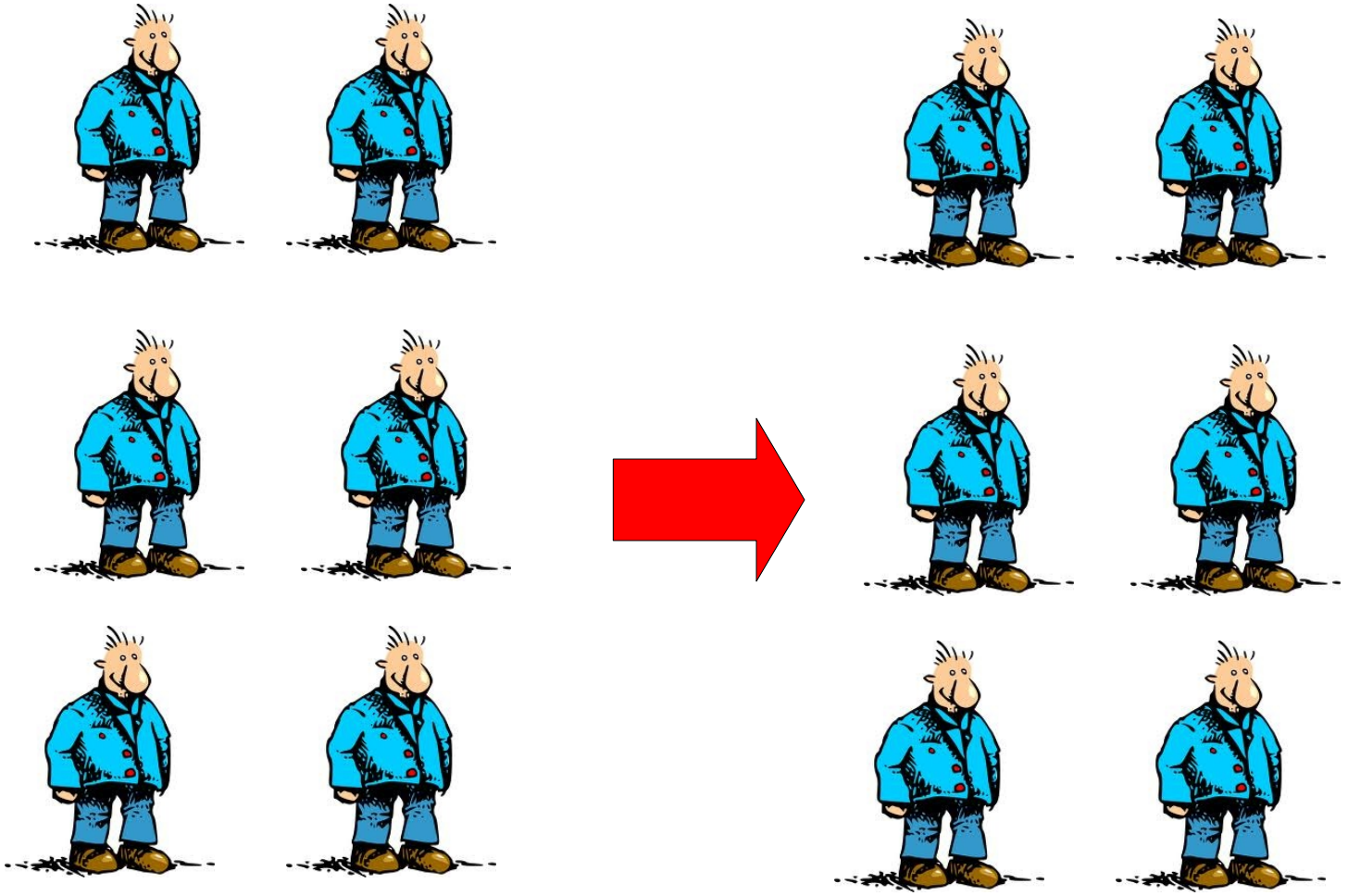
- tell a Story
- teach only Essentials
- create a Need
- Stimulate interest
- give good Examples
- if possible, teach a Skill

it was a dark and stormy night…

APPRAISAL OF RANDOMISED CONTROLLED TRIALS

Su May Liew
University of Malaya
CEBM Oxford

Dad’s proposed study



‘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!

create a Need…

Dibbing

teach only Essentials…

What are the essentials?

- The answerable question
- Relevancy
- Fairness
- Effect of Chance

• • •

Clinical question

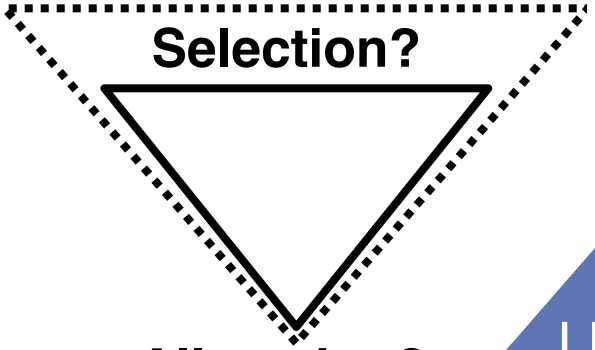
- Population
- Intervention
- Comparator
- Outcome/s

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

QUESTION:

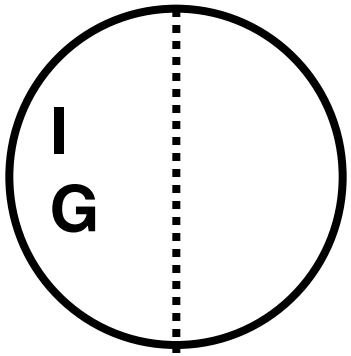
DESIGN:



Participants

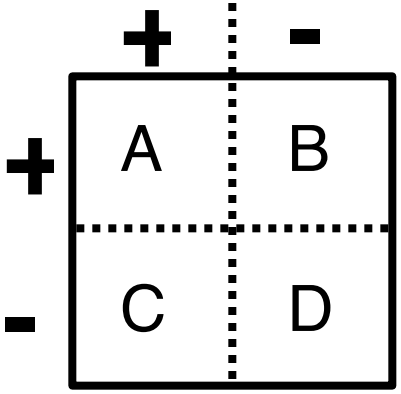
Use the GATE frame

Intervention Group (IG) & Comparison Group (CG)



Maintenance of allocation?

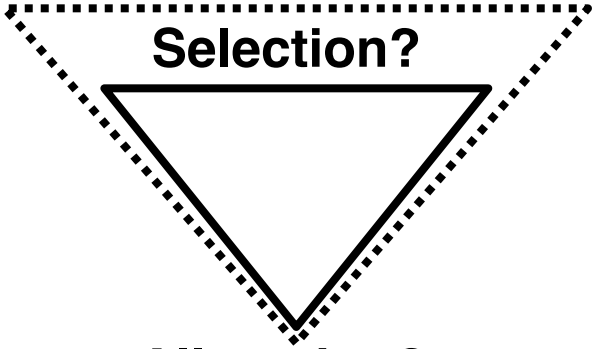
Outcome



Measurement of outcomes?

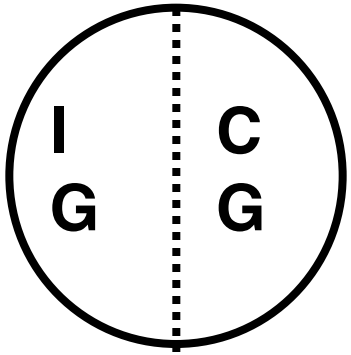
QUESTION: **DESIGN:** **VALIDITY**

Participants



1. Fair start?

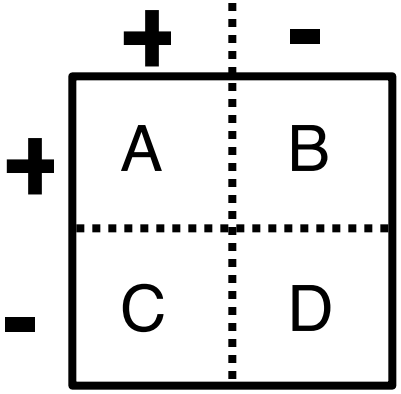
Intervention Group (IG) & Comparison Group (CG)



Maintenance of allocation?

2. Few drop outs?

Outcome

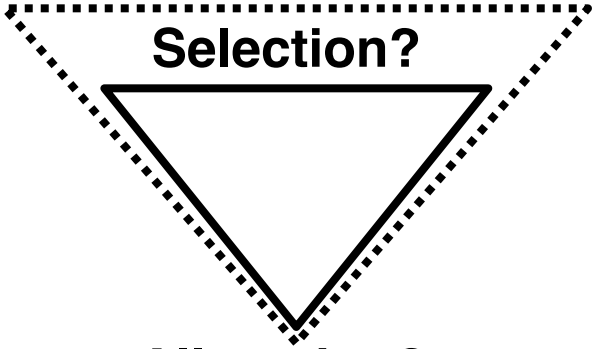


Measurement of outcomes?

3. Fair finish?

QUESTION:

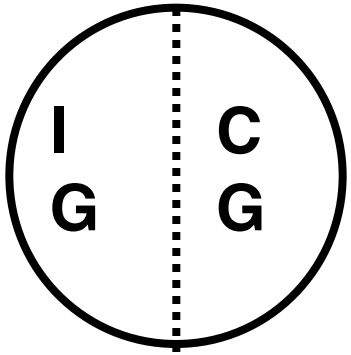
DESIGN:



Participants

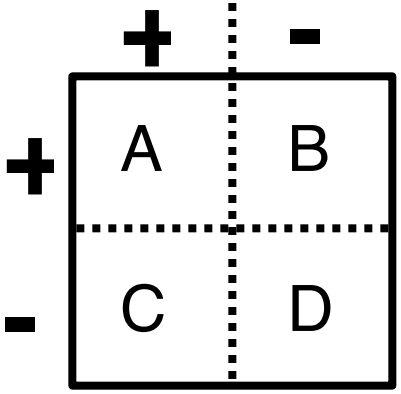
Allocation?

Intervention Group (IG) & Comparison Group (CG)



Maintenance of allocation?

Outcome



Measurement of outcomes?

VALIDITY

Representative?

Allocation?
Randomised?
comparable groups?

Maintenance?
treated equally?
compliant?

Measurements
blind subjective?
OR
objective?

Stimulate!

What is your question?

Search for an RCT

Does the PICO of the RCT fit that of your question?

ORIGINAL RESEARCH

jfponline

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

- Population
- Intervention
- Comparison
- Outcome(s)

Keep it consistent
Use the PICO format

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

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

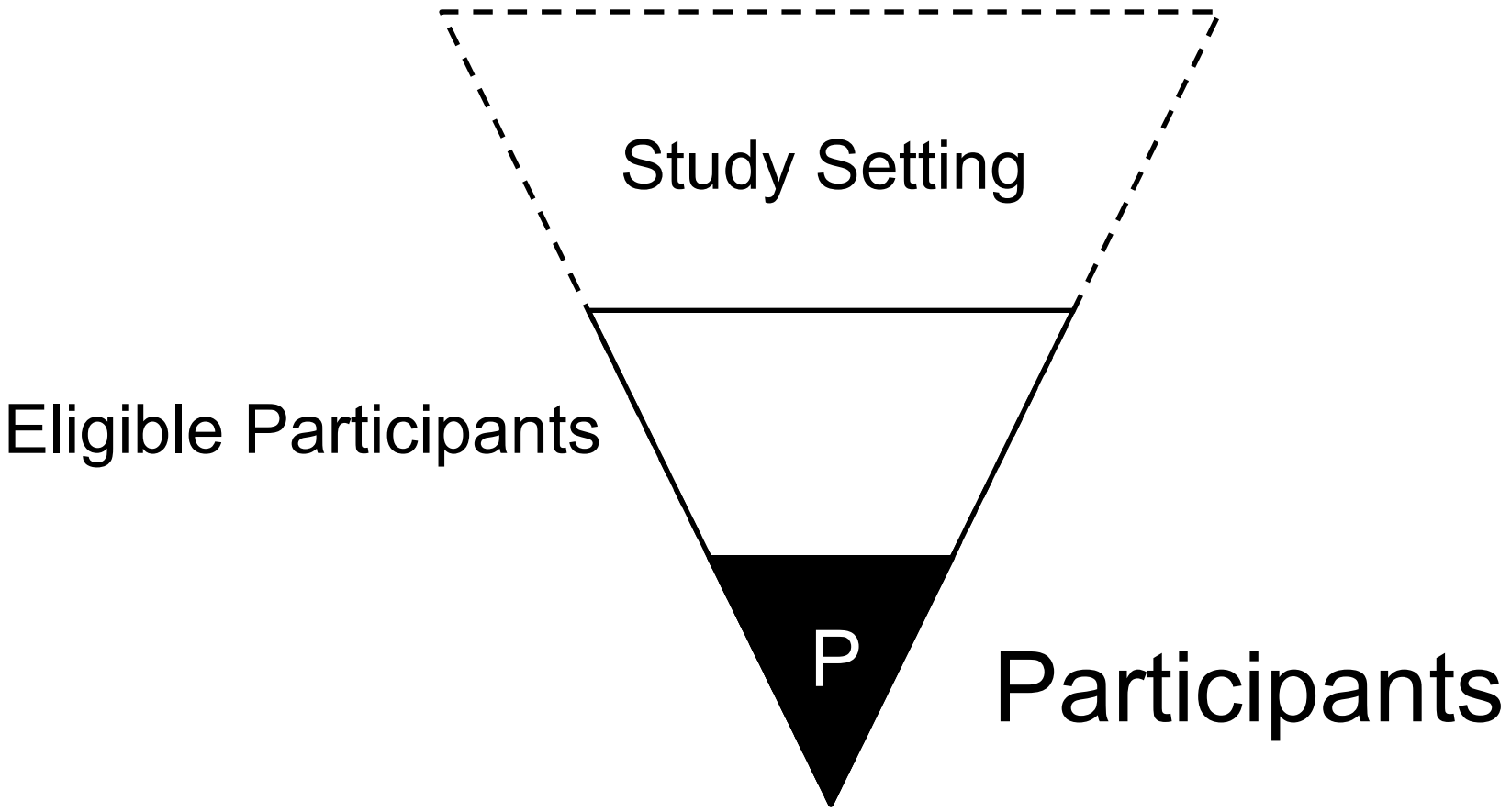
- Were the groups treated equally?
- Were outcomes ascertained & analysed for most patients?

Measurements **b**linded OR **o**bjective

- Were patients and clinicians "**b**linded" to treatment? OR
- Were measurements **o**bjective & standardised?

User Guide. JAMA, 1993

Participants



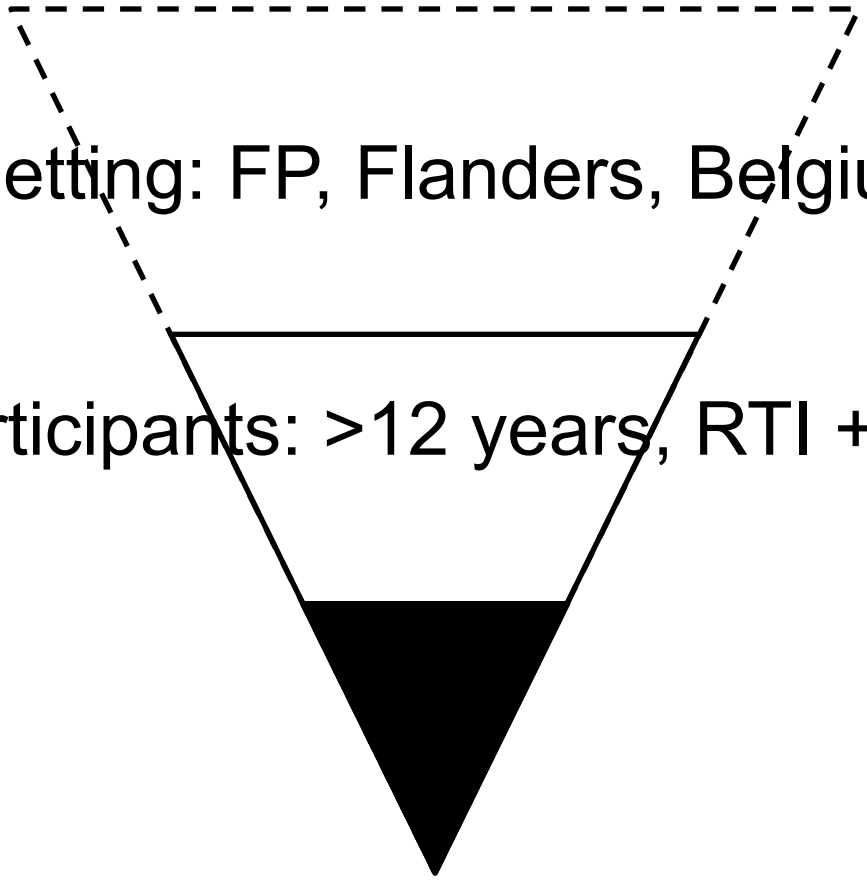
scribe antibiotics.[23-26] The trial was designed as a pragmatic effectiveness trial. Patient inclusion and evaluation were defined on a purely clinical basis to maximize relevance for routine daily practice.

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)[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?

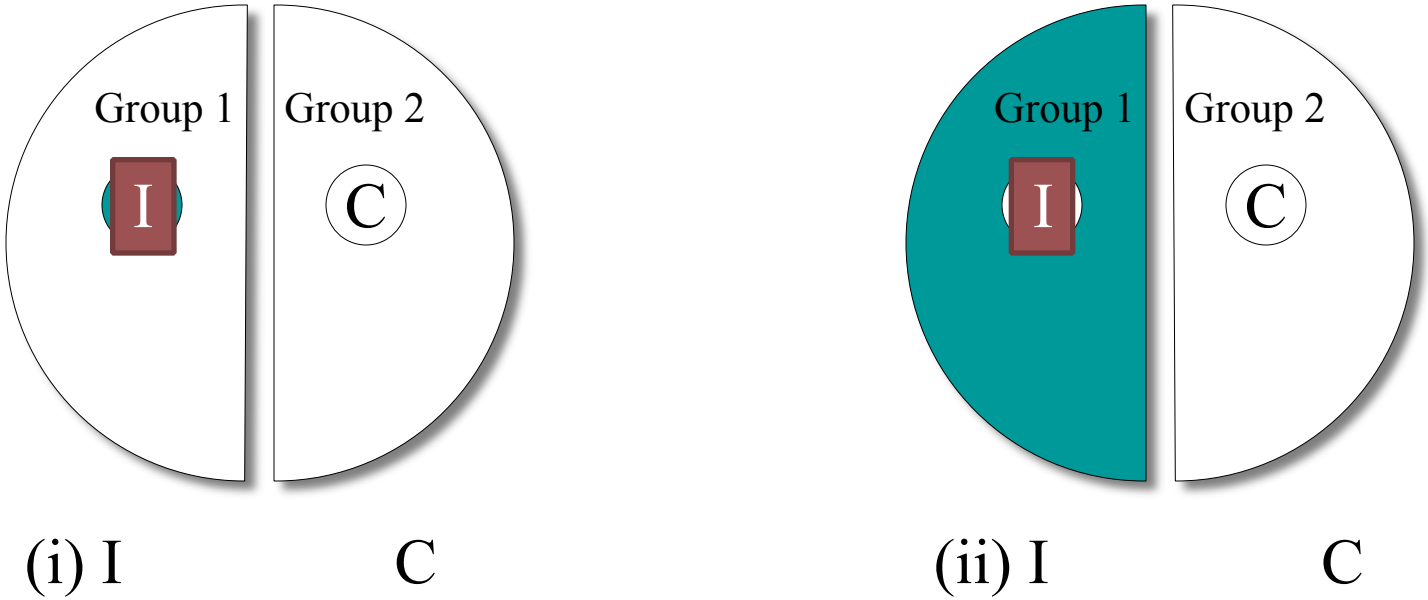
Measurements **b**linded OR **o**bjective

- Were patients and clinicians "**b**linded" to treatment? OR
- Were measurements **o**bjective & standardised?

User Guide. JAMA, 1993

Comparable Groups:
the only difference should be the treatment/s

Highlight this point!



(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?

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

use Examples…

"The only way to get rid of temptation is to yield to it." Oscar Wilde

There was a randomized trial of open vs laparoscopic appendectomy.At night, residents were reluctant to call the consultant needed for the laparoscopic procedure but not the open one. When an eligible patient appeared, the residents held the semiopaque envelopes containing the study assignment up to the light. They opened the first envelope that dictated an open procedure. The first eligible patient in the morning would then be allocated to the laparoscopic appendectomy group according to the passed-over envelope

(D. Wall, written communication, June 2000)

TABLE 1

BASELINE CHARACTERISTICS

General (placebo = 205, amoxicillin = 204)	**Placebo**	**Amoxicillin**
Mean age (SD)	39 (15)	37 (14)
Mean days of complaint before contact (SD	7.2 (5.5)	7.6 (5.4)
Women (%)	54	55
Mean Score on SNOT-20 (placebo = 196, amoxicillin = 192)	40.8 (SD 15.9)	38.4 (SD 16.1)
History (placebo = 196, amoxicillin = 192)		
Generally ill to very ill (%)	46	53
Unilateral facial pain (%)	56	53
Pain on bending forward (%)	70	66
Pain in upper teeth or when chewing (%)	44	41
Examination (placebo = 209, amoxicillin = 207)		
Sinus tenderness (%)	61	67
Pain on bending forward (%)	60	60
Postnasal discharge on throat inspection (%)	55	50
Purulent rhinorrhea on rhinoscopy (%)	47	40
Body temperature > 37°C (%)	38	41

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 **b**linded OR **o**bjective

- Were patients and clinicians "**b**linded" to treatment? OR
- Were measurements **o**bjective & standardised?

User Guide. JAMA, 1993

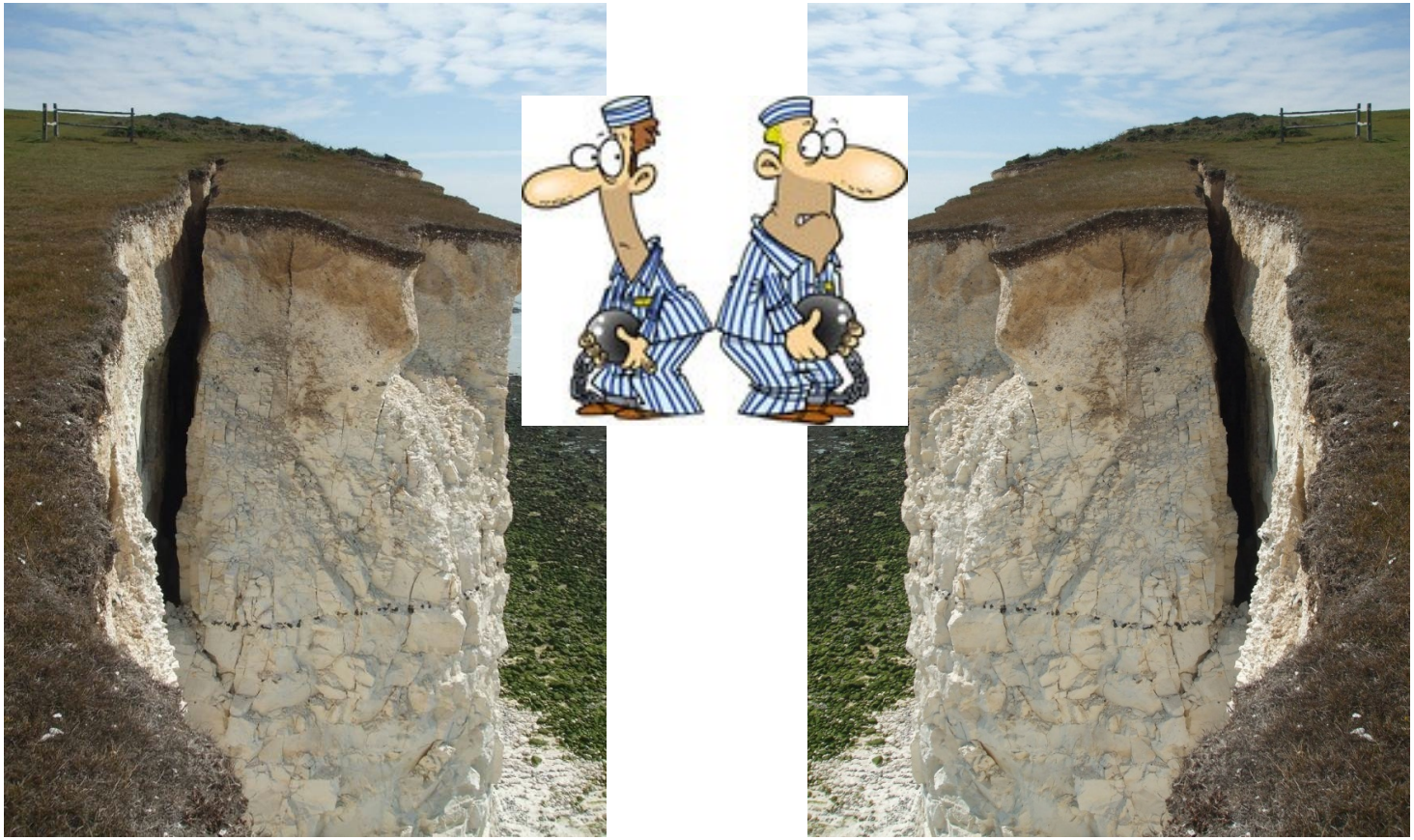
Effects of non-equal treatment

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 **b**linded OR **o**bjective

- Were patients and clinicians "**b**linded" to treatment? OR
- Were measurements **o**bjective & 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

Figure 1: The authors: double blinded versus single blinded

Figure 2: The authors blinded and masked

Schulz and Grimes. Lancet, 2002

TABLE 3

MEAN SYMPTOM CHANGE BETWEEN BASELINE AND 10-DAY FOLLOW-UP

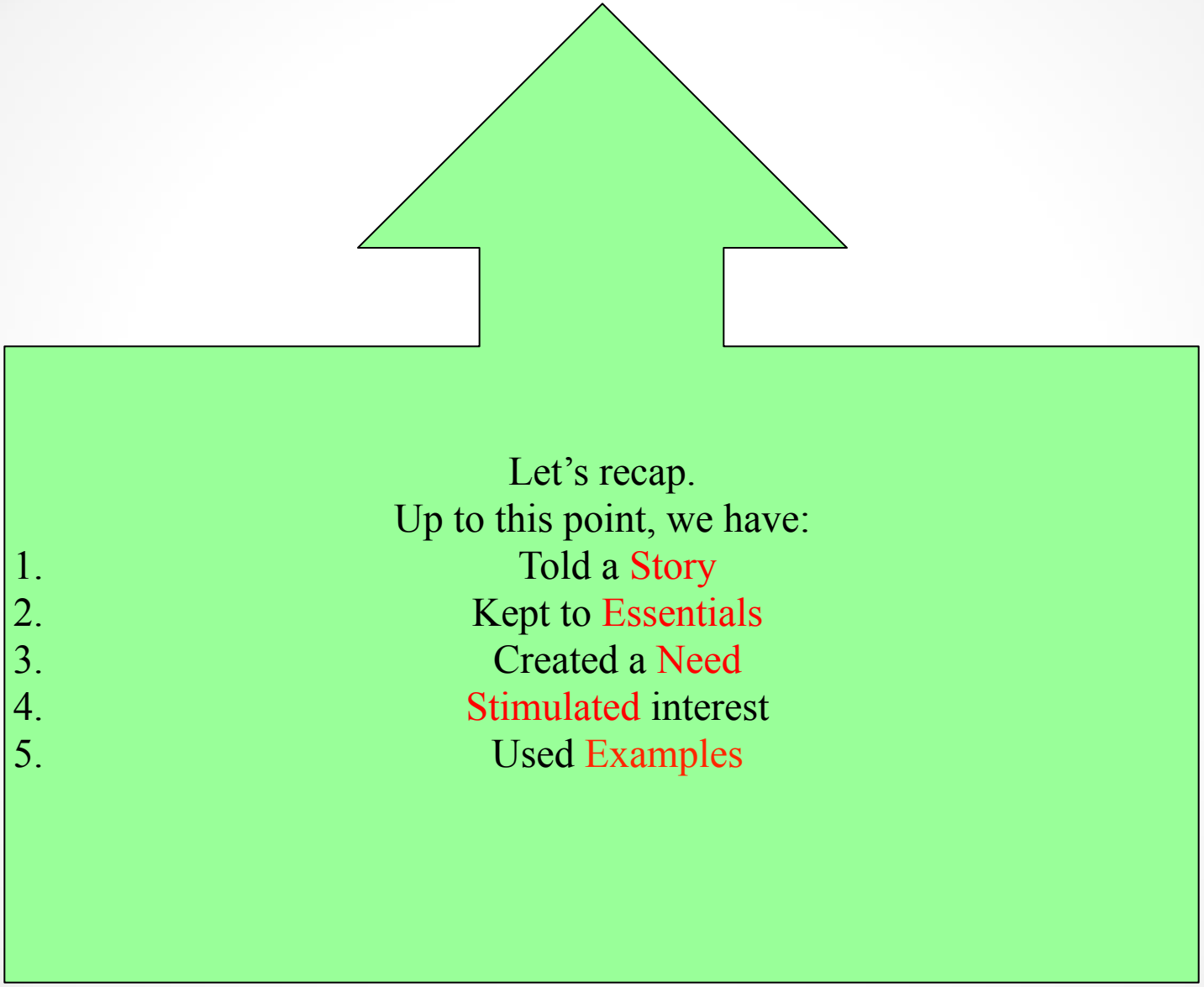
Symptom	Mean Score Reduction		*P* *
	Amoxicillin n = 170	Placebo n = 164	
Unilateral facial pain	1	1.1	.56
Pain on bending forward	1.21	1.32	.55
Pain in upper teeth or when chewing	0.7	0.93	.17
Need to blow nose	1.73	1.70	.85
Sneezing	1.13	1.05	.63
Runny nose	1.47	1.55	.33
Cough	1.0	1.11	.46
Thick nasal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

* 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.


Let's recap.
Up to this point, we have:
1. Told a Story
2. Kept to Essentials
3. Created a Need
4. Stimulated interest
5. Used Examples

Fundamental Equation of Error

Use good study design

Researcher

Use large numbers

- Measure = Truth + **Bias** + *Random Error*

Critically Appraise Design

Reader

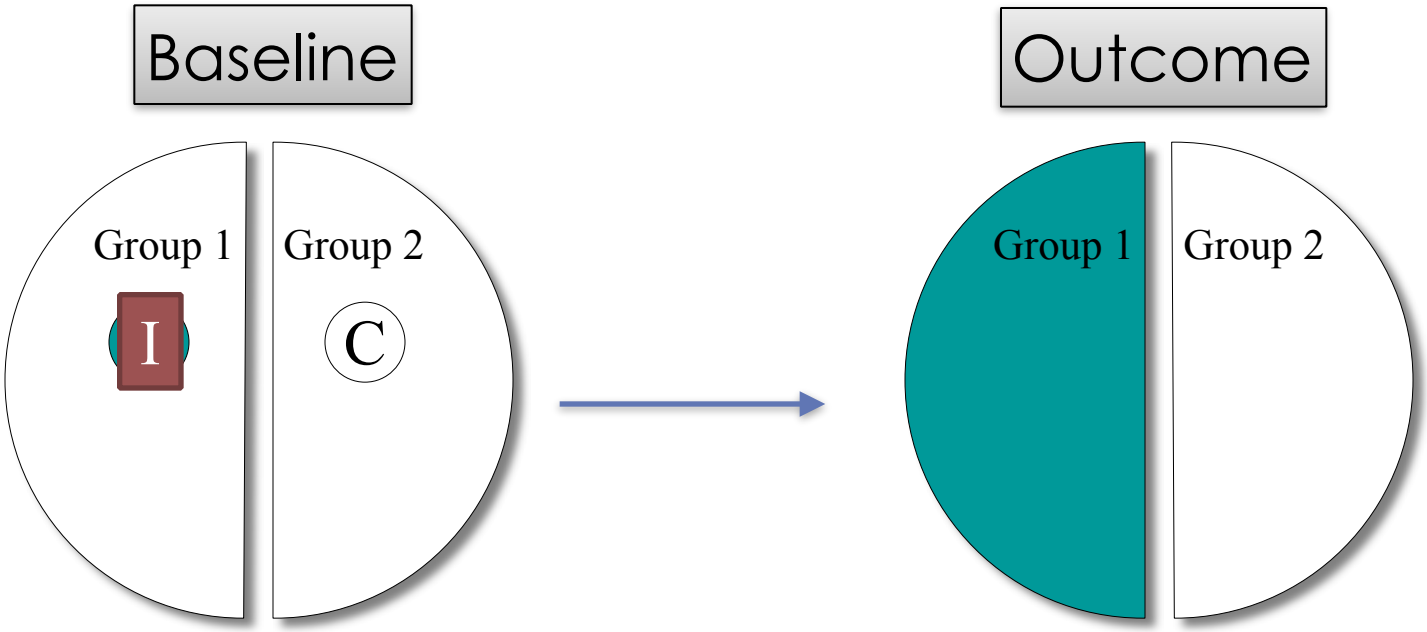
Confidence Intervals and P-values

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

Relationship between p-values and confidence intervals - if the value corresponding to 'no effect' (RR of 1 or treatment difference of 0) falls outside the CI then the result is statistically significant

Comparable Groups where the only difference is the treatment/s



If the difference is large enough that it will only occur by chance less than 5 times if the test is repeated 100 times (p less than 0.05), then we accept that it is statistically significant (not occurring by chance and due to the intervention).

That was the skill

- For RCTs - teaching RAMBO is key. Once the audience can understand how easy it is, they are then empowered to read RCTs.

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,