Unplanned versus pre-specified subgroup analysis reporting

Doug Altman

Centre for Statistics in Medicine
University of Oxford

ESMO, Copenhagen, 10 October 2016

Conflicts of interest – None

Subgroup analyses

“Properly performed, analysis of subgroups can yield useful insights into therapy; unfortunately, many commonly used approaches are often uninformative or misleading.”

[Yusuf et al, *JAMA* 1991]

Some questions

- What are (in)appropriate statistical approaches?
- How often is an appropriate method used?
- How good is current reporting?
- How does oncology compare with other specialties?
- What is good practice?

Main subgroup analysis strategies in RCTs

- **Treatment effect in subset of the participants (e.g. diabetics)**
  - Disregard the others (e.g. non-diabetics)

- **Treatment effects separately for 2 or more complementary subsets of participants (e.g. by diabetes; by cancer stage)**
  - Separate analyses (2+ P values)

- **Compare treatment effects across complementary subgroups – test of interaction**
  - One analysis (one P value)

Main subgroup analysis strategies in RCTs

- ~~**Treatment effect in subset of the participants (e.g. diabetics)**~~
  - ~~Disregard the others~~ (e.g. non-diabetics)

- ~~**Treatment effects separately for 2 or more complementary subsets of participants (e.g. by diabetes; by cancer stage)**~~
  - Separate analyses (2+ P values)

- **Compare treatment effects across complementary subgroups – test of interaction**
  - One analysis (one P value)

Published subgroup analyses

- **Subgroup analyses should be pre-planned**
  - Frequent discrepancies between trial protocols and subsequent publications
- **Even prespecified subgroup analyses lack power**
  - So most significant results will be false positives
- **Results of all subgroup analyses should be reported**
  - Publishing only significant results magnifies problems
  - Indicate whether pre-specified

BMJ 2014;349:g4539 doi: 10.1136/bmj.g4539 (Published 16 July 2014)

RESEARCH

Subgroup analyses in randomised controlled trials: cohort study on trial protocols and journal publications

[Kasenda *et al*, *BMJ* 2014]

Comparison of protocol and publication

Clinical discipline	No. of trials	No. (%) pre-planned
Oncology	155	42 (27%)
Cardiovascular	108	49 (45%)
Infectious disease	87	27 (31%)
Endocrinology	62	15 (24%)
Neurology	61	24 (39%)
Other	421	95 (23%)

[Kasenda *et al*, *BMJ* 2014]

An example of subgroup discrepancies

Outcome: **time to progression or death**

Subgroup analyses:

***Protocol*:** **baseline disease severity**

***Publication*: duration of previous treatment*,**
type of previous treatment*,
blood count*,
disease severity

***Described explicitly as pre-specified despite not appearing in the protocol**

Gabler *et al. Trials* (2016) 17:320
DOI 10.1186/s13063-016-1447-5

Trials

RESEARCH Open Access

No improvement in the reporting of clinical trial subgroup effects in high-impact general medical journals

Nicole B. Gabler[1*], Naihua Duan[2], Eli Raneses[1], Leah Suttner[1,3], Michael Ciarametaro[4], Elizabeth Cooney[1], Robert W. Dubois[4], Scott D. Halpern[1,5] and Richard L. Kravitz[6]

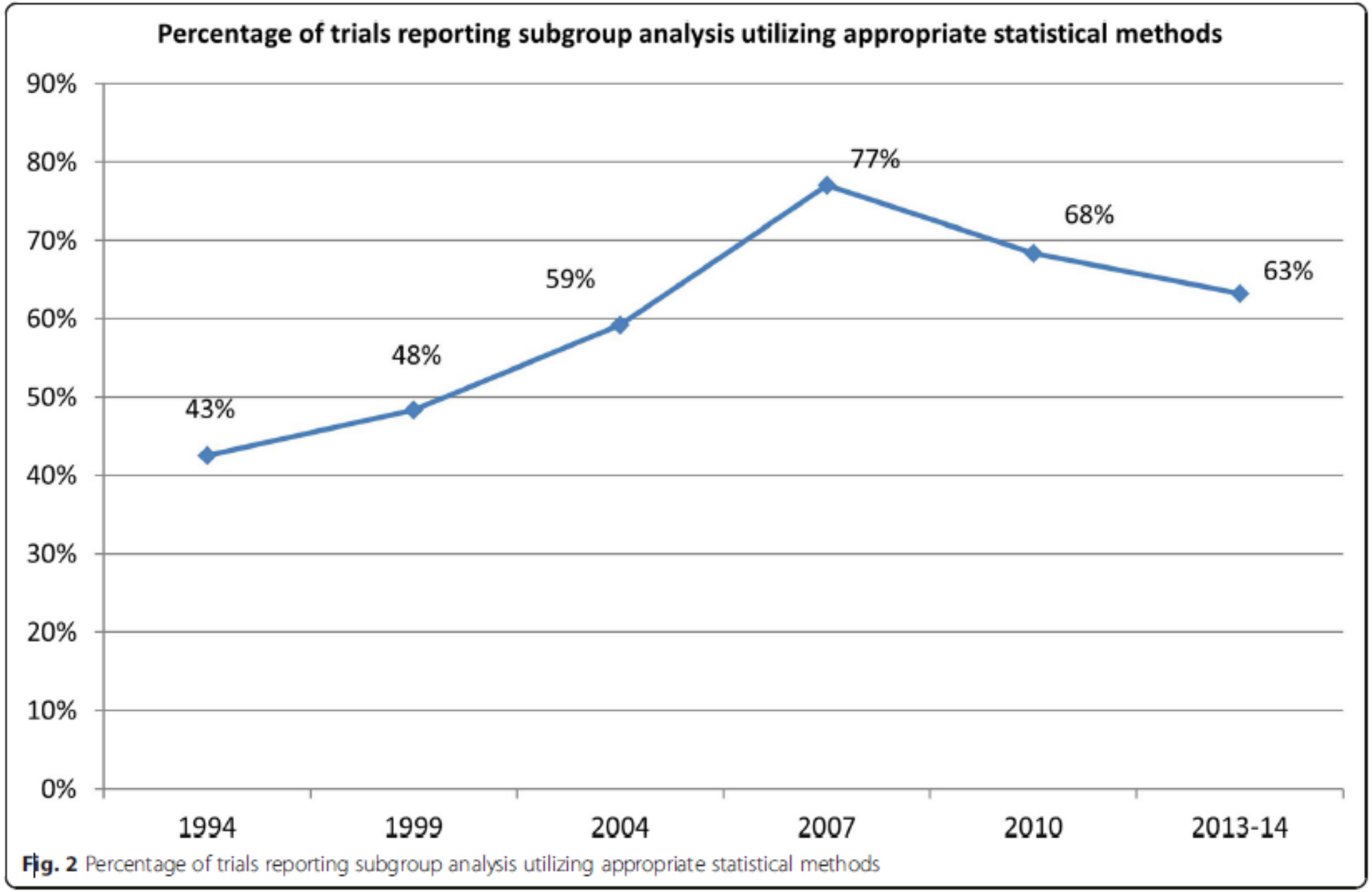


Fig. 2 Percentage of trials reporting subgroup analysis utilizing appropriate statistical methods

	No. of trials	Reported subgroup analyses	Used appropriate method
All trials	437	270 (62%)	185 (69%)

	No. of trials	Reported subgroup analyses	Used appropriate method
All trials	437	270 (62%)	185 (69%)
Cancer	59	45 (76%)	24 (53%)

Cancer was 2nd worst of 9 medical areas studied

	No. of trials	Reported subgroup analyses	Used appropriate method
All trials	437	270 (62%)	185 (69%)
Cancer	59	45 (76%)	24 (53%)
Cardiovascular	101	73 (72%)	63 (86%)

Cancer was 2nd worst of 9 medical areas studied

Bad approaches

- **Separate P values for each subset (or ignoring one subset)**
- **Post hoc comparison of subgroups based on observed results**
  - i.e. not pre-planned
- **Comparison of subset of active group with whole comparison group**
- **Subgroup defined by variable not known at randomisation and possibly influenced by treatment (“improper subgroups”)**
  - e.g. compliance, early clinical response

- **All the above methods can be extremely misleading**
- **Pre-specifying doesn’t guarantee an appropriate analysis**

Heterogeneity of treatment effect - a better approach [Kent et al, *Trials* 2010]

- **Conventional subgroup analyses do not account for the fact that patients have multiple characteristics simultaneously that affect the likelihood of treatment benefit**

- **A good approach is to estimate treatment effect in relation to multivariable risk score (risk stratification)**

Reporting subgroup analyses

Methods

- **Prespecified subgroup analyses (as in protocol)**
- **Postulated direction of effect**
- **Statistical method**

Results

- **Estimated effect size (with CI) in each subgroup**
- **Test of interaction (estimate of relative effect; P value)**

Carboplatin and weekly paclitaxel doublet chemotherapy compared with monotherapy in elderly patients with advanced non-small-cell lung cancer: IFCT-0501 randomised, phase 3 trial

Elisabeth Quoix, Gérard Zalcman, Jean-Philippe Oster, Virginie Westeel, Eric Pichon, Armelle Lavolé, Jérôme Dauba, Didier Debieuvre, Pierre-Jean Souquet, Laurence Bigay-Game, Eric Dansin, Michel Poudenx, Olivier Molinier, Fabien Vaylet, Denis Moro-Sibilot, Dominique Herman, Jaafar Bennouna, Jean Tredaniel, Alain Ducoloné, Marie-Paule Lebitasy, Laurence Baudrin, Silvy Laporte, Bernard Milleron, on behalf of Intergroupe Francophone de Cancérologie Thoracique

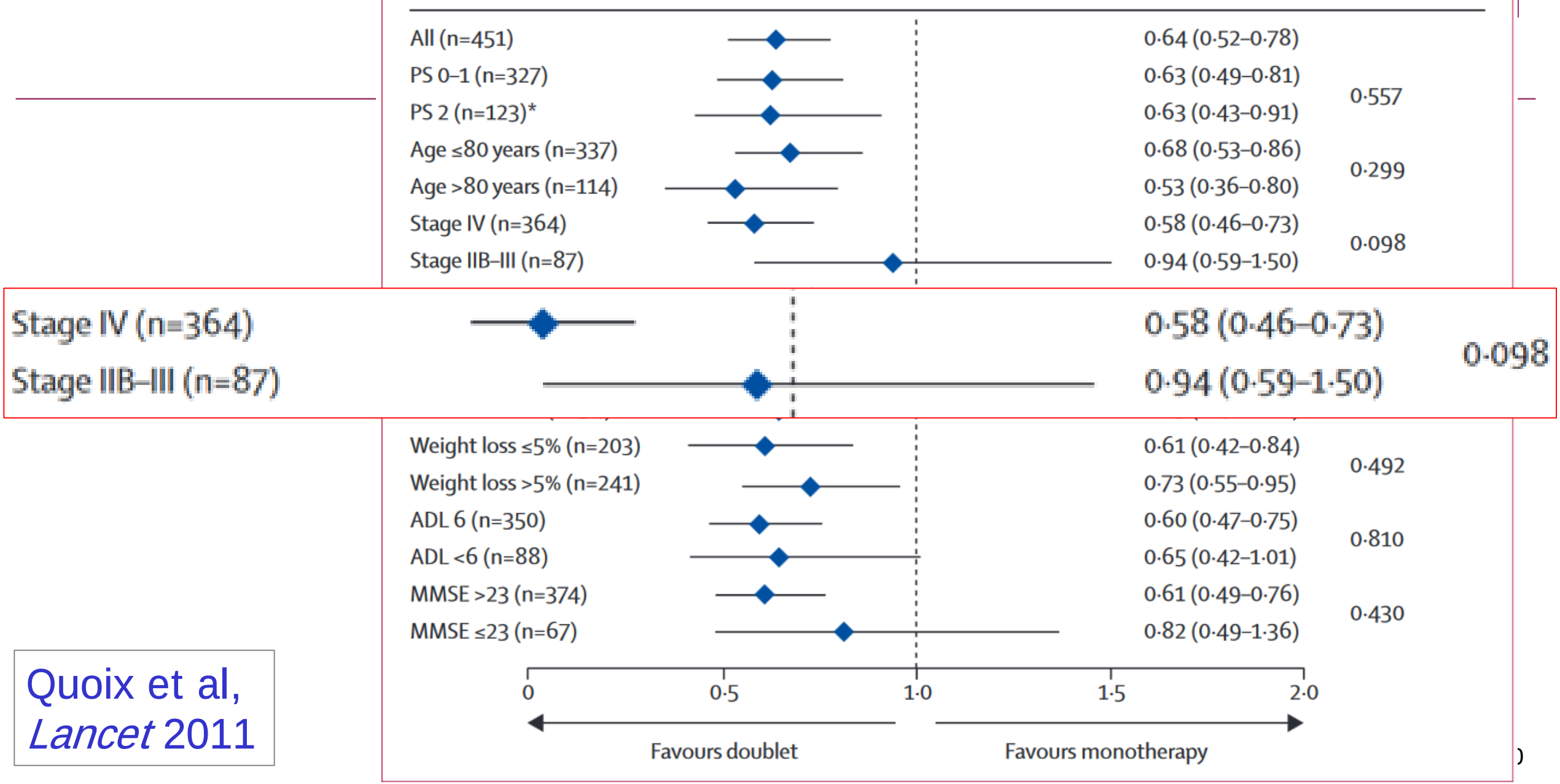
Summary

Background Platinum-based doublet chemotherapy is recommended to treat advanced non-small-cell lung cancer (NSCLC) in fit, non-elderly adults, but monotherapy is recommended for patients older than 70 years. We compared a carboplatin and paclitaxel doublet chemotherapy regimen with monotherapy in elderly patients with advanced NSCLC.

Lancet 2011; 378: 1079–88
Published Online
August 9, 2011
DOI:10.1016/S0140

Quoix et al,
Lancet 2011

***Figure 3*: Forest plot for subgroup analysis of overall survival**
Survival is for doublet chemotherapy *vs* monotherapy. Data are derived from Cox's analysis without covariates.



Quoix et al, *Lancet* 2011

Good practice

- **Pre-specify a (very) few planned subgroup analyses in trial protocol**
  - Preferably with rationale and direction of postulated difference
- **Use only variables known at baseline**
- **Use interaction analysis or multivariable risk score**
- **Indicate all subgroups analyses undertaken**
  - and whether prespecified or post hoc
- **Interpret subgroup findings very cautiously**
  - Exploratory analyses are good for hypothesis generating
  - Even pre-planned analyses may be misleading

Some key references

Altman DG. Subgroup analyses in randomized trials: more rigour needed. *Nature Rev Clin Oncol* 2015; 12: 506-7.

Gabler NB et al. No improvement in the reporting of clinical trial subgroup effects in high-impact general medical journals. *Trials* 2016; 17: 320.

Sun X et al. Credibility of claims of subgroup effects in randomised controlled trials: Systematic review. *BMJ* 2012; 344: e1553.

Sun X et al. How to use a subgroup analysis: Users' guide to the medical literature. *JAMA* 2014; 311: 405-411.

Yusuf S et al. Analysis and interpretation of treatment effects in subgroups of patients in randomized clinical trials. *JAMA* 1991; 266: 93-8.

Zhang S *et al.* Subgroup analyses in reporting of phase III clinical trials in solid tumours. *J Clin Oncol* 2015; 33: 1697-1702.