BMJ

RESEARCH

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

Ignacio Ferreira-González, research fellow,[1] Jason W Busse, research associate,[3] Diane Heels-Ansdell, statistician,[3] Victor M Montori, associate professor,[5] Elie A Akl, assistant professor,[6] Dianne M Bryant, clinical epidemiologist,[8] Pablo Alonso-Coello, general practitioner,[9] Jordi Alonso, general practitioner,[10] Andrew Worster, associate professor,[3] Suneel Upadhye, associate member,[3] Roman Jaeschke, clinical professor,[4] Holger J Schünemann, associate professor,[7] Gaietà Permanyer-Miralda, senior consultant,[2] Valeria Pacheco-Huergo, research fellow,[1] Antònia Domingo-Salvany, senior scientist,[10] Ping Wu, senior scientist,[11] Edward J Mills, assistant professor,[12] Gordon H Guyatt, professor[3]

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

[2]Cardiology Service, Epidemiology Unit, Hospital General Vall d'Hebron

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

[4]Department of Medicine, McMaster University

[5]Knowledge and Encounter Research Unit, Department of Medicine, Mayo Clinic College of Medicine, Rochester, MN 55905, USA

[6]Department of Medicine and of Social and Preventive Medicine, University at Buffalo, Buffalo, New York 14214, USA

[7]Unit of Clinical Research Development and INFORMAtion Translation/CLARITY Research Team, Department of Epidemiology, Italian National Cancer Institute Regina Elena, Rome 00144, Italy

[8]Faculty of Health Sciences, University of Western Ontario, Elborn College, London, ON, Canada N6G 1H1

[9]Iberoamerican Cochrane Centre, Department of Clinical Epidemiology and Public Health, Hospital de Sant Pau, Barcelona 08041

[10]Health Services Research Unit, Institut Municipal d'Investigació Mèdica (IMIM-hospital del mar), Barcelona E-08003

[11]College of Naturopathic Medicine, Toronto, ON, Canada M2K 1E2

[12]Global Health, Simon Fraser University, Vancouver, BC, Canada V5A 1S6

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

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]

Freemantle and colleagues have highlighted potential advantages and limitations of the use of composite end points.[1] 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 components of greater importance is small, and the magnitude of effect differs markedly across components.[3] For example, a statement that an intervention reduces a composite of cardiovascular mortality, myocardial infarction, and revascularisation procedures is problematic if most of the events were revascularisation procedures and investigators found a large apparent effect of treatment 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. In particular, 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

Eligibility criteria

We included parallel group randomised controlled trials that involved humans exposed to any cardiovascular therapeutic intervention and reported at least one composite end point. We defined a cardiovascular clinical trial as any randomised controlled trial in which the target population of the study had to
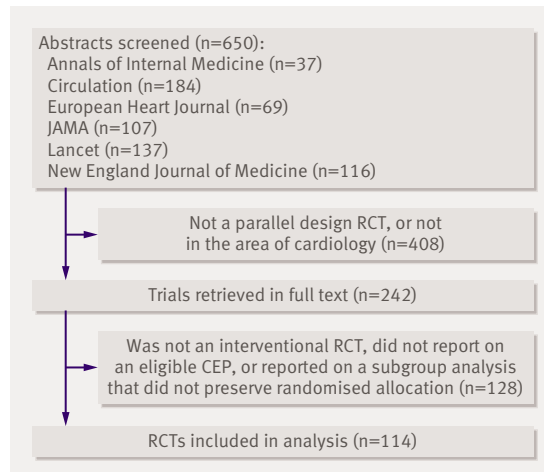
Fig 1 | Stages of systematic review of cardiovascular randomised controlled trials. CEP=composite end point; RCT=randomised controlled trial

have coronary artery disease, valvular heart disease, arrhythmia, cardiomyopathy, or congestive heart failure on entry. We also included randomised controlled trials investigating primary prevention of cardiovascular disease. We excluded trials that reported composite end points with components relating to toxicity or safety or with no outcomes important to patients (that is, including only surrogate outcomes) or subgroup analyses that ignored random allocation.

Literature search

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. We used the publication type function to restrict our search to "randomized controlled trial" and "human" subjects; the National Library of Medicine and the US Cochrane Center have collaborated to use these terms to accurately index randomised controlled trials in the Medline database.[4]

Study selection

Eight investigators (JWB, EAA, DMB, PA-C, AW, SU, VP-H, and AD-S), working in pairs, used standardised forms to establish if abstracts of articles identified in our electronic search were parallel group randomised controlled trials studying humans and covered a cardiology topic (as defined above). We retrieved the full text of all potentially eligible articles. The same reviewers independently assessed eligibility of the full text articles with standardised forms and resolved discrepancies by discussion. An arbitrator (VMM) resolved any discrepancies that remained.

Data extraction

Seven reviewers (JWB, EAA, DMB, PA-C, AW, SU, and IF-G) trained in health research methods worked in pairs to extract data independently and in duplicate, using a standardised form and a data collection manual. Reviewers collected information on content area and the type of interventions tested, sample size, the length of follow-up, the number of composite end points presented, and the declared source(s) of funding.

To avoid confounding, we explored only data associated with each trial's primary composite end point. When authors reported more than one composite end point, we established the primary one by using the following hierarchy: (a) authors' explicit declaration of primacy, (b) the composite end point used to calculate the sample size, (c) authors' attribution of importance to the composite end point in their description of the results, and (d) the composite end point that appeared first in the methods section. Two reviewers (JWB and IF-G) independently selected the primary composite end point, resolving discrepancies by discussion.

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. When authors reported results from the same composite end point for more than one time point, we used data from the longest interval. A statistician (DH-A) entered all data into an electronic database and reviewed them for errors and missing data.

Ranking of outcomes according to importance to patients

Patients will typically assign varying importance to different health outcomes.[5] We sought, but were unable to find, a published hierarchical categorisation of importance to patients for cardiovascular outcomes. Therefore, to explore the variability in importance to patients across components, we developed a hierarchical categorisation of importance to patients for the component end points included in eligible studies. Two cardiologists (GP-M and IF-G) and nine internists (GHG, HJS, VMM, EAA, RJ, JA, VP-H, PA-C, and AD-S) independently categorised each of 72 outcomes used as components of composite end points in the eligible trials into five categories (I to V, in descending order of importance): I=death, II=critical, III=major, IV=moderate, and V=minor. Estimates of utility associated with the outcomes guided the process.[6] Group members met to resolve disagreements and succeeded in coming to consensus.

Data analysis

The κ statistic provided a measure of interobserver agreement independent of chance on the eligibility of randomised controlled trials. 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 ensure independence of observations

Table 1 | Characteristics of 114 included studies

Characteristic	No (%)
Journal	
Annals of Internal Medicine	2 (2)
Circulation	45 (39)
European Heart Journal	17 (15)
JAMA	21 (18)
Lancet	21 (18)
New England Journal of Medicine	8 (7)
Cardiovascular area	
Coronary disease	91 (80)
Heart failure	12 (11)
Vascular disease	9 (8)
Other*	2 (2)
Intervention	
Drug	76 (67)
Coronary intervention (surgery/percutaneous)	13 (11)
Vascular intervention (surgery/percutaneous)	2 (2)
Other†	23 (20)
No of composite end points per trial (n=111)	
1	65 (59)
2	28 (25)
3	12 (11)
≥4	6 (5)
Total No of components per composite end point	
2	39 (34)
3	44 (39)
4	14 (12)
>4	17 (15)
Effect on components reported	
Every component	77 (68)
Some components	21 (18)
No components	16 (14)

*Valvular heart disease, primary cardiovascular prevention, arrhythmia, or cardiomyopathy.
†Rehabilitation or lifestyle intervention (diet, daily activity).

within categories of importance to patients, we selected only one end point in each category for each composite end point to make these calculations; where a composite included more than one end point in the same category, we selected the end point with the highest event rate in the control group. To estimate the effect of the intervention across trials and within each category, we used random effects meta-analyses with an inverse variance approach. This method is conservative, in that it considers both within study and between study differences in estimating the pooled estimate. We used the I^2 statistic, the percentage of between study variability that is due to true differences between studies (heterogeneity) rather than sampling error (chance), to quantify inconsistency among trials.[7]

To describe the gradient of importance to patients among component end points, we considered a large gradient 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 limited analysis of the gradient of efficacy across components to those composite end points that provided data on at least two of their individual end points. We assigned a large gradient in the effect of the intervention if the 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. Our approach was to quantify, in sequence, the impact of adding component end points from importance to patients category III (major end points) and categories IV and V (moderate and minor end points) to end points allocated to categories I and II (fatal and critical end points). This was only possible for those studies that supplied data for each component of the composite in categories I, II, and III (fatal, critical, and major). For each study, we first calculated the event rate in the control group and the relative risk reduction on the basis of a composite of all the end points in categories I and II included in the original composite. We then repeated these calculations for another composite including all end points in categories I, II, and III. When adding the moderate and minor components (categories IV and V), we used the data for the original composite end point reported in the paper to calculate the control event rate and the relative risk reduction. Thus, we did not need component data for end points in categories IV and V.

Calculation of the exact impact would require joint distributions for all the components; because authors did not provide this level of detail, we made estimations by using a conservative approach to assess the impact of the outcomes of moderate and minor importance to patients. To establish the effect on the event rate for the control group, we estimated the impact of fatal and critical end points under the assumption that no patient had both a critical and a fatal event or more than one critical event. For instance, if the rate of death for the control group was 1% and that of large stroke was 2%, we calculated an event rate for the end points within importance to patients categories I and II of 3%. We then estimated the effect of adding the events associated with end points grouped in category III of importance to patients, again assuming mutually exclusive events, to the more serious events. Thus, if the rate of non-fatal myocardial infarctions was 2%, the event rate for the control group would increase from 3% to 5%. We considered the end points grouped in categories IV and V of importance to patients to account for the total composite end point event rate left unaccounted. Thus, if the composite end point event rate

for the control group was 10%, the effect of adding the less important outcomes (categories IV and V) would increase the control event rate from 5% to 10%.

A similar approach allowed assessment of the impact of outcomes grouped according to importance to patients on the effect of the intervention. We calculated the median and associated interquartile range for both the control group event rate and the effect of the intervention. We calculated a test of proportions (χ^2 test) to explore associations between gradients in either importance to patients or of the effect of treatment on components within composite end points by declared source of funding (industry versus non-industry funded). We used SAS version 9.1 and S-PLUS version 6.2 (Insightful Corporation, Seattle, Washington) for analyses; we chose a 5% threshold for statistical significance for all analyses.

RESULTS

Results of literature search

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 (fig 1).

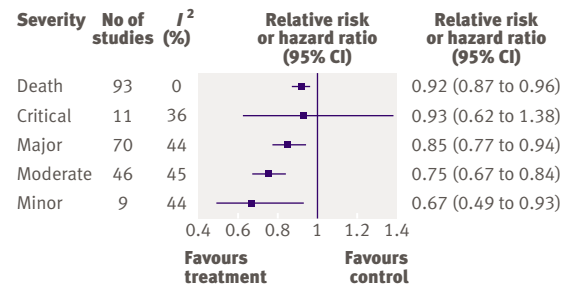| Severity	No of studies	I^2 (%)	Relative risk or hazard ratio (95% CI)	Relative risk or hazard ratio (95% CI)
Death	93	0		0.92 (0.87 to 0.96)
Critical	11	36		0.93 (0.62 to 1.38)
Major	70	44		0.85 (0.77 to 0.94)
Moderate	46	45		0.75 (0.67 to 0.84)
Minor	9	44		0.67 (0.49 to 0.93)

Fig 2 | Variability in magnitude of the effect of intervention across categories of importance to patients

Chance corrected agreement on eligibility was excellent ($\kappa=0.90$). Thus, approximately half of all parallel group cardiovascular randomised controlled trials identified reported an eligible composite end point.

Study characteristics

Table 1 describes the characteristics of the included studies. 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%) composite end points included fatal end points, usually reported as "all cause mortality" (table 2). 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). When a composite end point included only two components, reporting of individual event rates and the composite end point rate made the joint distribution apparent; no eligible trials with three or more components reported the joint distribution of component outcomes.

Most trials (69%; n=79) declared either direct financial industry funding (n=74) or industry having supplied the drug or device under investigation (n=5). Authors of 15 (13%) trials declared not for profit funding alone, and 20 (18%) did not declare a funding source. Of the 74 trials that declared industry funding, 27 trials also reported funding by not for profit sources.

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
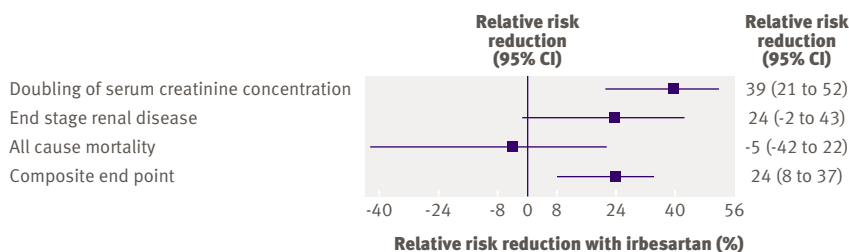
Table 2 | Component outcomes by category of importance to patients (114 composite end points)

End point	Prevalence—No (%)
I. Death	112 (98)
All cause mortality	43 (38)
Mortality not otherwise defined	32 (28)
Cardiac death not otherwise defined	13 (11)
Cardiac death due to coronary heart disease not otherwise defined	13 (11)
Other fatal end points (n=9)	24 (21)
II. Critical	18 (16)
Large myocardial infarction	14 (12)
Stroke leaving permanent moderate deficit	2 (2)
Cardiac arrest followed by resuscitation manoeuvres	2 (2)
Dissecting or ruptured aortic aneurysm	1 (1)
Other critical end points (n=3)	0 (0)
III. Major	95 (83)
Non-fatal myocardial infarction	83 (73)
Stroke leaving permanent deficit, severity not defined	16 (14)
Coronary artery bypass grafting	13 (11)
Cerebrovascular event not otherwise defined	13 (11)
Other major end points (n=13)	19 (17)
IV. Moderate	59 (52)
Coronary revascularisation not otherwise specified	22 (19)
Coronary revascularisation—angioplasty/stenting	12 (11)
Non-fatal angina needing hospital admission	11 (10)
Hospital admission not otherwise specified	5 (4)
Other moderate end points (n=17)	15 (13)
V. Minor	12 (11)
Non-fatal angina, not defined	7 (6)
Non-fatal arrhythmias, not otherwise specified	2 (2)
Dyspnoea, not otherwise defined	1 (1)
Change in blood pressure	1 (1)
Other minor end points (n=10)	2 (2)

Fig 3 | Comparison of irbesartan with amlodipine in the diabetic nephropathy study[8]

components (21%; n=18). Only 14% (n=12) of composite end points reviewed were composed of end points that exhibited either no gradient or a minor gradient in both importance to patients and the effect of treatment across components. Declared industry funding versus non-industry funding was not significantly associated with either the gradient in importance to patients or in the gradient of effect of treatment across components among composite end points.

Effect of end points of moderate and minor importance to patients

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 that pooled effects for fatal and critical outcomes were small, and end points of lesser importance to patients were associated with larger effects (fig 2).

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 3). The magnitude of

the treatment effect also rose substantially as less important components were included (table 3). 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.

Our results suggest that a naïve interpretation of composite end points may lead clinicians to overestimate the impact of treatments on preventing adverse events that matter most to patients. Consider, for example, the following statement: "In patients with in-stent stenosis of coronary artery bypass grafts, γ radiation reduced the composite end point of death from cardiac causes, Q wave myocardial infarction, and revascularisation of the target vessel." This result sounds impressive because it suggests that γ radiation reduces the incidence of death and myocardial infarction, as well as the need for revascularisation. The trial that produced this result randomised 120 patients with in-stent stenosis of a saphenous vein graft to γ radiation (iridium-192) or placebo.[8] Death or myocardial infarction contributed only 6/43 events in the placebo arm and 5/22 events in the iridium-192 arm. The investigators have shown the impact of the intervention on revascularisation. The trial provides, however, essentially no information about the effect of the intervention on myocardial infarction or death.

Consider another example—the irbesartan diabetic nephropathy trial that randomised 1715 hypertensive patients with nephropathy and type 2 diabetes to irbesartan, amlodipine, or placebo.[9] Results showed a benefit of irbesartan over amlodipine in the primary end point, a composite of a doubling of the baseline serum creatinine concentration, the onset of end stage renal disease (serum creatinine >6.0 mg/dl, initiation of dialysis, or transplantation), or death from any cause. Doubling of serum creatinine provided most events and was the only outcome for which irbesartan was convincingly beneficial. Indeed, in this instance, irbesartan both lowered the incidence of doubling of creatinine and showed a trend towards reduction in end stage renal disease, but it showed a slight trend towards

Table 3 | 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.

increased all cause mortality (fig 3). These examples highlight the challenges that clinicians face when making decisions on the basis of the results of cardiovascular trials that report composite end points.

DISCUSSION

Findings

Our analysis of a sample 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 one third of the articles. Most randomised controlled 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 review has potential limitations. Our conclusions depend on clinicians assigning importance to patients to cardiovascular end points, a challenging process. Previous analyses have considered component end points in one of two categories, fatal or non-fatal[1]; patients are, however, unlikely to attach similar importance to all non-fatal end points. Several authors have suggested weighting end points to reflect their relative importance when constructing composite end points,[10-14] and Lubsen and Kirwan have outlined a theoretical classification scheme for ranking components of composite end points.[15] Trialists, however, rarely use these strategies. Published data examining the utilities that patients attribute to a variety of cardiovascular outcomes guided our classification,[6] and 11 clinicians knowledgeable in cardiovascular care worked independently in generating the classification and were able to achieve consensus. Our decisions on categorisation are available (table 2 provides examples), and readers can independently judge their credibility.

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 analytic 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 end points of greater importance to patients.

One might question our application of meta-analytic approaches to data across a wide variety of interventions (fig 2). 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.[7] The increasing treatment effect with decreasing importance seems to be a real phenomenon.

Our work has additional strengths. Our sample of 114 composite end points is the result of a systematic search completed in duplicate with excellent agreement. Our data collection was comprehensive and careful, including independent judgment and abstraction of data at all stages by reviewers trained in the methodology and use of targeted, relevant analyses. We excluded composite end points containing efficacy and safety outcomes because their inclusion would have overestimated the proportion of composite end points with large gradients in the effect of the treatment. We included only one composite end point for each randomised controlled trial to avoid clustering, as multiple reported composite end points within a single trial commonly share components.

Implications

Use of composite end points appeals to clinical trialists because it increases event rates and statistical power. The fundamental problem with composite end points is, however, the difficulty in interpreting results when the gradient of importance to patients is substantial and a substantial gradient in the magnitude of the treatment effect also exists. Conversely, confident interpretation of composite end point results requires relatively small gradients of importance to patients and similar relative risk reductions across components.[3] 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. In particular, mortality outcomes, present in almost all cardiovascular composite end points, provide the lowest event rate and show the smallest treatment effects. Thus, 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.

The common use of inadequately reported composite end points with large gradients in importance to

patients, in which end points of least importance contribute most events, and in which treatment fundamentally affects these same components, is problematic. Trialists should report complete data on individual component end points to facilitate appropriate interpretation; clinicians should view with caution the results of cardiovascular trials that use composite end points to report their results. 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.

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, Permanyer-Miralda G, Busse JW, Bryant DM, Montori VM, Alonso-Coello P, et al. Rationale for using primary composite endpoints in clinical trials: a systematic review. *J Clin Epidemiol* (in press).

3   Montori VM, Permanyer-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   Cochrane Collaboration. Cochrane Handbook for Systematic Reviews of Interventions 4.2.6. Updated September 2006. www.cochrane.org/resources/handbook/Handbook4.2.6Sep2006.doc.

5   Guyatt G, Montori V, Devereaux PJ, Schunemann H, Bhandari M. Patients at the center: in our practice, and in our use of language. *ACP J Club* 2004;140:A11-2.

6   Tengs TO, Wallace A. One thousand health-related quality-of-life estimates. *Med Care* 2000;38:583-637.

7   Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557-60.

8   Waksman R, Ajani AE, White RL, Chan RC, Satler LF, Kent KM, et al. Intravascular gamma radiation for in-stent restenosis in saphenous-vein bypass grafts. *N Engl J Med* 2002;346:1194-9.

9   Lewis EJ, Hunsicker LG, Clarke WR, Berl T, Pohl MA, Lewis JB, et al. Renoprotective effect of the angiotensin-receptor antagonist irbesartan in patients with nephropathy due to type 2 diabetes. *N Engl J Med* 2001;345:851-60.

10 Neaton JD, Gray G, Zuckerman BD, Konstam MA. Key issues in end point selection for heart failure trials: composite end points. *J Card Fail* 2005;11:567-75.

11 Braunwald E, Cannon CP, McCabe CH. An approach to evaluating thrombolytic therapy in acute myocardial infarction: the 'unsatisfactory outcome' end point. *Circulation* 1992;86:683-7.

12 Moye LA. *Multiple analyses in clinical trials*. New York: Springer, 2003.

13 Hallstrom AP, Litwin PE, Weaver WD. A method of assigning scores to the components of a composite outcome: an example from the MITI trial. *Control Clin Trials* 1992;13:148-55.

14 DeMets DL, Califf RM. Lessons learned from recent cardiovascular clinical trials: part I. *Circulation* 2002;106:746-51.

15 Lubsen J, Kirwan BA. Combined endpoints: can we use them? *Stat Med* 2002;21:2959-70.