

Problems with use of composite end points in cardiovascular trials: systematic review of randomised controlled trials

I Ferreira-González,¹ J W Busse,² D Heels-Ansdell, V M Montori, E A Akl, D M Bryant, J Alonso, R Jaeschke, H J Schünemann, G Permanyer-Miralda, A Domingo-Salvany, G H Guyatt

EDITORIAL by Freemantle and Calvert

¹Departament de Medicina, Universitat Autònoma de Barcelona, and Hospital Vall d'Hebron, Barcelona 08035, Spain

²Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, ON, Canada L8N 3Z5

Correspondence to: J W Busse
j.busse@utoronto.ca

BMJ 2007;334:786-8

doi: 10.1136/bmj.39136.682083.AE

ABSTRACT

Objective To explore the extent to which components of composite end points in randomised controlled trials vary in importance to patients, the frequency of events in the more and less important components, and the extent of variability in the relative risk reductions across components.

Design Systematic review of randomised controlled trials.

Data sources Cardiovascular randomised controlled trials published in the *Lancet*, *Annals of Internal Medicine*, *Circulation*, *European Heart Journal*, *JAMA*, and *New England Journal of Medicine*, from 1 January 2002 to 30 June 2003. Component end points of composite end points were categorised according to importance to patients as fatal, critical, major, moderate, or minor.

Results Of 114 identified randomised controlled trials that included a composite end point of importance to patients, 68% (n=77) reported complete component data for the primary composite end point; almost all (98%; n=112) primary composite end points included a fatal end point. Of 84 composite end points for which component data were available, 54% (n=45) showed large or moderate gradients in both importance to patients and magnitude of effect across components. When analysed by categories of importance to patients, the most important components were associated with lower event rates in the control group (medians of 3.3-3.7% for fatal, critical, and major outcomes; 12.3% for moderate outcomes; and 8.0% for minor outcomes). Components of greater importance to patients were associated with smaller treatment effects than less important ones (relative risk reduction of 8% for death and 33% for components of minor importance to patients).

Conclusion The use of composite end points in cardiovascular trials is frequently complicated by large gradients in importance to patients and in magnitude of the effect of treatment across component end points. Higher event rates and larger treatment effects associated with less important components may result in misleading impressions of the impact of treatment.

INTRODUCTION

Composite end points capture the number of patients who have one or more of several events of interest. Clinical trials, particularly in cardiology,¹ often use composite end points to reduce sample size requirements and to capture the overall impact of therapeutic interventions.² Although composite end points may increase the event rate and thus the statistical power of the study, they may mislead if component end points are of widely differing importance to patients, the number of events in the more important components

is small, and the size of effect differs markedly across components.³ For example, a claim that an intervention reduces a composite of cardiovascular mortality, myocardial infarction, and revascularisation procedures is problematic if most events were revascularisation procedures and treatment had a large effect on revascularisation but not on death or infarction.

To explore the characteristics of composite end points in common use, we reviewed a consecutive sample of randomised controlled trials that investigated cardiology interventions and were published in six prominent journals. We were interested in the extent to which components of composite end points varied in importance to patients, the frequency of events in the more and less important components, and the extent of variability in the relative risk reductions across components.

METHODS

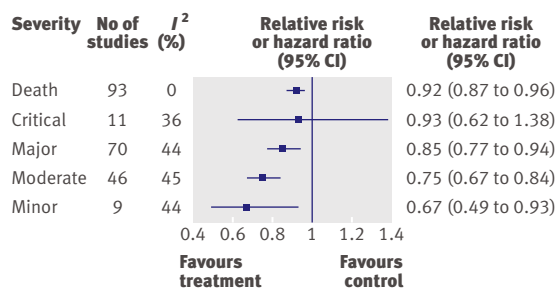
We included parallel group randomised controlled trials that involved humans exposed to any cardiovascular therapeutic intervention and reported at least one composite end point. The target population of the study had to have coronary artery disease, valvular heart disease, arrhythmia, cardiomyopathy, or congestive heart failure on entry. We also included trials investigating primary prevention of cardiovascular disease.

We used Medline to search electronically four high impact general medicine journals (*Lancet*, *Annals of Internal Medicine*, *JAMA*, and *New England Journal of Medicine*) and two leading cardiology journals (*Circulation* and *European Heart Journal*), from 1 January 2002 to 30 June 2003. Eight investigators, working in pairs, used standardised forms to establish if abstracts met the inclusion criteria. The same reviewers independently assessed eligibility of the full text articles with standardised forms and resolved discrepancies by discussion.

Seven reviewers worked in pairs to extract data independently and in duplicate, using a standardised form. Reviewers collected information on content area, interventions tested, sample size, the length of follow-up, and the number of composite end points presented. To avoid confounding, we explored only data associated with each trial's primary composite end point. For each composite end point we extracted the number of components, the effect of the intervention on the composite end point, and the number of events attributed to the composite end point. For the component end points of each composite end point we recorded the effect of the intervention and the number of patients who achieved the outcome.

The full list of authors, their affiliations, and funding are given in the full version of the paper on bmj.com

This article is an abridged version of a paper that was published on bmj.com on 2 April 2007. Cite this version as: *BMJ* 2 April 2007; doi: 10.1136/bmj.39136.682083.AE (abridged text, in print: *BMJ* 2007;334:786-8)



Variability in magnitude of the effect of intervention across categories of importance to patients

Two cardiologists and nine internists independently categorised each of 72 outcomes used as components of composite end points in the eligible trials into five categories according to their importance to patients: I=death, II=critical, III=major, IV=moderate, and V=minor. Estimates of utility associated with the outcomes guided the process.⁴

Data analysis

We calculated, for each of the five categories of outcomes, the median event rate and the interquartile range for the control group as well as the effect of the intervention within the category by using the authors' reporting of relative risk, odds ratio, or hazard ratio. To estimate the effect of the intervention across trials and within each category, we used random effects meta-analyses with an inverse variance approach. We used the I^2 statistic to test heterogeneity to quantify inconsistency among trials.⁵

We considered a large gradient in importance to patients to be present in composite end points combining outcomes from categories I or II (fatal and critical) with outcomes from category V (minor). We considered a moderate gradient to be present when composite end points combined outcomes from categories I or II with outcomes from category IV (moderate) without any component from category V. We assigned a minor or absent gradient to composite end points not included in the other two categories. We assigned a large gradient in the effect of the intervention if the

Additive effect of component end points on composite end points, based on category of importance to patients (in composite end points with moderate/large gradient in importance) (n=46)

Outcome	Most important components (CI, CII)*	Moderately important components added (CIII)†	Less important components added (CIV, CV)‡
Median % (IQR) event rate of composite end point for control group	2.5 (0.8-6.1)	8.7 (5.2-14.5)	21.7 (13.1-33.2)
Median % (IQR) effect of treatment (RRR)	13 (-14-41)	17 (-9-36)	24 (8-42)

CI=category I (death); CII=category II (critical); CIII=category III (major); CIV=category IV (moderate); CV=category V (minor); IQR=interquartile range; RRR=relative risk reduction.

*Reflects event rate in control group and RRR for only most important components included in composite end point.

†Moderate components, when present (n=35), have been included with most important components to calculate event rate in control group and RRR.

‡Event rate in control group and RRR results for full composite end point; reflects addition of least important components to data in previous column.

difference between the smallest and largest reported treatment effects (relative risk, odds ratio, or hazard ratio) was >0.4 , a moderate gradient when the difference was 0.2 to 0.4, and a small gradient when the difference was <0.2 .

Among composite end points with moderate or large gradients in importance to patients, we explored the impact of the outcomes with less importance to patients on both the total event rate for the composite end point in the control group and the magnitude of effect of the intervention (see bmj.com).

RESULTS

Our literature search generated 650 abstracts, from which we identified 242 potentially eligible randomised controlled trials, of which 114 proved eligible on consensus review of the full text publications. Chance corrected agreement on eligibility was excellent ($\kappa=0.90$).

Most trials appeared in *Circulation*, the *Lancet*, and *JAMA*; focused on treatment of coronary disease primarily through pharmacological intervention; and reported only one primary composite end point. Almost all (98%; n=112) composite end points included fatal end points, usually reported as "all cause mortality" (see bmj.com). The median sample size of eligible randomised controlled trials was 840 (interquartile range 238-2334), and the median follow-up time was 1 year (90 days-3.5 years).

Gradient in importance to patients and effect of intervention across components

Most composite end points (56%; n=64) showed either a large (10%; n=11) or moderate (47%; n=53) gradient in importance to patients. Among the 84 composite end points that reported data for at least two of their component end points, the gradient in the effect of the intervention across component end points was usually large (57%; n=48) or moderate (18%; n=15). Of these 84 randomised controlled trials, 45 (54%) included a composite end point with components that exhibited large or moderate gradients in both importance to patients and effect of intervention across components. Many remaining composite end points (32%; n=27) included a large or moderate gradient in either importance to patients (11%; n=9) or treatment effect across components (21%; n=18). Only 14% (n=12) of composite end points reviewed were composed of end points that exhibited either no or a minor gradient in both these aspects.

End points of moderate and minor importance

When analysed by categories of importance to patients, the most important components were associated with lower control group event rates, with medians of 3.3% (interquartile range 1.4-6.9%) for fatal outcomes, 3.3% (2.2-5.2%) for critical end points, and 3.7% (1.6-8.5%) for major outcomes. End points of moderate and minor importance to patients had higher event rates: median 12.3% (2.9-26.7%) for moderate and 8.0% (4.5-26.8%) for minor. Similarly, we found

WHAT IS ALREADY KNOWN ON THIS TOPIC

Clinical trialists use composite end points to increase event rates and statistical power. When the gradient of importance to patients is large, and the more important events are uncommon and show negligible treatment effects, use of composite end points can be misleading.

WHAT THIS STUDY ADDS

Almost half of a sample of recent prominently published cardiovascular trials used composite end points, which often showed large gradients in importance to patients. End points of least importance to patients typically contributed most events.

that pooled effects for fatal and critical outcomes were small, and end points of lesser importance to patients were associated with larger effects (fig).

Of 64 composite end points with moderate or large gradients in importance to patients, 46 had sufficient data to quantify the impact of the less important end points on both the event rates for the composite end point in the control group and the effect of the intervention on the composite end point. The median event rate for the control group when we considered only the most important patient outcomes (fatal and critical end points) was 2.5%. The event rate rose to 8.7% when we added end points of major importance to the composite and to 21.7% when we added end points of moderate and minor importance (table). The magnitude of the treatment effect also rose substantially as less important components were included (table). Of the 46 composite end points with a large or moderate gradient in importance to patients that provided data on component end points, 59% (n=27) were statistically significant ($P<0.05$). However, only seven achieved statistical significance when we considered only components of greater importance to patients (fatal, critical, and major end points), whereas most (20/27) achieved statistical significance only when we added end points of moderate or minor importance to the composite.

DISCUSSION

Our analysis of 114 randomised controlled trials on cardiovascular interventions found that the use of composite end points is common. Reporting is not optimal: authors failed to provide the effect of treatment for all the components in almost a third of the articles. Most trials showed a large or moderate gradient in importance of end points to patients, and in 54% of the 84 trials in which data were available the component end points exhibited substantial gradients in both importance to patients and the effect of treatment across components. Less important components showed higher event rates and larger treatment effects.

Limitations and strengths

Our conclusions depend on clinicians assigning importance to patients to cardiovascular end points. Published data on the utilities that patients attribute to cardiovascular outcomes guided our classification,⁴ and 11 cardiovascular clinicians worked independently on the classification and achieved consensus.

Our analysis of the effect of treatment across components was limited to trials that reported data on component end points in categories I (fatal) to III (major) of importance to patients. This may have led to a biased sample. Our approach was, however, conservative in that when three or more components were present and the joint distribution of results was unavailable we assumed distributions that attributed the maximum number of events to the more important end points.

One might question our application of meta-analytic approaches to data across a wide variety of interventions (fig). The variability in results, represented by the I^2 , proved to be 0% for fatal end points and was below a commonly used threshold of 50% for other end points.⁵

Implications

Use of composite end points appeals to clinical trialists because it increases event rates and statistical power. The problem is the difficulty in interpreting results when the gradient of importance to patients is substantial and a substantial gradient in the size of the treatment effect also exists. Our findings suggest that most composite end points used in cardiovascular randomised controlled trials have substantial gradients in both importance to patients and treatment effects across component end points. Furthermore, less important outcomes provide larger contributions to the composite end point event rate and show larger treatment effects. An important and plausible risk of misleading conclusions associated with the use of composite end points is to attribute reductions in mortality to interventions that do not, in fact, reduce death rates.

Trialists should report complete data on individual component end points to facilitate interpretation; clinicians should view with caution the results of cardiovascular trials that use composite end points. Clinicians and patients are best served when trialists restrict their use of composite end points to end points of similar importance to patients and contexts in which they anticipate that more important end points will contribute a large proportion of study events. If they do not, they risk misleading their audience.

We thank Lisa Buckingham for assistance in creating the database used to capture the information extracted from eligible trials.

Contributors: See bmj.com.

Competing interests: None declared.

Ethics approval: Not needed.

- 1 Freemantle N, Calvert M, Wood J, Eastaugh J, Griffin C. Composite outcomes in randomized trials: greater precision but with greater uncertainty? *JAMA* 2003;289:2554-9.
- 2 Ferreira-González I, Permyer-Miralda G, Busse JW, Bryant DM, Montori VM, Alonso-Coello P, et al. Rational for using primary composite endpoints in clinical trials: a systematic review. *J Clin Epidemiol* (in press).
- 3 Montori VM, Permyer-Miralda G, Ferreira-Gonzalez I, Busse JW, Pacheco-Huergo V, Bryant D, et al. Validity of composite end points in clinical trials. *BMJ* 2005;330:594-6.
- 4 Tengs TO, Wallace A. One thousand health-related quality-of-life estimates. *Med Care* 2000;38:583-637.
- 5 Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557-60.

Accepted: 29 January 2007