Sub-Group Analysis in Randomized Controlled Trials

2013 Anesthesia Interest Group Meeting

J.W. Busse, DC, PhD
Assistant Professor
Dept. of Anesthesia
McMaster University
The influence of study characteristics on reporting of subgroup analyses in randomised controlled trials: systematic review

Xin Sun, research fellow, 1,2 Matthias Briel, assistant professor, 1,3 Jason W Busse, scientist, 1,4 John J You, assistant professor, 1,5 Elie A Akl, associate professor, 1,6 Filip Mejza, research fellow, 1,7 Malgorzata M Bala, research fellow, 1,8 Dirk Bassler, associate professor, 1,9 Dominik Mertz, research fellow, 1,10 Natalia Diaz-Granados, doctoral candidate, 1,11 Per Olav Vandvik, researcher, 11,12 German Malaga, associate professor, 13 Sadeesh K Srinathan, assistant professor, 14 Philipp Dahm, associate professor, 15 Bradley C Johnston, postdoctoral fellow, 1 Pablo Alonso-Coello, researcher, 16 Basil Hassouneh, research fellow, 1 Jessica Truong, undergraduate student, 17 Neil D Dattani, medical student, 18 Stephen D Walter, professor, 1 Diane Heels-Ansdell, statistician, 1 Neera Bhatnagar, librarian, 19 Douglas G Altman, professor, 20 Gordon H Guyatt, professor
Background

- Investigators carry out subgroup analyses to examine if observed treatment effects differ across baseline characteristics.

- Subgroup analyses are common in randomised controlled trials:
  - 60% of trials published in high impact general medical journals;
  - 61% of cardiovascular trials; and
  - 37% of surgical trials report subgroup analyses.
Study Focus

- We systematically reviewed RCTs to investigate the association of prespecified study characteristics with reporting of subgroup analyses.

- In particular we examined the impact of industry funding on the reporting of subgroups.
We found that studies with declared industry funding received approximately 22 more citations per year only if their results were industry-favoring.

The added influence appears to be the quantitative equivalent of having an extra publication in a high-impact journal.
Flow of Study Screening

Random sample of searched citations: 1140 journal reports (570 in high impact, 570 in lower impact journals)*

Excluded studies (n=505):
- Not a study in humans (n=3)
- Not a randomised controlled trial (n=502)

Potentially eligible reports after screening of title and abstract (n=635)

Excluded studies† (n=166):
- Not a randomised controlled trial (n=120)
- Explicitly labelled as phase I trial (n=4)
- Exclusively investigated pharmacokinetic mechanisms (n=12)
- Research letter (n=1)
- Not a study in humans (n=2)
- Focus on subset of enrolled participants (n=30)

Final inclusion of trials after full text screening (n=469):
- High impact journals (n=219)
- Lower impact journals (n=250)

Trials reporting subgroup analysis (n=207):
- High impact journals (n=41, 64%)
- Lower impact journals (n=66, 26%)

* Annals of Internal Medicine, BMJ, JAMA, Lancet, and New England Journal of Medicine
† Studies may be excluded for multiple reasons
# Regression Analysis of Factors Associated with Reporting Subgroup Analyses

<table>
<thead>
<tr>
<th>Study characteristics</th>
<th>Univariable analyses</th>
<th>Multivariable analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Odds ratio (95% CI)</td>
<td>P value</td>
</tr>
<tr>
<td>High impact v lower impact journals*</td>
<td>5.04 (3.39 to 7.48)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Non-surgical v surgical trial</td>
<td>3.01 (1.91 to 4.74)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Sample size per arm (fourths):</td>
<td>1 (reference)</td>
<td></td>
</tr>
<tr>
<td>3-32</td>
<td>2.38 (1.30 to 4.36)</td>
<td>0.005</td>
</tr>
<tr>
<td>33-101</td>
<td>5.85 (3.21 to 10.65)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>102-301</td>
<td>8.64 (4.70 to 15.86)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>≥302</td>
<td>1.18 (0.97 to 1.43)</td>
<td>0.098</td>
</tr>
<tr>
<td>No of prespecified primary outcomes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistical significance of result for primary outcome: non-significant v significant</td>
<td>1.38 (0.94 to 2.01)†</td>
<td>0.092</td>
</tr>
<tr>
<td>Industry funding v other</td>
<td>1.91 (1.31 to 2.77)†</td>
<td>0.001</td>
</tr>
<tr>
<td>Statistical significance×trial funding</td>
<td>1.89 (0.85 to 4.25)§</td>
<td>0.12</td>
</tr>
<tr>
<td>Association of trial funding (industry v other) with reporting subgroup analyses**:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>With non-significant primary outcome</td>
<td>3.00 (1.56 to 5.76)</td>
<td>0.001</td>
</tr>
<tr>
<td>With significant primary outcome</td>
<td>1.58 (0.99 to 2.53)</td>
<td>0.057</td>
</tr>
</tbody>
</table>
Other Results

- Industry-funded RCTs are less likely to:
  - Pre-specify subgroup hypotheses
  - Use a test of interaction for subgroup effects
Credibility of claims of subgroup effects in randomised controlled trials: systematic review

Xin Sun assistant professor ¹ ² ³, Matthias Briel assistant professor ² ³, Jason W Busse scientist ² ⁴, John J You assistant professor ² ⁵, Elie A Akl associate professor ² ⁶, Filip Mejza research fellow ⁷, Malgorzata M Bala research fellow ⁸, Dirk Bassler associate professor ⁹, Dominik Mertz assistant professor ² ⁵ ¹⁰, Natalia Diaz-Granados doctoral candidate ², Per Olav Vandvik researcher ¹¹ ¹², German Malaga associate professor ¹³, Sadeesh K Srinathan assistant professor ¹⁴, Philipp Dahm professor ¹⁵, Bradley C Johnston assistant professor ² ¹⁶, Pablo Alonso-Coello researcher ¹⁷, Basil Hassouneh research fellow ², Stephen D Walter professor ², Diane Heels-Ansdell statistician ², Neera Bhatnagar librarian ¹⁸, Douglas G Altman professor ¹⁹, Gordon H Guyatt professor ²

BMJ 2012;344:e1553 doi: 10.1136/bmj.e1553 (Published 15 March 2012)
Background

- The credibility of a proposed subgroup effect ranges from extremely unlikely to highly plausible.

- There are criteria to evaluate the credibility of reported subgroup effects.
Study Focus

- We systematically reviewed RCTs to investigate the degree to which reported subgroup effects adhered to credibility criteria.
Criteria to Establish the Credibility of Subgroup Effects

**Design**
- Was the subgroup variable a baseline characteristic?
- Was the subgroup variable a stratification factor at randomisation?*
- Was the subgroup hypothesis specified a priori?
- Was the subgroup analysis one of a small number of subgroup hypotheses tested (≤5)?

**Analysis**
- Was the test of interaction significant (interaction P<0.05)?
- Was the significant interaction effect independent, if there were multiple significant interactions?

**Context**
- Was the direction of subgroup effect correctly prespecified?
- Was the subgroup effect consistent with evidence from previous related studies?
- Was the subgroup effect consistent across related outcomes?
- Was there any indirect evidence to support the apparent subgroup effect—for example, biological rationale, laboratory tests, animal studies?
64/207 RCTs Reporting Subgroup Analyses
Claimed Effects for the Primary Outcome

Measured subgroup variables at baseline – 94%
Used subgroup variable as a stratification factor – 20%
Pre-specified their subgroup hypothesis – 44%
Correctly pre-specified direction of effect – 6%
Tested a small number of hypotheses – 44%
Carried out a test of interaction that was significant – 9%
Documented replication of subgroup effect – 33%
Showed consistency of subgroup effect across outcomes – 30%
Provided indirect evidence for the effect – 22%

In 19 trials making more than 1 claim, only 1 (5%) checked the independence of the interaction
Results

- Of the 64 claims for subgroup effects, 54 (84%) met 4 or fewer of the 10 credibility criteria.
Hot off the Presses...

- An important credibility criteria for reported subgroup effects is pre-specification

- We are limited to interpreting what author’s have reported ...or are we?
Study Focus

- We investigated subgroup planning and reporting in a consecutive series of RCT protocols that were approved by 6 research ethics committees in 3 countries.

- We specifically examined the concordance between statements about pre-specification in the publication and corresponding statements in the protocols.
894 trial protocols included (protocol set)

362 trials not published in journal
- No subgroup analyses planned N=295
- Subgroup analyses planned N=66

532 trials published as full journal publication
- 515 journal publications considered for reporting analyses (publication set)
  - Subgroup analyses reported N=246
  - No subgroup analyses reported N=269

6 reported pooled data
5 no RCT results
6 letters to editor
Results

- In 81 of 246 (32.9%) publications, authors stated that at least one of their reported subgroup analyses was pre-specified, but only “X”% of underlying protocols reported a planned subgroup analysis.

60%
Conclusions

- Subgroup effects are much more likely to be reported by industry-funded trials when the primary outcome is non-significant.

- Claims of subgroup effects rarely meet even half of criteria for credibility.

- 40% of trials that claim pre-specified subgroup effects are not being truthful.
Conclusions

- If you choose to disbelieve a reported subgroup effect, you will be right far more often than you are wrong.
Thanks for Your Time