Rubin Chapter 5: Critically appraising experiments

1. Classic Pretest-Posttest Control Group Design
2. Posttest-Only Control Group Design
3. Solomon Four-Group Design
4. Alternative Treatment Designs (can add TAU)
5. Dismantling Designs

O AB O

O A O

O B O

O TAU O

1. Placebo Control Group Designs

* **Placebo Effects**
* **Novelty and Disruption Effects**

1. Experimental Demand and Experimenter Expectancies
2. Obtrusive versus Unobtrusive Observation
3. Compensatory Equalization and Compensatory Rivalry
4. Resentful Demoralization
5. Treatment Diffusion
6. Treatment Fidelity
7. Practitioner Equivalence
8. Differential Attrition

Trisha Greenhalgh: *How to read a paper:* Assessing the methodological quality of published papers

1. Was the study original?

* Is this study bigger, continued for longer, or otherwise more substantial than the previous one(s)?
* Is the methodology of this study any more rigorous (in particular, does it address any specific methodological criticisms of previous studies)?
* Will the numerical results of this study add significantly to a meta-analysis of previous studies?
* Is the population that was studied different in any way (has the study looked at different ages, sex, or ethnic groups than previous studies)?
* Is the clinical issue addressed of sufficient importance, and is there sufficient doubt in the minds of the public or key decision makers, to make new evidence “politically” desirable even when it is not strictly scientifically necessary?

1. Whom is the study about?

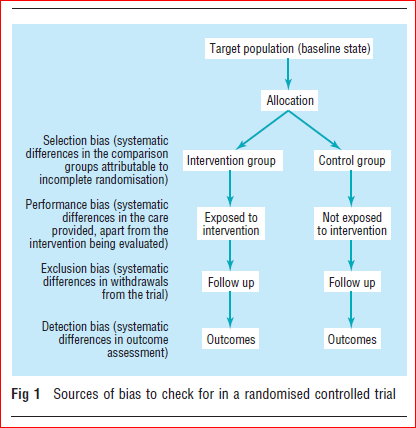
* *How were the subjects recruited?*
* *Who was included in the study?*
* *Who was excluded from the study?*
* *Were the subjects studied in “real life” circumstances?*

1. Was the design of the study sensible?

* *What specific intervention or other manoeuvre was being considered, and what was it being compared with?*
* *What outcome was measured, and how?*

1. Was systematic bias avoided or minimised?

* **Randomised controlled trials:** In a randomised controlled trial, systematic bias is (in theory) avoided by selecting a sample of participants from a particular population and allocating them randomly to the different groups. Figure 1 summarises sources of bias to check for.



* **Non-randomised controlled clinical trials:** As a general rule, if the paper you are looking at is a non-randomised controlled clinical trial, you must use your common sense to decide if the baseline differences between the intervention and control groups are likely to have been so great as to invalidate any differences ascribed to the effects of the intervention. This is, in fact, almost always the case.
* **Cohort studies:** The selection of a comparable control group is one of the most difficult decisions facing the authors of an observational (cohort or case-control) study. Few, if any,cohort studies, for example, succeed in identifying two groups of subjects who are equal in age, sex mix, socioeconomic status, presence of coexisting illness, and so on, with the single difference being their exposure to the agent being studied. In practice, much of the “controlling” in cohort studies occurs at the analysis stage, where complex statistical adjustment is made for baseline differences in key variables. Unless this is done adequately, statistical tests of probability and confidence intervals will be dangerously misleading. (NB: These are “quasiexperimental designs.”)
* **Case-control studies :** In case-control studies (in which the experiences of individuals with and without a particular disease are analysed retrospectively to identify putative causative events), the process that is most open to bias is not the assessment of outcome, but the diagnosis of “caseness” and the decision as to when the individual became a case.

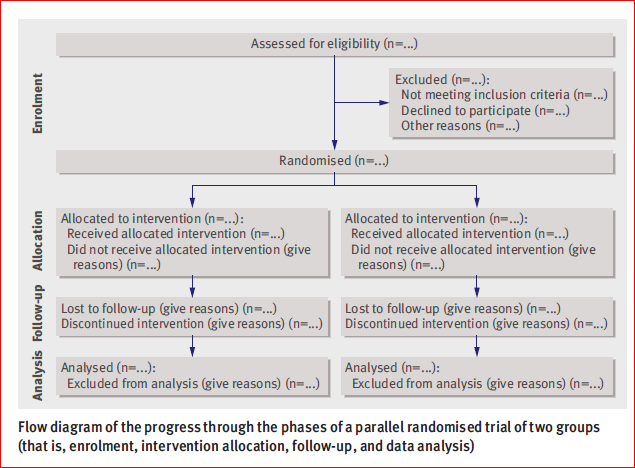
1. Was assessment “blind”?

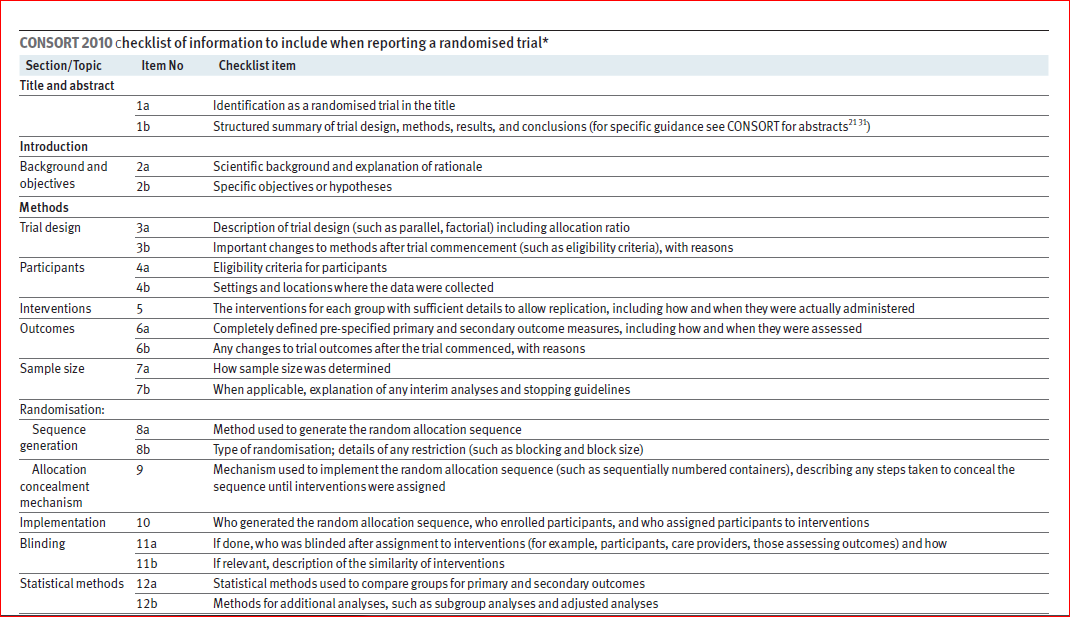
* Even the most rigorous attempt to achieve a comparable control group will be wasted effort if the people who assess outcome (for example, those who judge whether someone is still clinically in heart failure, or who say whether an *x* ray is “improved” from last time) know which group the patient they are assessing was allocated to.

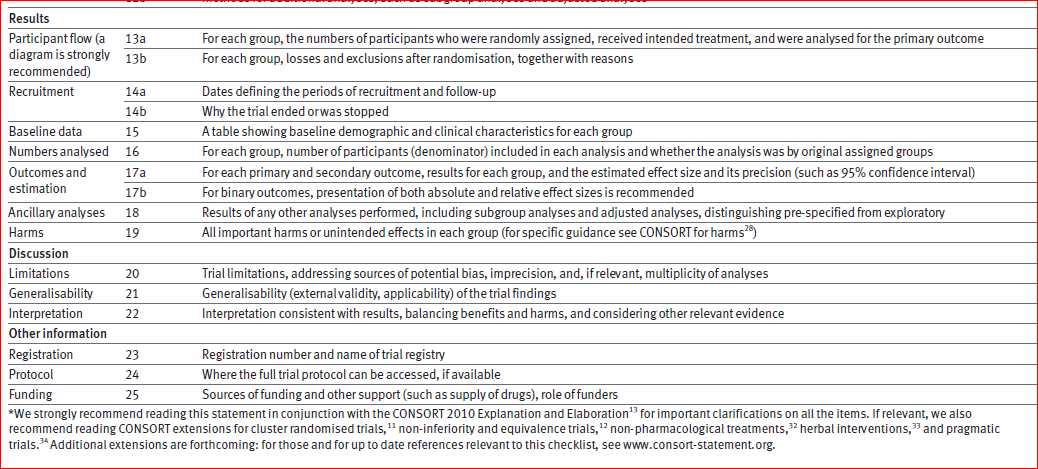
1. Were preliminary statistical questions dealt with?

* **Sample size :** A trial should be big enough to have a high chance of detecting, as statistically significant, a worthwhile effect if it exists, and thus to be reasonably sure that no benefit exists if it is not found in the trial. To calculate sample size, the clinician must decide two things. The first is what level of difference between the two groups would constitute a clinically significant effect. Note that this may not be the same as a statistically significant effect.
* **Duration of follow up:** Even if the sample size was adequate, a study must continue long enough for the effect of the intervention to be reflected in the outcome variable.
* **Completeness of follow up:** Subjects who withdraw from (“drop out of”) research studies are less likely to have taken their tablets as directed, more likely to have missed their interim checkups, and more likely to have experienced side effects when taking medication, than those who do not withdraw.

Kenneth F Schulz, Douglas G Altman, David Moher, for the CONSORT Group: CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials







David Moher, Sally Hopewell, Kenneth F Schulz, Victor Montori, Peter C Gøtzsche, P J Devereaux, Diana Elbourne, Matthias Egger, Douglas G Altman: CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials

