

RESEARCH ARTICLE

Open Access



# Cumulative subgroup analysis to reduce waste in clinical research for individualised medicine

Fujian Song\* and Max O. Bachmann

## Abstract

**Background:** Although subgroup analyses in clinical trials may provide evidence for individualised medicine, their conduct and interpretation remain controversial.

**Methods:** Subgroup effect can be defined as the difference in treatment effect across patient subgroups. Cumulative subgroup analysis refers to a series of repeated pooling of subgroup effects after adding data from each of related trials chronologically, to investigate the accumulating evidence for subgroup effects. We illustrated the clinical relevance of cumulative subgroup analysis in two case studies using data from published individual patient data (IPD) meta-analyses. Computer simulations were also conducted to examine the statistical properties of cumulative subgroup analysis.

**Results:** In case study 1, an IPD meta-analysis of 10 randomised trials (RCTs) on beta blockers for heart failure reported significant interaction of treatment effects with baseline rhythm. Cumulative subgroup analysis could have detected the subgroup effect 15 years earlier, with five fewer trials and 71% less patients, than the IPD meta-analysis which first reported it. Case study 2 involved an IPD meta-analysis of 11 RCTs on treatments for pulmonary arterial hypertension that reported significant subgroup effect by aetiology. Cumulative subgroup analysis could have detected the subgroup effect 6 years earlier, with three fewer trials and 40% less patients than the IPD meta-analysis. Computer simulations have indicated that cumulative subgroup analysis increases the statistical power and is not associated with inflated false positives.

**Conclusions:** To reduce waste of research data, subgroup analyses in clinical trials should be more widely conducted and adequately reported so that cumulative subgroup analyses could be timely performed to inform clinical practice and further research.

**Keywords:** Subgroup analysis, Individual patient data, Cumulative meta-analysis, Randomised controlled trials, Individualised medicine

## Background

Randomised controlled trials (RCTs) provide the most valid evidence on effects of treatments and healthcare interventions, and results of RCTs are usually reported as estimated average effects. However, average effects from RCTs may not be generalisable to individual patients in clinical practice, because of heterogeneity across individual patients in terms of demographic characteristics, genetic features, disease severity, co-morbidities and other

factors [1, 2]. Subgroup analysis is often used in RCTs to investigate differences in treatment effect between patients with different characteristics [3]. A study found that subgroup analyses were reported in 44% of the 469 RCTs published in core clinical journals in 2007 [4].

There are well known limitations with subgroup analyses in RCTs, including possible false positive subgroup effects due to multiple testing, and false negative subgroup results due to inadequate statistical power [5]. It has been recommended that only a small number of pre-specified subgroup analyses in RCTs should be conducted using appropriate statistical methods [6, 7].

\* Correspondence: [fujian.song@uea.ac.uk](mailto:fujian.song@uea.ac.uk)  
Norwich Medical School, Faculty of Medicine and Health Science, University of East Anglia, Research Park, Norwich, Norfolk NR4 7TJ, UK

Methods have been recommended for assessing credibility of subgroup analyses in RCTs [8, 9]. However, the conduct, reporting and interpretation of subgroup analysis in RCTs are still controversial [10, 11].

Although meta-analyses of RCTs are generally accepted for estimating the overall treatment effects, subgroup analyses in meta-analysis are hampered because of inadequate data on subgroups in published trials [12]. Meta-analysis using individual patient data (IPD) has been increasingly used to conduct more statistically powerful subgroup analyses by pooling data from multiple trials [13]. A recent study found that significant subgroup effects were identified in 44 IPD meta-analyses and in only three aggregate data meta-analyses among 204 paired IPD meta-analyses and aggregate data meta-analyses [14].

Previous studies regarding subgroup analysis methodology have focused mainly on the results of individual trials, and research data may have been wasted for detecting clinically important differences in treatment effects between patient subgroups. The objective of the current study is to illustrate the usefulness and statistical properties of cumulative subgroup analysis by providing empirical evidence from case studies and computing simulations.

## Methods

Patient characteristics at baseline are relevant subgroup variables in the current study, although subgroup analyses may also be conducted by other types of variables. Subgroup effect in a trial is defined as the interaction of treatment with a subgroup variable, or the difference in treatment effect across subgroups [5]. In the current study, subgroup effect is measured by the ratio of relative treatment effects between subgroups (e.g. ratio of odds ratios, ratio of hazard ratios) for binary outcomes, and the difference in absolute mean differences or standardised mean differences between subgroups for continuous outcomes.

### Cumulative subgroup analysis

For estimating overall treatment effects, cumulative meta-analysis consists of a series of repeated meta-analyses after adding data from each new trial chronologically [15]. It can be used to reveal the contribution of individual trials to the overall estimate, and to identify the earliest time at which the pooled effect becomes statistically significant. We applied the method of cumulative meta-analysis to investigate changes in estimates of subgroup effects over time. Statistical methods for cumulative subgroup analysis are presented in Additional file 1.

### Case studies

We searched PubMed to identify recently published IPD meta-analyses with sufficient data to perform cumulative subgroup analysis according to year of publication

(see Additional file 2 for search strategy). Two cases were presented in detail to illustrate the usefulness of cumulative subgroup analysis. The first case study used data from an IPD meta-analysis of beta blockers for heart failure [16], and the second case study used data from an IPD meta-analysis that compared treatment effects for connective tissue disease-associated pulmonary arterial hypertension (PAH) and for idiopathic PAH [17]. Data from the two case studies were used to conduct cumulative subgroup analyses. The relevant subgroup analyses in publications of clinical trials and aggregate data meta-analyses were also examined, and compared with the results of the IPD meta-analyses.

### Computer simulations

Computer simulations were conducted to assess true and false positive rates of conventional and cumulative subgroup analyses in trials using different input parameters in terms of the assumed subgroup effects and sample size (see Additional file 3 for simulation methods and input values). A series of sequentially ordered RCTs were simulated. For each of the simulated trials, the computer programme randomly generated individual patients according to distributions of assumed values of input parameters, and the simulated individual patients were randomly allocated to a control and a treatment arm. Simulated patients belong to different subgroups according to four independent characteristics ( $X_1$ ,  $X_2$ ,  $X_3$  and  $X_4$ ) at baseline, and subgroup analyses were conducted in each simulated trial by these baseline characteristics. It is assumed that only  $X_1$  is an effect modifier, and the other three factors ( $X_2$ ,  $X_3$ , and  $X_4$ ) are not truly associated with the treatment effect. The treatment effect for patients with  $X_1$  (i.e.  $X_1 = 1$ ) is greater than patients without  $X_1$  (i.e.  $X_1 = 0$ ). The simulated subgroup effect is measured by ratio of odds ratios (ROR), which is assumed to be ROR less than 1.0 for patients with  $X_1 = 1$  versus  $X_1 = 0$ , and ROR = 1.0 for the same comparisons of patients for  $X_2$ ,  $X_3$  and  $X_4$ . For example, if ROR = 0.8 for patients with  $X_1$  and OR = 0.7 for patients without  $X_1$ , the treatment effect for patients with  $X_1$  will be OR =  $0.7 \times 0.8 = 0.56$ . We used logistic regression model with an interaction term to conduct subgroup analyses in the simulated trials. In the simulation we calculated  $P$  values for subgroup effect each time a new trial was added to the analysis, and the observation of any statistically significant subgroup effect was not used to stop the cumulative subgroup analysis early before all relevant trials were included. For each scenario, 5000 replications were simulated to estimate the statistical power and the rate of type I error in subgroup analyses.

Data analyses and computer simulations were performed using Stata/IC 13 and R language [18]. Statistical significance was defined as two-sided  $P \leq 0.05$ .

**Results**

**Case study 1: beta blockers for heart failure**

Published in 2014, an IPD meta-analysis of 10 RCTs found that the use of beta blockers reduced all-cause mortality for patients with heart failure and sinus rhythm, but not for those with heart failure plus atrial fibrillation at baseline [16]. The subgroup effect was statistically significant ( $P = 0.002$  for interaction with baseline rhythm). Using reported hazard ratios for patients with sinus rhythm and those with atrial fibrillation, we estimated the ratio of hazard ratios (RHR) between the two subgroups in each of the 10 RCTs (Fig. 1). There was no statistically significant heterogeneity in RHRs across RCTs ( $I^2 = 22.7\%$ ,  $P = 0.23$ ), with a pooled RHR of 0.77 (95% CI, 0.64–0.93;  $P = 0.007$ ). Results of cumulative subgroup analysis revealed that the subgroup effect reached statistical significance by 1999 (pooled RHR = 0.67; 95% CI, 0.46–0.98;  $P < 0.05$ ), when five RCTs had been published with a total of 5180 patients. The overall subgroup effect had not been materially changed since 1999, with five more RCTs and 13,075 more patients (Fig. 1). That is, the significant subgroup effect could have been detected in cumulative subgroup analysis 15 years earlier and with 71.6% less patients than the IPD meta-analysis published in 2014 [16] which first reported it.

Subgroup analysis stratified by baseline rhythm was not reported in publications of the main results in the 10 RCTs [19–28]. After the publication of the main results, subgroup results by cardiac rhythm at baseline were presented in subsequent reports of the four trials [29–32]. The first was published in 2001 [32], and found statistically significant interaction between treatment and rhythm ( $P < 0.01$ ) using data from the CIBIS-II trial [23]. Although there were numerous aggregate data meta-analyses on beta blockers for heart failure, we identified only one

meta-analysis, published in 2013, that performed subgroup analysis by baseline rhythm [33], and which included published data from four trials [29–32].

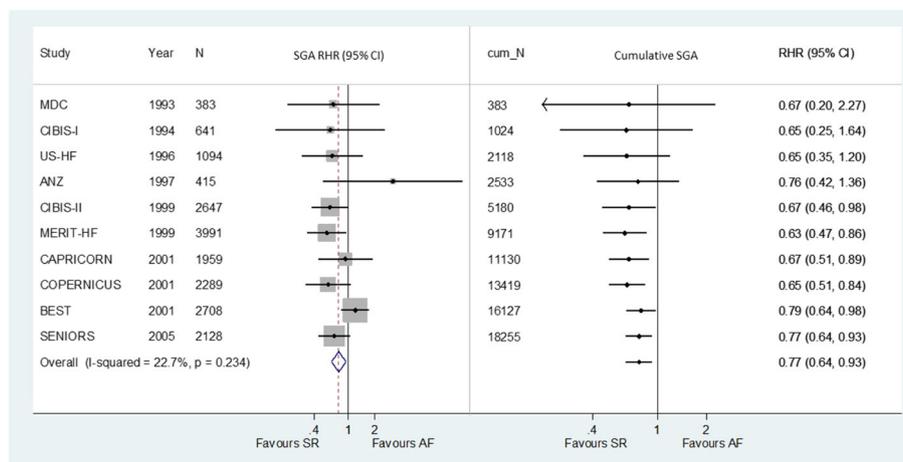
**Case study 2: treatment for pulmonary arterial hypertension**

An IPD meta-analysis included 11 placebo-controlled RCTs that evaluated drugs for pulmonary arterial hypertension (PAH) [17]. For the improvement in 6-minute walk distance, the IPD meta-analysis found that treatment for connective tissue disease-associated PAH was less effective than for idiopathic PAH (Fig. 2). Using data from this IPD meta-analysis, cumulative subgroup analysis revealed that the interaction between treatment and diagnosis became statistically significant ( $P < 0.05$ ) by 2008, when eight RCTs involving a total of 1644 patients had been published (Fig. 2). The significant subgroup effect was detectable 6 years earlier, and with 1118 less patients, than the IPD meta-analysis which first reported it.

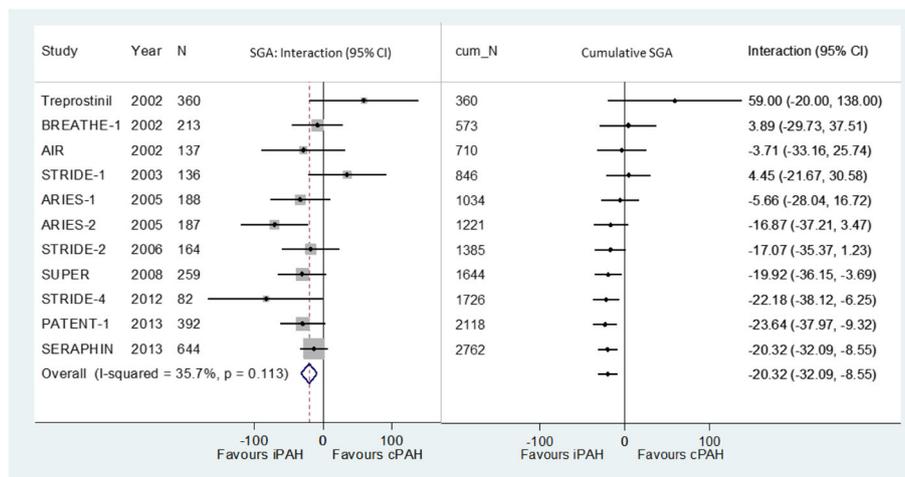
Although subgroup analysis by aetiology was mentioned in most of the included trials, no significant subgroup effect by diagnosis was reported in original publications of the included RCTs [34–43]. For example, data from the ARIES-2 trial showed a statistically significant subgroup effect (Fig. 2), which was not reported in the original publication of the trial [40]. Subgroup analysis by aetiology was not conducted in any of aggregate data meta-analyses of trials that recruited patients with both connective tissue disease-associated PAH and idiopathic PAH [44–47].

**Results of computer simulations**

Table 1 shows rates of true and false positive subgroup effects ( $P \leq 0.05$ ) observed in simulated trials under various scenarios. As expected, the true positive rates



**Fig. 1** Beta blockers for heart failure: meta-analysis and cumulative subgroup analysis for differences in treatment effect between patients with sinus rhythm (SR) and those with atrial fibrillation (AF) at baseline



**Fig. 2** Meta-analysis and cumulative subgroup analysis of differences in treatment effect between idiopathic (iPAH) and connective tissue disease-associated pulmonary arterial hypertension (cPAH)

(statistical power) are associated with the subgroup effect (ROR), sample size, event rate in the control arm, and the proportion of patients in the  $X_1$  subgroup. The false positive rates (type I error) for each of the three variables ( $X_2$ ,  $X_3$  and  $X_4$ ) are about 5%, corresponding to the definition of statistical significance used ( $P \leq 0.05$ ). However, the false positive rate will be inflated (to about 14.3%) if it is calculated by identifying at least one false positive subgroup effect among all subgroup analyses by  $X_2$ ,  $X_3$  and  $X_4$ .

**Table 1** True and false positive rates of significant subgroup effects in simulated trials under different scenarios

Variables	Input values	True positive	False positive				Any of $X_2$ , $X_3$ , $X_4$
			$X_1$	$X_2$	$X_3$	$X_4$	
Ratio of odds ratios (ROR)	0.90	6.2%	4.9%	5.0%	5.0%	14.2%	
	0.80	10.3%	5.0%	5.1%	4.9%	14.2%	
	0.70	17.9%	4.8%	5.0%	5.0%	14.1%	
No. of patients per arm	200	7.5%	4.9%	4.9%	5.1%	14.1%	
	400	10.1%	5.0%	5.0%	4.8%	14.2%	
	600	12.6%	5.1%	5.0%	4.8%	14.2%	
Event rate in the control arm	0.20	8.7%	4.9%	5.1%	5.0%	14.3%	
	0.30	10.2%	5.2%	5.1%	5.1%	14.6%	
	0.40	11.3%	5.1%	4.9%	5.1%	14.4%	
Proportion of patients belonging to a subgroup	20%	8.0%	4.6%	5.0%	4.9%	14.2%	
	30%	9.3%	5.1%	4.9%	4.9%	14.4%	
	40%	10.0%	5.3%	4.9%	5.1%	14.5%	

Except for varying input values shown in the table, the following basic input parameters were used for other variables: No. of trials = 10; No. of patients per arm = 400; OR = 0.7; event rate in the control arm = 0.3; heterogeneity variance = 0.01; ROR = 0.8; proportion of patients with  $X_1 = 0.4$ . Positive rates were calculated based on 5000 simulations for each scenario (with a given set of input parameters)

Results of the cumulative subgroup analyses that used data from simulated trials are shown in Table 2. Given the assumed subgroup effect and sample size, the true positive rate was only about 10% with a single trial, 15% with two trials, and increased to about 80% with 18 trials (Table 2). For individual variables that are not associated with the treatment effect, the false positive rates in cumulative subgroup analyses were about 5%, corresponding to the defined statistical significance at  $P \leq 0.05$  (Table 2).

### Discussion

Although subgroup analyses in randomised trials have well-known limitations, such as inadequate statistical power and inflated false positive rate, identification of subgroup effects is important for clinical practice and further research. The results of the two case studies and computer simulations presented in this paper indicate that cumulative subgroup analysis should be used to overcome limitations of isolated subgroup analyses in trials, and to encourage appropriate conduct, complete reporting and timely synthesis of subgroup analyses in clinical trials.

The detection of differences in effect between subgroups usually requires larger sample sizes than the evaluation of the overall treatment effect, unless the effects are in opposite directions for different subgroups. The statistical power to detect meaningful subgroup effects is unlikely to be sufficient in a single trial, and it is necessary to shift the focus from subgroup analyses in separate trials to subgroup analyses involving all related trials. Although IPD meta-analyses have been increasingly used for this purpose, cumulative subgroup analysis may detect subgroup effects much earlier than IPD meta-analyses. The early identification of important

**Table 2** True and false positive rates in cumulative subgroup analysis using data from simulated trials

No. of trials sequentially included	No. of total patients involved	True positive	False positive		
		X1	X2	X3	X4
1	800	9.8%	4.8%	5.4%	5.2%
2	1600	14.8%	4.7%	5.3%	4.5%
3	2400	19.7%	5.1%	5.5%	5.0%
4	3200	25.7%	5.1%	5.5%	5.3%
5	4000	30.7%	4.9%	5.4%	4.6%
6	4800	36.4%	5.0%	5.3%	4.9%
7	5600	41.3%	5.4%	5.5%	4.6%
8	6400	46.0%	5.5%	5.2%	4.7%
9	7200	50.7%	5.5%	5.4%	4.5%
10	8000	54.4%	5.2%	5.2%	4.9%
11	8800	58.3%	5.5%	5.3%	4.9%
12	9600	61.6%	5.2%	4.9%	4.5%
13	10,400	65.2%	5.2%	5.3%	4.7%
14	11,200	68.9%	5.5%	5.3%	4.9%
15	12,000	71.9%	5.3%	5.5%	4.9%
16	12,800	74.4%	5.5%	5.4%	4.6%
17	13,600	77.1%	5.7%	5.5%	4.5%
18	14,400	79.6%	5.7%	5.1%	4.8%
19	15,200	82.0%	5.9%	4.7%	4.7%
20	16,000	83.8%	5.7%	4.6%	4.7%

Except for varying input values shown in the table, the following basic input parameters were used for other variables: No. of patients per arm = 400; OR = 0.7; event rate in the control arm = 0.3; heterogeneity variance = 0.01; ROR = 0.8; proportion of patients with  $X_1$ ,  $X_2$ ,  $X_3$  and  $X_4$  = 0.4. Positive rates were based on 5000 simulations

subgroup effects may help clinicians in patient care, and to inform the design and analysis of further research [48]. After the publication of IPD meta-analysis on beta blockers for heart failure in 2014 [16], there are still debates about the use of beta blockers for patients with heart failure and atrial fibrillation. For example, the results of the IPD meta-analysis was dismissed as a “retrospective subgroup analysis” in the European Society of Cardiology 2016 Guidelines for the Diagnosis and Treatment of Acute and Chronic Heart Failure [49]. Had the cumulative subgroup analysis been reported by 1999, the observed subgroup effect would have been prospectively investigated in the subsequent large scale trials, and potential mechanisms might have been better investigated.

Inflated false positive rate (type I error) is often used as a reason to restrict the conduct of multiple subgroup analyses in clinical trials. However, inflated false positive findings depend on the following two conditions: selective reporting of statistically significant results of multiple subgroup analyses, and a variety of significant subgroup effects when defined by different patient characteristics.

The first problem, selective reporting, is not unique to subgroup analysis [50]. It can be addressed by complete reporting of all subgroup analyses conducted, and by clearly stating whether the subgroup analyses reported are pre-specified or post hoc. In addition, the cumulative subgroup analysis should not be stopped early when a statistically significant subgroup effect is observed, particularly at its early stage with only a few included trials. Irrespective of currently estimated subgroup effects, data from all subsequent trials conducted for various reasons should be continuously added to the cumulative subgroup analysis. The second problem may be of limited relevance in practice because it makes little sense to merge subgroup effects defined by different patient variables. For example, purely by chance, the treatment effect may be associated with age in a trial, and with different baseline variables in other trials. The results of subgroup effect by age, or by another variable, should be interpreted separately from other subgroup effects and based on pooled data from all related trials. For the same subgroup variable, the rate of false positive subgroup effect will not be inflated, and will correspond well with the statistical significance level adopted in both conventional and cumulative subgroup analyses (Tables 1 and 2). In addition, the possible harms due to false positive subgroup effects in individual trials will be minimal in practice when clinical guidelines are developed after rigorously assessing the validity of all available evidence [51].

Subgroup analyses in clinical trials may be used to test or generate hypotheses on subgroup effects [2, 52]. Given limited statistical power and lack of clear prior understanding of important subgroup variables, subgroup analyses should be generally considered as hypothesis-generating when single trials are considered in isolation. With the concept of cumulative subgroup analysis, a subgroup analysis in the first trial or a few early trials is for the purpose of hypothesis generation, but the same subgroup analysis in subsequent trials may be considered as hypothesis testing. Because it may be difficult to decide whether a subgroup analysis is hypothesis-testing or hypothesis-generating, a Bayesian approach may provide a more convenient theoretical framework for cumulative subgroup analysis [53]. Analogous to the Bayesian method of combining prior and new evidence, a cumulative subgroup analysis continuously incorporates existing information with data from a new trial.

Data on patient characteristics at baseline are routinely collected in randomised controlled trials for multiple purposes, including a description of study population, assessment of comparability of trial groups, adjusting for possible confounding factors, and conduct of subgroup analyses [54]. During the past several decades, subgroup analyses in trials have not been encouraged [10]. Consequently, data on baseline characteristics collected in

trials have been under-used, or completely wasted, for the purpose of subgroup analysis. According to the recommended criteria for credible subgroup analyses [9], the number of subgroup analyses in a trial should be no more than five. Another recommended criterion for credible subgroup analyses is whether subgroup effects across related studies are consistent [9], which will be impossible to assess if the same subgroup analysis has not been conducted and reported in other related studies. The argument that only a few pre-specified subgroup analyses should be conducted in a trial conflicts with the need to compare and combine results of the same subgroup analysis from all related trials [3]. The current emphasis on avoiding false positive subgroup effects has restricted the conduct and reporting of exploratory subgroup analyses in trials, resulting in a waste of research data and missing opportunities of detecting subgroup effects that are meaningful for clinical practice or additional research [2, 48].

Inaccessibility and lack of full information are avoidable waste in biomedical research [55]. Subgroup analyses in meta-analyses using published data are often very limited or impossible due to inadequate reporting of results of subgroup analysis in trials [2, 12]. The development of IPD meta-analyses has facilitated the identification of important subgroup effects. However, trial data on subgroup effects have been wasted before the conduct of IPD meta-analyses, and continue to be wasted where IPD meta-analyses remain unavailable, although the magnitude of such waste is currently unclear. Therefore, more IPD meta-analyses of existing trials should be conducted to identify meaningful subgroup effects. In future, exploratory subgroup analyses using full data on patient characteristics at baseline should be encouraged.

Subgroup analyses in clinical trials should be conducted using appropriate statistical tests of interactions, and reporting of subgroup analyses should be complete, with sufficient information to be included in cumulative analyses. To conduct cumulative subgroup analysis, the same or similar definitions of subgroups of interest need to be adopted in related clinical trials. First, cumulative subgroup analyses should be taken into account in making decisions about data collection at the design stage. Patient subgroups could be defined according to patient baseline characteristics using data routinely collected in clinical trials. Ideally, increased sharing of trial data may enable prospective IPD meta-analysis with cumulative subgroup analyses that starts when data from the first two RCTs are available and is repeated when a new RCT is completed. Prospective IPD meta-analysis will also allow the subgroups to be defined in the same way, for example, using the same cut-points between subgroups.

The usefulness of cumulative subgroup analysis will be limited when the number of related trials is very small.

Our search of PubMed (see Additional file 2 for search strategy) identified 60 IPD meta-analyses published in 2014 and only three provided sufficient data for cumulative subgroup analysis. We discussed in detail only two cases to illustrate the usefulness of cumulative subgroup analyses for clinical practice and further research. We believe that our study will inspire others to conduct more cumulative subgroup analyses using data collected in existing and future IPD meta-analyses.

## Conclusions

Without selective reporting, multiple tests of the same interaction between a treatment and a subgroup variable in related trials are not inherently associated with inflated rate of false subgroup effects. To avoid waste of research data and facilitate the early detection of important subgroup effects, subgroup analyses should be more widely conducted in clinical trials and completely reported. Subgroup analyses need to be consistently conducted across related trials using appropriate statistical methods, and their reporting should be complete, with sufficient data to be included in cumulative subgroup analysis.

## Additional files

**Additional file 1:** Methods of cumulative subgroup analysis. (DOCX 14 kb)

**Additional file 2:** Search strategy to identify cases from published individual patient data (IPD) meta-analyses. (DOCX 26 kb)

**Additional file 3:** Methods for computer simulations of subgroup analysis in the context of cumulative meta-analysis of randomised controlled trials. (DOCX 22 kb)

## Abbreviations

IPD: Individual patient data; OR: Odds ratio; PAH: Pulmonary arterial hypertension; RCT: Randomised controlled trial; RHR: Ratio of hazard ratios; ROR: Ratio of odds ratios

## Acknowledgements

Not applicable.

## Funding

There was no specific funding for this study.

## Availability of data and material

All data generated or analysed during this study are included in this published article and its supplementary information files.

## Authors' contributions

FS conceived the idea, searched and assessed individual patient data meta-analyses, analysed data, conducted computer simulations, and drafted the manuscript. MOB provided methodological advice and critically commented on the draft manuscript. Both authors read and approved the final manuscript.

## Competing interests

The authors declare that they have no competing interests.

## Consent for publication

Not applicable.

## Ethics approval and consent to participate

Not applicable.

Received: 19 September 2016 Accepted: 9 November 2016

Published online: 15 December 2016

## References

- Rothwell PM. Can overall results of clinical trials be applied to all patients? *Lancet*. 1995;345(8965):1616–9.
- Varadhan R, Segal JB, Boyd CM, Wu AW, Weiss CO. A framework for the analysis of heterogeneity of treatment effect in patient-centered outcomes research. *J Clin Epidemiol*. 2013;66(8):818–25.
- Rothwell PM. Treating individuals 2. Subgroup analysis in randomised controlled trials: importance, indications, and interpretation. *Lancet*. 2005;365(9454):176–86.
- Sun X, Briel M, Busse JW, You JJ, Akl EA, Mejza F, Bala MM, Bassler D, Mertz D, Diaz-Granados N, et al. The influence of study characteristics on reporting of subgroup analyses in randomised controlled trials: systematic review. *BMJ*. 2011;342:d1569.
- Sun X, Ioannidis JP, Agoritsas T, Alba AC, Guyatt G. How to use a subgroup analysis: users' guide to the medical literature. *JAMA*. 2014;311(4):405–11.
- Yusuf S, Wittes J, Probstfield J, Tyroler HA. Analysis and interpretation of treatment effects in subgroups of patients in randomized clinical trials. *JAMA*. 1991;266(1):93–8.
- Pocock SJ, Assmann SE, Enos LE, Kasten LE. Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems. *Stat Med*. 2002;21(19):2917–30.
- Oxman AD, Guyatt GH. A consumer's guide to subgroup analyses. *Ann Intern Med*. 1992;116(1):78–84.
- Sun X, Briel M, Busse JW, You JJ, Akl EA, Mejza F, Bala MM, Bassler D, Mertz D, Diaz-Granados N, et al. Credibility of claims of subgroup effects in randomised controlled trials: systematic review. *BMJ*. 2012;344, e1553.
- Feinstein AR. The problem of cogent subgroups: a clinicostatistical tragedy. *J Clin Epidemiol*. 1998;51(4):297–9.
- Oxman AD. Subgroup analyses: the devil is in the interpretation. *BMJ*. 2012;344, e2022.
- Thompson SG, Higgins JP. Treating individuals 4: can meta-analysis help target interventions at individuals most likely to benefit? *Lancet*. 2005;365(9456):341–6.
- Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. *BMJ*. 2010;340:c221.
- Huang Y, Tang J, Tam WW, Mao C, Yuan J, Di M, Yang Z. Comparing the overall result and interaction in aggregate data meta-analysis and individual patient data meta-analysis. *Medicine (Baltimore)*. 2016;95(14), e3312.
- Lau J, Antman EM, Jimenez-Silva J, Kupelnick B, Mosteller F, Chalmers TC. Cumulative meta-analysis of therapeutic trials for myocardial infarction. *N Engl J Med*. 1992;327(4):248–54.
- Kotecha D, Holmes J, Krum H, Altman DG, Manzano L, Cleland JG, Lip GY, Coats AJ, Andersson B, Kirchhof P, et al. Efficacy of beta blockers in patients with heart failure plus atrial fibrillation: an individual-patient data meta-analysis. *Lancet*. 2014;384(9961):2235–43.
- Rhee RL, Gabler NB, Sangani S, Praestgaard A, Merkel PA, Kawut SM. Comparison of treatment response in idiopathic and connective tissue disease-associated pulmonary arterial hypertension. *Am J Respir Crit Care Med*. 2015;192(9):1111–7.
- R Development Core Team: R. A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna. 2008. <http://www.R-project.org>. Accessed 21 Nov 2016.
- Waagstein F, Bristow MR, Swedberg K, Camerini F, Fowler MB, Silver MA, Gilbert EM, Johnson MR, Goss FG, Hjalmarson A. Beneficial effects of metoprolol in idiopathic dilated cardiomyopathy. Metoprolol in Dilated Cardiomyopathy (MDC) Trial Study Group. *Lancet*. 1993;342(8885):1441–6.
- CIBIS-I. A randomized trial of beta-blockade in heart failure. The Cardiac Insufficiency Bisoprolol Study (CIBIS). CIBIS Investigators and Committees. *Circulation*. 1994;90(4):1765–73.
- Packer M, Bristow MR, Cohn JN, Colucci WS, Fowler MB, Gilbert EM, Shusterman NH. The effect of carvedilol on morbidity and mortality in patients with chronic heart failure. U.S. Carvedilol Heart Failure Study Group. *N Engl J Med*. 1996;334(21):1349–55.
- Australia/New Zealand Heart Failure Research Collaborative Group. Randomised, placebo-controlled trial of carvedilol in patients with congestive heart failure due to ischaemic heart disease. *Lancet*. 1997;349(9049):375–80.
- CIBIS-II. The Cardiac Insufficiency Bisoprolol Study II (CIBIS-II): a randomised trial. *Lancet*. 1999;353(9146):9–13.
- MERIT-HF. Effect of metoprolol CR/XL in chronic heart failure: Metoprolol CR/XL Randomised Intervention Trial in Congestive Heart Failure (MERIT-HF). *Lancet*. 1999;353(9169):2001–7.
- CAPRICORN. Effect of carvedilol on outcome after myocardial infarction in patients with left-ventricular dysfunction: the CAPRICORN randomised trial. *Lancet*. 2001;357(9266):1385–90.
- Packer M, Coats AJ, Fowler MB, Katus HA, Krum H, Mohacs P, Rouleau JL, Tendera M, Castaigne A, Roecker EB, et al. Effect of carvedilol on survival in severe chronic heart failure. *N Engl J Med*. 2001;344(22):1651–8.
- BEST. A trial of the beta-blocker bucindolol in patients with advanced chronic heart failure. *N Engl J Med*. 2001;344(22):1659–67.
- Flather MD, Shibata MC, Coats AJ, Van Veldhuisen DJ, Parkhomenko A, Borbola J, Cohen-Solal A, Dumitrascu D, Ferrari R, Lechat P, et al. Randomized trial to determine the effect of nebivolol on mortality and cardiovascular hospital admission in elderly patients with heart failure (SENIORS). *Eur Heart J*. 2005;26(3):215–25.
- Mulder BA, van Veldhuisen DJ, Crijns HJ, Bohm M, Cohen-Solal A, Babalis D, Roughton M, Flather MD, Coats AJ, Van Gelder IC. Effect of nebivolol on outcome in elderly patients with heart failure and atrial fibrillation: insights from SENIORS. *Eur J Heart Fail*. 2012;14(10):1171–8.
- Joglar JA, Acosta AP, Shusterman NH, Ramaswamy K, Kowal RC, Barbera SJ, Hamdan MH, Page RL. Effect of carvedilol on survival and hemodynamics in patients with atrial fibrillation and left ventricular dysfunction: retrospective analysis of the US Carvedilol Heart Failure Trials Program. *Am Heart J*. 2001;142(3):498–501.
- van Veldhuisen DJ, Aass H, El Alaff D, Dunselman PH, Gullestad L, Halinen M, Kjekshus J, Ohlsson L, Wedel H, Wikstrand J. Presence and development of atrial fibrillation in chronic heart failure. Experiences from the MERIT-HF Study. *Eur J Heart Fail*. 2006;8(5):539–46.
- Lechat P, Hulot JS, Escolano S, Mallet A, Leizorovicz A, Werhlen-Grandjean M, Pochmalicki G, Dargie H. Heart rate and cardiac rhythm relationships with bisoprolol benefit in chronic heart failure in CIBIS II Trial. *Circulation*. 2001;103(10):1428–33.
- Rienstra M, Damman K, Mulder BA, Van Gelder IC, McMurray JJ, Van Veldhuisen DJ. Beta-blockers and outcome in heart failure and atrial fibrillation: a meta-analysis. *JACC Heart Fail*. 2013;1(1):21–8.
- Simonneau G, Barst RJ, Galie N, Naeije R, Rich S, Bourge RC, Keogh A, Oudiz R, Frost A, Blackburn SD, et al. Continuous subcutaneous infusion of treprostinil, a prostacyclin analogue, in patients with pulmonary arterial hypertension: a double-blind, randomized, placebo-controlled trial. *Am J Respir Crit Care Med*. 2002;165(6):800–4.
- Olschewski H, Simonneau G, Galie N, Higenbottam T, Naeije R, Rubin LJ, Nikkho S, Speich R, Hoepfer MM, Behr J, et al. Inhaled iloprost for severe pulmonary hypertension. *N Engl J Med*. 2002;347(5):322–9.
- Rubin LJ, Badesch DB, Barst RJ, Galie N, Black CM, Keogh A, Pulido T, Frost A, Roux S, Leconte I, et al. Bosentan therapy for pulmonary arterial hypertension. *N Engl J Med*. 2002;346(12):896–903.
- Barst RJ, Langleben D, Frost A, Horn EM, Oudiz R, Shapiro S, McLaughlin V, Hill N, Tapson VF, Robbins IM, et al. Sitaxsentan therapy for pulmonary arterial hypertension. *Am J Respir Crit Care Med*. 2004;169(4):441–7.
- Galie N, Ghofrani HA, Torbicki A, Barst RJ, Rubin LJ, Badesch D, Fleming T, Parpia T, Burgess G, Branzi A, et al. Sildenafil citrate therapy for pulmonary arterial hypertension. *N Engl J Med*. 2005;353(20):2148–57.
- Barst RJ, Langleben D, Badesch D, Frost A, Lawrence EC, Shapiro S, Naeije R, Galie N. Treatment of pulmonary arterial hypertension with the selective endothelin-A receptor antagonist sitaxsentan. *J Am Coll Cardiol*. 2006;47(10):2049–56.
- Galie N, Olschewski H, Oudiz RJ, Torres F, Frost A, Ghofrani HA, Badesch DB, McGoon MD, McLaughlin VV, Roecker EB, et al. Ambrisentan for the treatment of pulmonary arterial hypertension: results of the ambrisentan in pulmonary arterial hypertension, randomized, double-blind, placebo-controlled, multicenter, efficacy (ARIES) study 1 and 2. *Circulation*. 2008;117(23):3010–9.
- Sandoval J, Torbicki A, Souza R, Ramirez A, Kurzyna M, Jardim C, Jerjes-Sanchez Diaz C, Teal SA, Hwang LJ, Pulido T. Safety and efficacy of sitaxsentan 50 and 100 mg in patients with pulmonary arterial hypertension. *Pulm Pharmacol Ther*. 2012;25(1):33–9.
- Ghofrani HA, Galie N, Grimminger F, Grunig E, Humbert M, Jing ZC, Keogh AM, Langleben D, Kilama MO, Fritsch A, et al. Riociguat for the treatment of pulmonary arterial hypertension. *N Engl J Med*. 2013;369(4):330–40.

43. Pulido T, Adzerikho I, Channick RN, Delcroix M, Galie N, Ghofrani HA, Jansa P, Jing ZC, Le Brun FO, Mehta S, et al. Macitentan and morbidity and mortality in pulmonary arterial hypertension. *N Engl J Med*. 2013;369(9):809–18.
44. Liu HL, Chen XY, Li JR, Su SW, Ding T, Shi CX, Jiang YF, Zhu ZN. Efficacy and safety of PAH-specific therapy in pulmonary arterial hypertension: a meta-analysis of randomized clinical trials. *Chest*. 2016;150(2):353–66.
45. Galie N, Manes A, Negro L, Palazzini M, Bacchi-Reggiani ML, Branzi A. A meta-analysis of randomized controlled trials in pulmonary arterial hypertension. *Eur Heart J*. 2009;30(4):394–403.
46. Coeytaux RR, Schmit KM, Kraft BD, Kosinski AS, Mingo AM, Vann LM, Gilstrap DL, Hargett CW, Heidenfelder B, Dolor RJ, et al. Comparative effectiveness and safety of drug therapy for pulmonary arterial hypertension: a systematic review and meta-analysis. *Chest*. 2014;145(5):1055–63.
47. Zhang HD, Zhang R, Jiang X, Sun K, Wu DC, Jing ZC. Effects of oral treatments on clinical outcomes in pulmonary arterial hypertension: a systematic review and meta-analysis. *Am Heart J*. 2015;170(1):96–103. 103.e1–14.
48. Hemila H, Kaprio J. Subgroup analysis of large trials can guide further research: a case study of vitamin E and pneumonia. *Clin Epidemiol*. 2011;3:51–9.
49. Ponikowski P, Voors AA, Anker SD, Bueno H, Cleland JG, Coats AJ, Falk V, Gonzalez-Juanatey JR, Harjola VP, Jankowska EA, et al. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC) Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. *Eur Heart J*. 2016;37(27):2129–200.
50. Chan AW, Hrobjartsson A, Haahr MT, Gotzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA*. 2004;291(20):2457–65.
51. Alonso-Coello P, Oxman AD, Moberg J, Brignardello-Petersen R, Akl EA, Davoli M, Treweek S, Mustafa RA, Vandvik PO, Meerpohl J, et al. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 2: Clinical practice guidelines. *BMJ*. 2016;353:i2089.
52. Burke JF, Sussman JB, Kent DM, Hayward RA. Three simple rules to ensure reasonably credible subgroup analyses. *BMJ*. 2015;351:h5651.
53. Lau J, Schmid CH, Chalmers TC. Cumulative meta-analysis of clinical trials builds evidence for exemplary medical care. *J Clin Epidemiol*. 1995;48(1):45–57. discussion 59–60.
54. Assmann SF, Pocock SJ, Enos LE, Kasten LE. Subgroup analysis and other (mis)uses of baseline data in clinical trials. *Lancet*. 2000;355(9209):1064–9.
55. Chan AW, Song F, Vickers A, Jefferson T, Dickersin K, Gotzsche PC, Krumholz HM, Ghersi D, van der Worp HB. Increasing value and reducing waste: addressing inaccessible research. *Lancet*. 2014;383(9913):257–66.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

