

Critical appraisal of the medical literature

Jane Hirst

MBBS, MPH, PhD FRANZCOG

University of Oxford

Geneva Workshop 2016



What makes a study reliable?

**NOW, CRACK THE CODE
TO YOUNGER ACTING SKIN.**

**NEW
YOUTH CODE™
Youth Regenerating Skincare**

ONE DROP
INSTANTLY IMPROVES SKIN QUALITY

ONE WEEK
SKIN BEGINS TO LOOK YOUNGER

ONE MONTH
REVEAL THE NEW YOUTH OF YOUR SKIN**

**10 YEARS OF GENE RESEARCH
INTERNATIONAL PATENT**

**L'ORÉAL
PARIS**

**THE NEW ERA OF SKINCARE:
GENE SCIENCE.**

Imagine what if you could turn yourself every great discovery begins by pushing the boundaries of science. After 10 years of research, now we know that recovery genes in youthful skin respond to stress in a different way than aging skin does. So even though you can't grow young, we now have the knowledge to help you begin cracking the code to younger acting skin.

CLINICAL STUDY

**5X
FASTER**

5X FASTER RECOVERY

5X FASTER RECOVERY

Discover all of the Youth Code products and learn more about gene science.

© 2013 L'Oréal Paris. All rights reserved. *See product packaging for details.

Nine out of ten women we asked believe that "Youth Code" makes their skin firmer and younger looking

Which design maximises the chance of getting the result the company wants?

- a) Ask women buying "Youth Code" in the shops whether they agree their skin is firmer and younger looking?

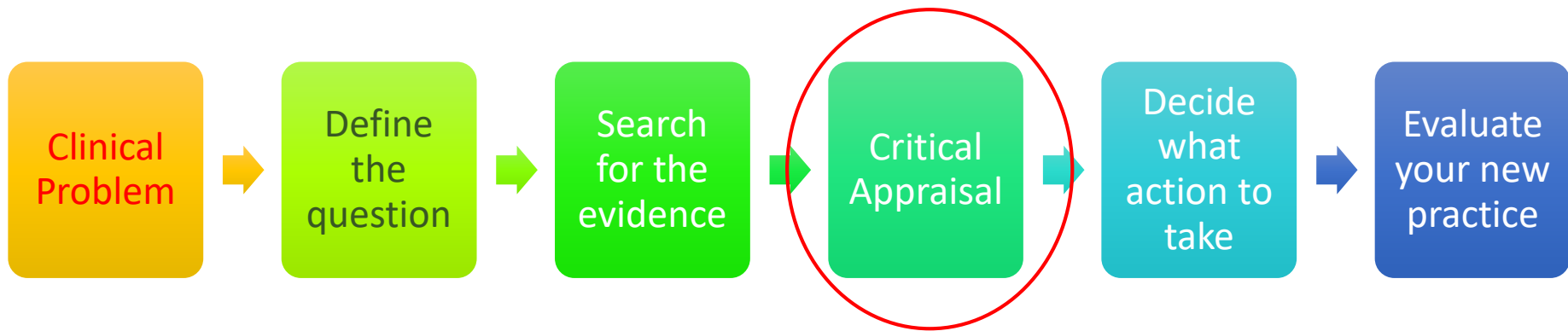
- a) Ask a random sample of women to try "Youth Code" and then comment on whether they agree that their skin is firmer and younger looking?

What is critical appraisal?

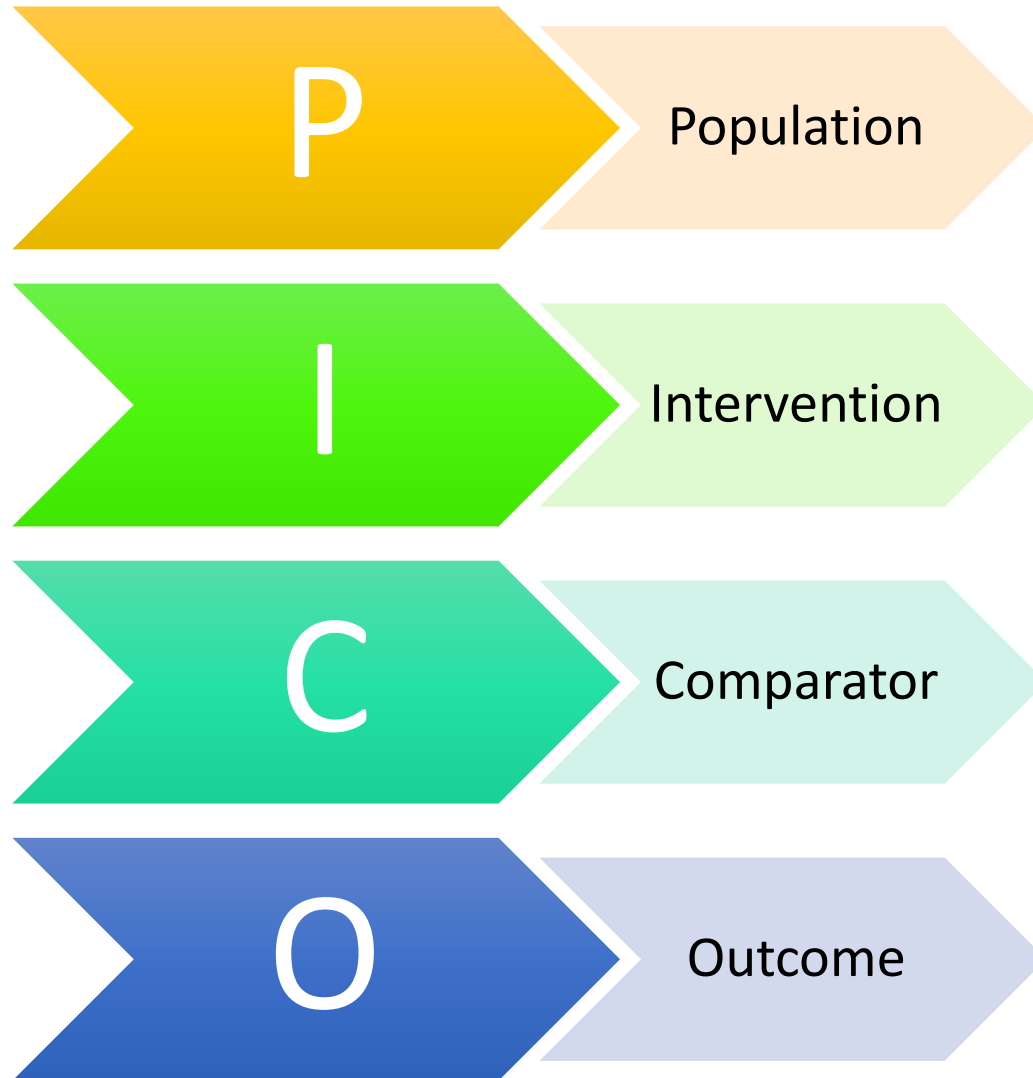
- Carefully and systematically evaluate research to assess:
 - Validity (is these findings trustworthy?)
 - Value (what do the results show?)
 - Relevance (How do these results relate to my clinical practice?)



Critical appraisal: a key component of evidence based medicine



Asking the right question



	1	2	3	4
	Patient or Problem	Intervention (a cause, prognostic factor, treatment, etc.)	Comparison Intervention (if necessary)	Outcomes
Tips for Building	Starting with your patient, ask “How would I describe a group of patients similar to mine?” Balance precision with brevity.	Ask “Which main intervention am I considering?” Be specific.	Ask “What is the main alternative to compare with the intervention?” Again, be specific.	Ask “What can I hope to accomplish?” or “What could this exposure really affect?” Again, be specific.
Example	“In patients with heart failure from dilated cardiomyopathy who are in sinus rhythm ...”	“... would adding anticoagulation with warfarin to standard heart failure therapy ...”	“... when compared with standard therapy alone ...”	“... lead to lower mortality or morbidity from thromboembolism. Is this enough to be worth the increased risk of bleeding?”

Choosing right study design

- Some study designs are not appropriate to answer certain questions
- All study designs are prone to different biases



Pyramid of evidence



So are RCTs the gold standard for evidence?



....depends

Limitations of RCTs

- Excellent vs Poor RCTs – quality varies
 - Impact on interpretation of result (external validity)?
- Expensive and time consuming
 - £250k - £millions over 2-5 years+
- May not always be the right study design to answer that question

A RCT to examine if smoking causes lung cancer

- 30 healthy Oxford Students
- Randomise to 2 groups
 - Gp1 smokes 20 cigarettes per day every day
 - Gp2 no smoking



wellcometrust

NHS
National Institute for
Health Research

MRC | Medical
Research
Council

Types of research

- What is the best study design for answering this type of question?
 - Aetiology
 - Diagnosis
 - Prognosis
 - Harm
 - Effectiveness
 - Qualitative

How to critically appraise an article

- **Validity:** methods to check that the **biases** for which that particular study design is prone have been minimised
- **Results**
- **Clinical relevance**

Validity

Internal

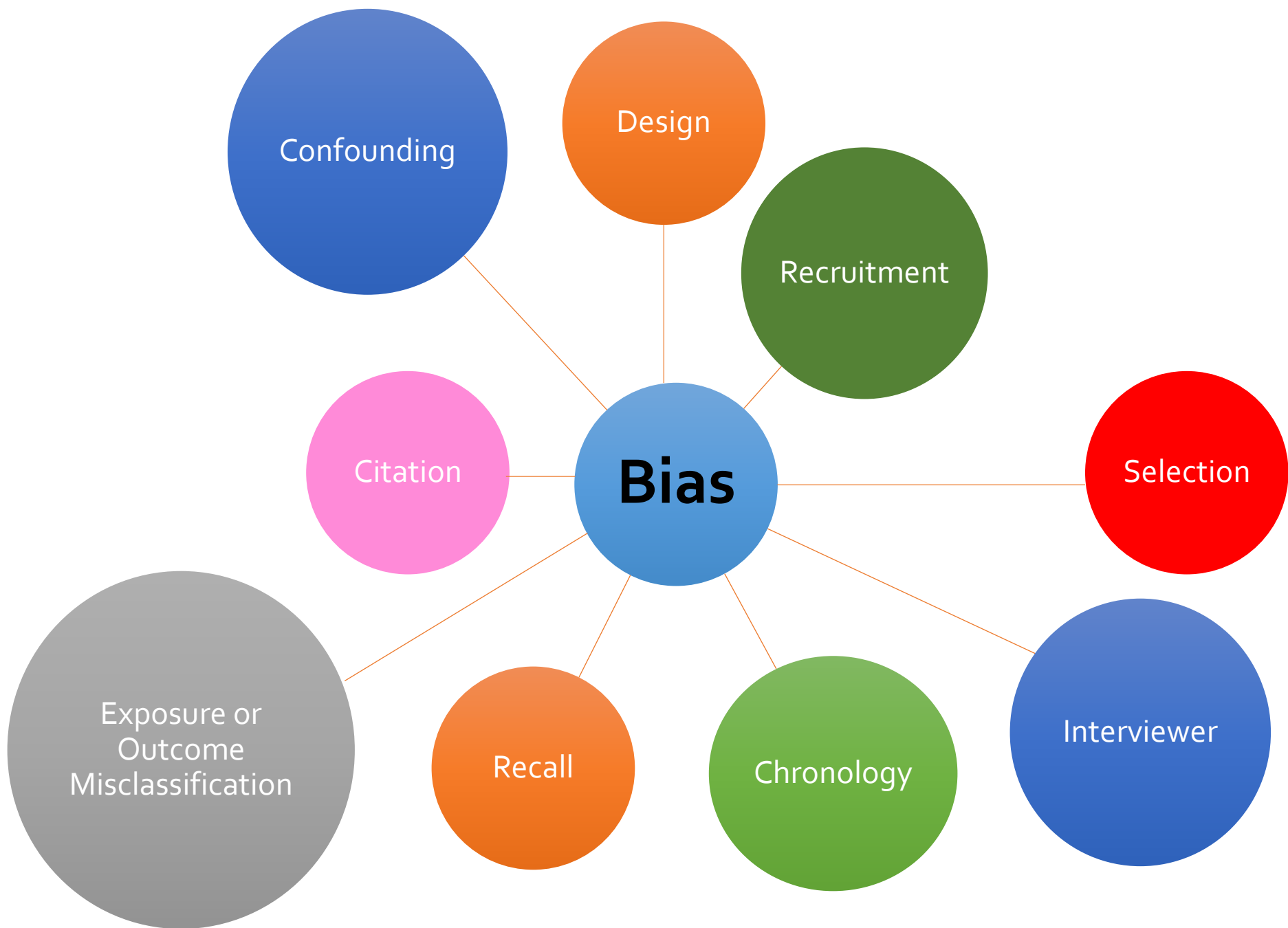
External



Bias

“the systematic deviation of the results of a study from the truth because of the way it has been conducted, analysed or reported”





Sources of bias in clinical trials

Table 1. Key sources of bias in clinical trials²

Selection bias	Biased allocation to comparison groups
Performance bias	Unequal provision of care apart from treatment under evaluation
Detection bias	Biased assessment of outcome
Attrition bias	Biased occurrence and handling of deviations from protocol and loss to follow up

Assessing Trials of effectiveness

Questions to ask:

1. Are the results of the trial valid?
2. What are the results?
3. Will the results help locally?

Checklists for clinical trials



The screenshot shows the CONSORT website homepage. At the top, there is a navigation bar with the CONSORT logo and the tagline 'TRANSPARENT REPORTING of TRIALS'. Below the navigation bar, there is a main content area with a large banner for the CONSORT 2010 checklist and a sidebar with links to key documents. The banner includes the text 'Welcome to the CONSORT Website' and a brief description of the organization's mission.

CONSORT TRANSPARENT REPORTING of TRIALS

Home CONSORT 2010 Extensions Downloads Examples Resources About CONSORT

CONSORT 2010 Key Documents

- CONSORT 2010 Checklist
- CONSORT 2010 Flow Diagram
- CONSORT 2010 Statement
- CONSORT 2010 Explanation and Elaboration Document

Welcome to the CONSORT Website

CONSORT stands for Consolidated Standards of Reporting Trials and encompasses various initiatives developed by the CONSORT Group to alleviate the problems arising from inadequate reporting of randomized controlled trials.



CEBM



Critical
Appraisal
Skills
Programme

Critical Appraisal Skills Programme (CASP)

Making sense of evidence

11 useful questions for critical appraisal of a randomised trial

1. Did the trial address a clearly focused issue?

Yes

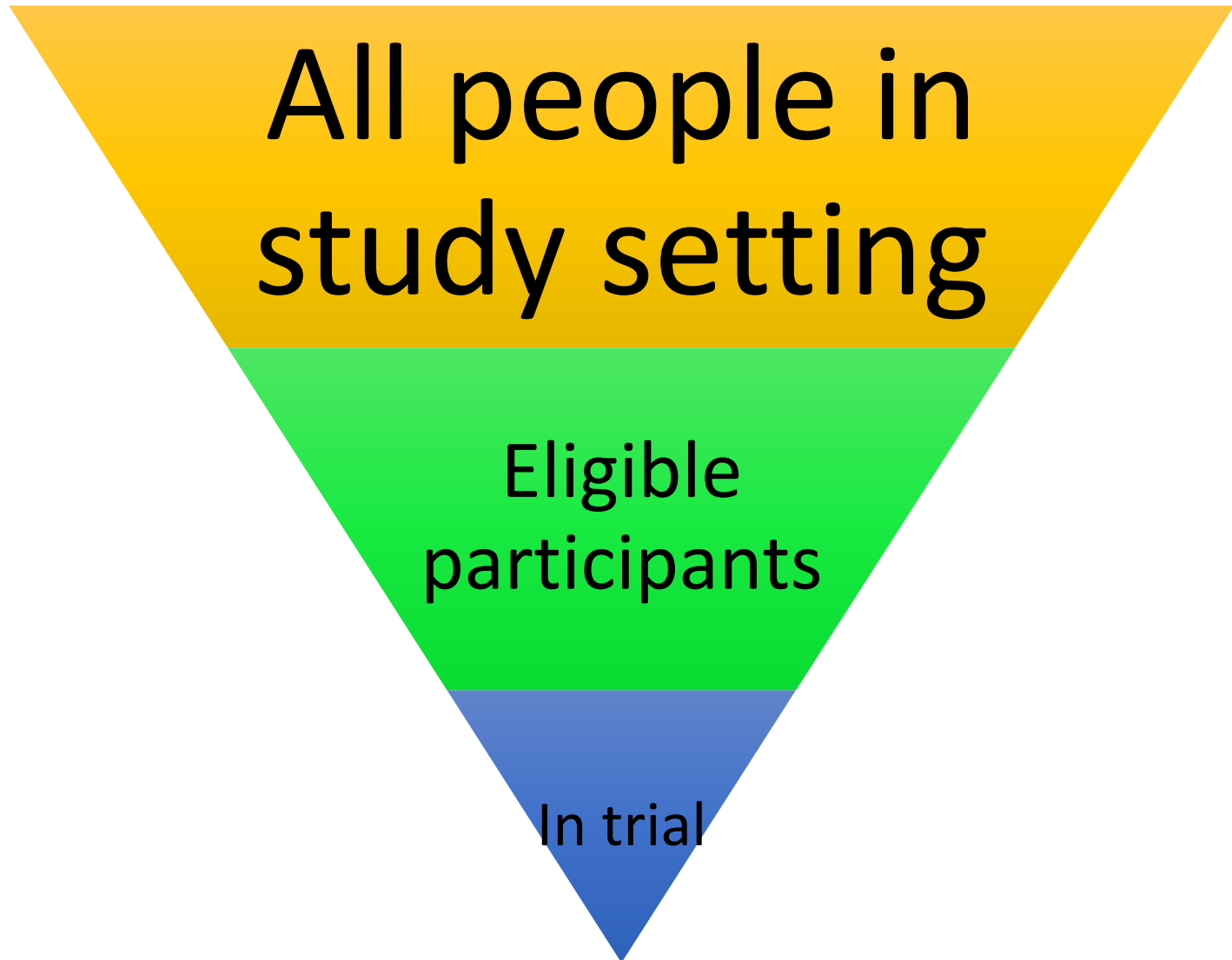
Can't tell

No

Consider: An issue can be 'focused' In terms of

- The population studied
- The intervention given
- The comparator given
- The outcomes considered

Representative: Are the trial subjects representative of patients in this setting?



2. Was the assignment of patients to treatments randomised?

Yes

Can't tell

No

Consider:

- How was this carried out, some methods may produce broken allocation concealment
- Was the allocation concealed from researchers?

Why randomise?

- Minimises measured and unmeasured confounding



Minimising allocation bias

- Centralised computer randomisation the best
- Other methods such as sealed envelopes doubtful
- Non randomised: date of birth, alternate patients alternate days, etc



If answer to first two questions
is no....



Detailed questions

3. Were patients, health workers and study personnel blinded?

Consider:

- Health workers could be; clinicians, nurses etc
- Study personnel – especially outcome assessors

Yes Can't tell No



4. Were the groups similar at the start of the trial?

Yes

Can't tell

No

Consider: Look at

- Other factors that might affect the outcome such as age, sex, social class, these may be called baseline characteristics



Maintenance: Were the groups treated equally?

5. Aside from the experimental intervention, were the groups treated equally?

Yes

Can't tell

No

6. Were all of the patients who entered the trial properly accounted for at its conclusion?

Yes

Can't tell

No

Consider:

- Was the trial stopped early?
- Were patients analysed in the groups to which they were randomised?



From: Critical Appraisal Skills Program, Oxford
www.casp-uk.net

(B) What are the results?

7. How large was the treatment effect?

Consider:

- What outcomes were measured?
- Is the primary outcome clearly specified?
- What results were found for each outcome?
- Is there evidence of selective reporting of outcomes?

8. How precise was the estimate of the treatment effect?

Consider:

- What are the confidence limits?
- Were they statistically significant?

Intention to treat

- Once a participant is randomised, they should be analysed to the group they were assigned to
- Pros
 - Reflects “real life” e.g non compliance
 - Unbiased estimate of true effect
 - Maintains sample size thus maintaining statistical power
- Cons
 - Noncompliance provides little data on efficacy
 - Treatment effect may be conservative
 - Dropouts/non-compliant/compliant subjects are different

What does this study tell us?

- **P values** (hypothesis testing):
 - Tests to exclude the null hypothesis
- **Confidence intervals** (estimation of effect)
 - Range of values within which the true effect is likely to lie
 - Wider the confidence interval, less precision in result
- **Relative Risk**
- **Absolute Risk**
- **Odds Ratios**
- **Number needed to treat**

(C) Will the results help locally?

9. Can the results be applied in your context?
(or to the local population?)

Yes Can't tell No

Consider:

- Do you have reason to believe that your population of interest is different to that in the trial
- If so, in what way?



10. Were all clinically important outcomes considered?

Yes Can't tell No

Consider:

- Is there other information you would like to have seen?
- Was the need for this trial clearly described?

11. Are the benefits worth the harms and costs?

Yes Can't tell No

Consider:

- Even if this is not addressed by the trial, what do you think?

Example to work through

- OPPTIMUM study: vaginal progesterone prophylaxis for preterm birth: a multicentre, randomised, double blind trial
- Norman et al, The Lancet, vol 387 May 2016

Conclusion

- Critical appraisal helps us decide whether evidence is valid, what the results tell us and whether the study is relevant to our setting
- Checklists are available to help
- Don't believe everything you read in journals!