

Challenges in applying guidance on subgroup analysis

Mark Petticrew

London School of Hygiene and Tropical Medicine



The problem

- Systematic reviews of social interventions are expected to present evidence not just on overall population effects, but also on differential effects
- “*What works for whom and in what circumstances*”
- i.e. the “differential” effects of interventions across subgroups
- Also important for reasons of equity: (“are the effects of this intervention differentiated by socioeconomic group... ethnicity...gender...age...”)

- This poses the problem that, even where such evidence of differential effects exists, it usually derives from post-hoc subgroup analyses
- Usually described as weak and potentially misleading evidence:
- Altman (1993) *Sub-group analyses pose problems in interpretation...it is reasonable to carry out a small number of subgroup analyses if these were specified in the protocol [author's emphasis] but on no account should the data be analysed in numerous different ways in the hope of discovering some significant comparison".*

The Cochrane handbook:

- *“Subgroup analyses are observational by nature* and are not based on randomized comparisons. False negative and false positive significance tests increase in likelihood rapidly as more subgroup analyses are performed. If their findings are presented as definitive conclusions there is clearly a risk of patients being denied an effective intervention or treated with an ineffective (or even harmful) intervention. Subgroup analyses can also generate misleading recommendations about directions for future research that, if followed, would waste scarce resources”.*
- *So, just like systematic reviews, then...

It gets worse...

- “*Subgroups kill people*” (Rothwell (2005), quoting statistician Richard Peto)

The effects may not be real anyway

- In meta-analyses the results observed in subgroups may differ *by chance* from the overall effect, and the subgroup findings may not be confirmed by subsequent large trials.
- In this case, the best estimate of the outcome of the intervention in that sub-group will come paradoxically, by discounting the results from that sub-group, and using the results of the overall meta-analysis (i.e., Stein’s paradox).
- But for some social interventions it is entirely plausible that the effects vary between groups

Guidance on subgroup analysis emphasises that...

- Subgroup analyses should be restricted to those proposed before data collection, driven by theory (or a specific logic model);
- Any chosen after this should be clearly identified
- Trials should ideally be powered with subgroups in mind
- Subgroup analyses should always be based on formal tests of interaction, although even these should be interpreted with caution
- The results of subgroup analyses should not be over-interpreted. Unless there is strong supporting evidence, they are best viewed as a hypothesis-generating exercise.

Questions

- Does this sort of guidance also apply equally to “social” systematic reviews?
- Given SGA is key to many reviews, are there ways we can enhance its reliability?
- How do we tell present them to decision-makers?