ORIGINAL RESEARCH

Endpoint Selection and Relative (Versus Absolute) Risk Reporting in Published Medication Trials

Michael Hochman, MD[1] *and Danny McCormick, MD, MPH*[2]

[1]The Robert Wood Johnson Clinical Scholars® Program, University of California, Los Angeles, and the U.S. Department of Veterans Affairs, Los Angeles, CA, USA; [2]Department of Medicine, Cambridge Health Alliance, and Harvard Medical School, Cambridge, MA, USA.

BACKGROUND: The use of surrogate and composite endpoints, disease-specific mortality as an endpoint, and relative (rather than absolute) risk reporting in clinical trials may produce results that are misleading or difficult to interpret.

OBJECTIVE: To describe the prevalence of these endpoints and of relative risk reporting in medication trials.

DESIGN AND MAIN MEASURES: We analyzed all randomized medication trials published in the six highest impact general medicine journals between June 1, 2008 and September 30, 2010 and determined the percentage using these endpoints and the percentage reporting results in the abstract exclusively in relative terms.

KEY RESULTS: We identified 316 medication trials, of which 116 (37%) used a surrogate primary endpoint and 106 (34%) used a composite primary endpoint. Among 118 trials in which the primary endpoint involved mortality, 32 (27%) used disease-specific mortality rather than all-cause mortality. Among 157 trials with positive results, 69 (44%) reported these results in the abstract exclusively in relative terms. Trials using surrogate endpoints and disease-specific mortality as an endpoint were more likely to be exclusively commercially funded (45% vs. 29%, difference 15% [95% CI 5%–26%], P=0.004, and 39% vs. 16%, difference 22% [95% CI 6%-37%], P=0.007, respectively). Trials using surrogate endpoints were more likely to report positive results (66% vs. 49%, difference 17% [95% CI 5%–28%], P=0.006) while those using mortality endpoints were less likely to be positive (46% vs. 62%, difference −16% [95% CI −27%–−4%], P=0.01).

CONCLUSIONS: The use of surrogate and composite endpoints, endpoints involving disease-specific mortality, and relative risk reporting is common. Articles should highlight the limitations of these endpoints and should report results in absolute terms.

The views expressed in this article are solely those of the author and do not necessarily represent those of the Department of Veterans Affairs or of the Robert Wood Johnson Foundation

Electronic supplementary material *The online version of this article (doi:10.1007/s11606-011-1813-7) contains supplementary material, which is available to authorized users.*

Received April 13, 2011
Revised June 8, 2011
Accepted July 1, 2011
Published online August 13, 2011

KEY WORDS: surrogate endpoints; composite endpoints; disease-specific mortality; relative risk reduction.
J Gen Intern Med 26(11):1246–52
DOI: 10.1007/s11606-011-1813-7

INTRODUCTION

Medical interventions are judged as successes or failures in clinical studies based on their impact on the study endpoints. Because interventions have a variety of effects, and may appear effective using one set of endpoints but harmful using another set, the selection of endpoints is critical.[1]

Recently, there have been growing concerns about the selection of endpoints and the reporting of results in clinical trials. In particular, experts have suggested that the use of surrogate endpoints,[2–6] and composite endpoints,[7–11] and the use of disease-specific—rather than all-cause—mortality as an endpoint,[12,13] may mislead readers if the limitations of these endpoints are not adequately explained. Similar concerns have also been raised about the reporting of trial results in relative—rather than absolute—numbers,[14,15] Table 1 describes these concerns.

A limited number of studies have examined the prevalence of the use of surrogate[3–5] and composite endpoints[7,9,16] the use of disease specific mortality as an endpoint[12,17] and the reporting of results in relative numbers[14,18,19] in clinical trials. However, data from these studies are several years old, typically reflect trials involving treatments for only a limited spectrum of diseases, and provide only limited details about factors that may be associated with these endpoints and with relative risk reporting.

In this study we examine the prevalence of the use of surrogate endpoints, composite endpoints, disease-specific mortality as an endpoint, and the use of relative risk reporting in recently published medication trials from the six highest impact general medicine and internal medicine journals. We also examine the association between trial funding source and outcome (i.e. positive vs. negative results) and these trial characteristics.

METHODS

Details of the process by which studies were identified have previously been published.[20] In summary, we identified all

Table 1. Definitions of and Concerns with Different Trial Endpoints and with Relative Risk Reporting

Characteristic	Definition	Concerns	Impact on a Study's Power
Surrogate endpoint	A laboratory, radiologic, or physician-assessed measurement that is believed to correlate with a clinically important outcome, such as mortality	Surrogate outcomes may not correlate with clinically important outcomes	Generally increased since surrogate endpoints like blood pressure are frequently more responsive to interventions than hard clinical endpoints like strokes
Composite endpoint	An endpoint consisting of two or more individual endpoints grouped together	Composite endpoints may amalgamate different endpoints of unequal clinical importance, making it difficult to interpret the effects of an intervention	Increased since multiple endpoints are amalgamated together, increasing event rates
Disease-specific mortality as an endpoint	The use of death due to a specific condition, rather than death due to any cause, as an endpoint	The effect of an intervention on disease-specific mortality may be counterbalanced by an opposite effects on other causes of mortality	Increased relative to all-cause mortality because components of all-cause mortality not impacted by the intervention dilute the measured impact
Relative risk reporting	The reporting of results in relative, rather than absolute, numbers	Relative risk reporting may distort the magnitude of the effect of an intervention, e.g. reporting that an intervention reduces mortality by 50% when the intervention reduces death rates from 0.0002% to 0.0001%	NA

randomized trials involving medications that reported on at least one pre-specified primary endpoint published in the 28-month period between June 1, 2008, and September 30, 2010, in the 6 general medicine and internal medicine journals with the highest impact factor (*New England Journal of Medicine*, *Lancet*, *JAMA*, *Annals of Internal Medicine*, *BMJ*, and *Archives of Internal Medicine*).[21] One author (M.H.) performed the literature review, abstracted information from included studies, and characterized each study with regard to outcomes of interest according to pre-specified criteria (outlined below and in the Appendix—available online). The studies were identified by manually reviewing all original articles with a formal abstract published during the specified time period.

All randomized trials involving humans in which systemically active medications (either a specific medication or a class of medications) were compared with either an active treatment (such as another medication or a non-pharmacologic therapy) or an inactive control (either a placebo or no therapy) were selected for inclusion. Phase 1 and phase 2 trials were excluded since these studies are preliminary assessments, and are not generally powered to assess clinically important endpoints. Trials that were not assigned to a phase were excluded if there were fewer than 100 subjects, since many of these trials would likely have been classified as phase 1 or 2 trials (only seven trials fit into this category and inclusion of these studies did not affect the overall study results).

DETERMINING STUDY CHARACTERISTICS

Each selected study was reviewed to determine the sample size, the trial phase (i.e. phase 3 or 4 or unassigned as reported either in the text of the study, the trial registry, or both), whether the study was commercially funded (i.e. received funding aside from free study medications from a for-profit company according to the Methods or Acknowledgements section), whether the study received funding from a government agency (either in the United States or another country according to the Methods or Acknowledgements section), whether the trial was an active comparator vs. inactive comparator trial, whether a non-inferiority data analysis (either alone or in addition to a superiority analysis) was performed, and whether the study medications were US Food and Drug Administration (FDA)–approved when trial enrollment began. Trials were also reviewed to determine if there was a statistically significant positive result (i.e. a favorable result such as better efficacy or a lower rate of adverse events) for the newer therapy relative to the control group with regard to any primary trial endpoint. Trials were considered ineligible for this latter analysis if there were three or more treatment groups because, in these cases, it was not possible to determine which two groups should be considered when determining if there was a positive result. In addition, trials comparing different medication administration strategies were considered ineligible since it was not clear which treatment arm represented the "newer" therapy.

In addition to the above characterizations, the primary endpoint(s) of each trial—as indicated in the article or, if not indicated in the article, as indicated in the trial registry—were reviewed to determine whether they were surrogate vs. clinical endpoints, whether they were composite endpoints, whether any primary endpoint involved disease-specific mortality or all-cause mortality, and whether the primary endpoint analyses were reported in the abstract exclusively using relative rather than absolute numbers (i.e. the trial was categorized as reporting results in relative terms only if all references to the result in the abstract were in relative numbers; relative vs. absolute reporting in the main article text was not assessed). The latter determination was only made for trials with positive results since relative risk reporting is most problematic when it causes an intervention to appear more effective than it actually is.

Further details of how surrogate and composite endpoints were classified are available in the Appendix (available online). Briefly, we defined surrogate endpoints as those using "a laboratory measurement or a physical sign used as a substitute for a clinically meaningful outcome that measures directly how a patient feels, functions or survives."[22,23] This judgment was made by the reviewing author. Prior

studies have also classified endpoints as surrogate vs. clinical based on the authors' judgment.[3–5] In the Appendix (available online), we list specific examples of endpoints that we classified as surrogate vs. clinical.

QUALITY CONTROL ASSESSMENT

Because of the potential subjectivity in determining whether trial endpoints should be classified as surrogate vs. clinical endpoints, a quality control review was performed. One author (D.M.) independently reviewed all studies published during the 2nd, 6th, 10th, 14th, 18th, 22nd, and 26th months of the 28-month study period and independently classified studies with regard to the use of surrogate vs. clinical endpoints. Unlike the primary data extractor (M.H.), D.M. was blinded to each study's funding source during the quality control review. Inter-rater agreement was assessed by calculating a κ statistic. Differences were not reconciled since the quality check was only performed on a subset of studies. For all analyses reported in this paper, M.H.'s determinations were used.

STATISTICAL ANALYSIS

We tested for associations between trial funding source and the use of each of the examined trial endpoints, as well as for associations between trial funding source and relative risk reporting. We also tested for associations between trial results (i.e. whether the trial had statistically significant positive results with respect to the newer therapy) and the use of each examined endpoint. Since we only examined relative risk reporting among trials with positive results, it was not possible to test for associations between the use of relative risk reporting and trial results. We used two-sided χ^2 tests with an a priori level of significance of $P \leq 0.05$ to test for all of the above associations. All of the above analyses were pre-specified.

We performed post-hoc calculations using two-tailed tests with an α level of 0.05 to assess the power of our study to detect statistically significant associations between the use of each examined trial characteristic and the reliance on exclusive commercial funding. We based these calculations on the actual prevalence of each endpoint, and we calculated the power to detect an absolute difference of 15% (e.g. if the prevalence of a particular endpoint was 30%, we determined the power to detect a difference in the use of that endpoint between commercially and non-commercially funded of 22.5% vs. 37.5%). The power to detect a significant difference for commercially vs. non-commercially funded trials for each characteristic was: 79% for surrogate endpoints; 81% for composite endpoints; 96.5% for multiple primary endpoints; 79% for any mortality endpoint; 41% for disease-specific mortality vs. all-cause mortality; and 45% for relative risk reporting.

We also performed post-hoc power calculations using two-tailed tests with an α level of 0.05 to assess the power of our study to detect statistically significant associations between a trial's outcome (i.e. positive vs. negative/neutral results) and the use of each examined endpoint. We based these calculations on the actual prevalence of positive results, 55%, and we calculated the power to detect an absolute difference of 15% (i.e. a rate of positive results of 47.5% among trials using a particular endpoint and a rate of positive results of 62.5% among trials not using that endpoint). The power to detect a significant difference was: 69% for surrogate endpoints; 68% for composite endpoints; 40% for multiple primary endpoints; 70% for any mortality endpoint; and 29% for disease-specific mortality vs. all-cause mortality.

All statistical analyses were performed using SAS version 9.2 (SAS Institute Inc, Cary, North Carolina) or STATA version 11.0 (StataCorp LP, College Station, Texas).

RESULTS

2,592 original articles were reviewed in the 6 medical journals, of which 316 reported on a pre-specified primary endpoint and were eligible for the analysis. Characteristics of these 316 trials are shown in Table 2.

Table 3 shows the percentage of trials using each type of endpoint as well as the percentage reporting results in the

Table 2. Trial Characteristics

Characteristic	Number of Trials N=316	
	N	%
Journal		
NEJM	139	44
Lancet	101	32
JAMA	28	9
Annals of Internal Medicine	22	7
British Medical Journal	16	5
Archives of Internal Medicine	10	3
Number of study subjects		
<150	28	9
150 – 499	87	28
>500	201	64
Randomized trial phase		
Phase 3	202	64
Phase 4	49	16
Trial phase unassigned	65	21
Commercial funding		
Exclusively commercial	152	48
Joint commercial and non-commercial	55	17
Exclusively non-commercial	109	34
Government funding		
U.S. government	64	20
Other government	56	18
No governmental funding	196	62
Type of comparator		
Active comparator	156	49
Inactive comparator	160	51
Non-inferiority trials		
Non-inferiority analysis	46	15
No non-inferiority analysis	270	85
FDA approval status of trial medications		
All FDA-approved medications	231	73
≥ 1 non-FDA approved medications	85	27
Trial results*		
Statistically significant positive result	157	55
Negative / neutral result	126	45

Applies to the 283 trials eligible for this analysis

Table 3. Endpoint Selection, Relative Risk Reporting, and Trial Funding

	Total (No, %)	Commercial Funding			Any Government Funding		
		Exclusive Commercial Funding No (%)	Some Non-Commercial Funding* No (%)	Difference % (95% CI)	Yes No (%)	No No (%)	Difference % (95% CI)
	N=316	N=152	N=164		N=120	N=196	
Surrogate Endpoints	116 (37)	68 (45)	48 (29)	15 (5–26)†	36 (30)	80 (41)	−11 (−21 – 0.1)
Composite Endpoints	106 (34)	57 (38)	49 (30)	8 (−3 - 18)	36 (30)	70 (36)	−6 (−16 – 5)
Any Mortality Endpoint	118 (37)	57 (38)	61 (37)	0.3 (−10 - 11)	46 (38)	72 (37)	2 (−9 – 13)
Disease specific mortality among all trials with mortality endpoints‡	32 (27)	22 (39)	10 (16)	22 (6–37)†	9 (20)	23 (32)	−12 (−27 – 4)
Exclusive relative risk reporting in abstract§	69 (44)	43 (43)	26 (46)	3 (−13 – 18)	21 (47)	48 (43)	4 (−13 – 21)

* *Non-commercial funding sources include government agencies and non-profit organizations such as foundations*
† *Statistically significant (P<0.05)*
‡ *Applies to the 118 trials with a mortality endpoint, 57 of which received exclusive commercial funding, 61 of which received non-commercial funding (i.e. at least some funding from a non-commercial entity), 46 of which received government funding, and 72 of which did not receive government funding*
§*Applies to the 157 trials reporting statistically significant positive results, 100 of which received exclusive commercial funding, 57 of which received non-commercial funding (i.e. at least some funding from a non-commercial entity), 45 of which received government funding, and 112 of which did not receive government funding*

abstract exclusively in relative terms. The percentages are stratified by trial funding source. Trials using surrogate endpoints and disease-specific mortality as an endpoint were more likely to be exclusively commercially funded than trials receiving some non-commercial funding (45% vs. 29%, difference 15% [95% CI 5%-26%] P=0.004, and 39% vs. 16%, difference 22%, [95% CI 6%-37%], P=0.007, respectively).

Table 4 shows the relationship between each of the examined endpoints and the trial results (positive vs. negative with respect to the newer therapy). Trials using surrogate endpoints were more likely to report positive results (66% vs. 49%, difference 17% [95% CI 5%–28%], P=0.006) while those using mortality endpoints were less likely to be positive (46% vs. 62%, difference −16% [95% CI −27% – −4%], P=0.01). In addition, trials using composite endpoints were less likely to generate positive results (47% vs. 60%, difference −12% [95% CI −24% – 0.0%], P=0.05), though this difference was of borderline statistical significance.

Among the 106 trials using a composite endpoint, 86 (81%) included a mortality endpoint (either disease-specific or all-cause) as part of the composite, while 30 (28%) included a surrogate endpoint as part of the composite. Among the 118 trials using a mortality endpoint, in 86 cases (73%) the mortality endpoint was part of a composite.

A total of 67 of the 316 trials were analyzed for the quality control review, and there was agreement between the primary data analysis and the quality review for 59 trials with regard to classification of the primary trial endpoint as surrogate vs. clinical (88%; 95% CI, 80%-96% [κ 0.76; 95% CI, 0.61–0.92]).

Table 4. Endpoint Selection and Trial Results*

Characteristic	Positive Results No (%)	Difference % (95% CI)	P-value
Surrogate Endpoints			
Yes	70 (66)	17 (5 – 28)	0.006
No	87 (49)		
Composite Endpoints			
Yes	46 (47)	−12 (−24 – 0.0)	0.05
No	111 (60)		
Any Mortality Endpoint			
Yes	51 (46)	−16 (−27 – −4)	0.01
No	106 (62)		
Type of Mortality Endpoint†			
Disease-specific	12 (39)	−10 (−28 – 10)	0.34
All-Cause	39 (49)		

**Applies to the 283 trials eligible for this analysis. Of these 283 trials, 106 used surrogate endpoints, 97 used composite endpoints, and 111 used any mortality endpoint*
† *Applies to the 111 trials with two treatment arms and any mortality endpoint. Of these trials, 31 used disease specific mortality and 80 used all-cause mortality*

DISCUSSION

Among 316 randomized trials published in top medical journals, 37% used surrogate endpoints and 34% used composite endpoints. Among trials using mortality as an endpoint, 27% used disease-specific mortality rather than all-cause mortality. In addition, 44% of trials reported results in the abstract exclusively in relative numbers. Our study represents the largest and most comprehensive analysis of trial endpoints and relative risk reporting that we are aware of. The results of several smaller analyses are summarized in Table 5.

We believe the high prevalence of these endpoints is notable because it demonstrates the complexity of interpreting medical research for patients, doctors, and policy makers. Without appropriate explanation of the limitations of these endpoints, readers may draw inappropriate conclusions. Determination of whether articles adequately explained the limitations was beyond the scope of our analysis.

Despite the concerns with these endpoints, however, we do not intend to suggest that these endpoints should never be used. An important advantage of these endpoints is that they frequently increase the power of a study to detect significant

Table 5. Previous Studies Examining the Prevalence of Different Endpoints and of Relative Risk Reporting in Clinical Trials

Studies	Findings
Surrogate Endpoints	
Ridker et al.[4]	An analysis of cardiovascular trials published in three top journals between 2000–2005 showed that 35% used surrogate endpoints. Commercially funded trials were less likely than non-commercially funded cardiovascular trials to use surrogate endpoints, which seems to contradict the findings of our analysis. Consistent with our results, however, trials using surrogate endpoints were more likely to generate positive results than those using clinical endpoints
Bero et al.[3]	An analysis of statin trials identified in Pubmed between 1999–2005 found that 98% used surrogate endpoints
la Cour et al.[5]	An analysis of 626 randomized clinical trials published in 2005 and 2006 in six top general medicine journals found that 17% used a surrogate primary endpoint, though this analysis was not a true prevalence estimate since the estimate was obtained using a Pubmed search rather than by hand-searching and non-randomized studies were automatically counted as using clinical endpoints. In addition, the definition of a surrogate endpoint was somewhat different than ours: studies using cost effectiveness and other measures of resource utilization as a primary endpoint, as well as those with a composite endpoint in which any of the components was a clinical endpoint, were classified as using clinical endpoints
Composite Endpoints	
Lim et al.[16]	A study of cardiovascular trials published in 14 top journals between 2000–2007 showed that 37% used composite endpoints. Trials using composite endpoints were more likely to report P-values of <0.05 than they were to report non-significant P-values, but whether the significant P-value referred to a positive vs. negative result with respect to the newer therapy was not specified. In addition, the study did not report whether trials using composite endpoints were more or less likely than trials not using composite endpoints to report significant results. Therefore, it is not clear that these results contradict our finding that trials using composite endpoints are less likely than trials not using composite endpoints to show positive results
Disease-Specific Mortality	
Kip et al.[17]	An analysis of 20 trials comparing bare metal vs. drug-eluting stents showed that 55% used cardiovascular mortality as an endpoint while 45% used all-cause mortality; another analysis reported in the same paper showed that, among 27 trials published in 2006 in the Journal of the American College of Cardiology, 19% used cardiovascular death as an endpoint while 74% used all-cause mortality
Relative Risk Reporting	
Nuovo et al.[14]	Fewer than 10% of trials published in top medical journals prior to 1998 reported risk reductions in absolute terms anywhere in the article
Dryver et al.[18]	Among articles published in eight journals in 1996, 12% of the abstracts provided a measure of absolute risk difference; no abstracts published in these journals in 1986 did
Schwartz et al.[19]	Among articles published in 2003 and 2004 in six top journals, 68% of failed to report the underlying absolute risks for the first ratio measure in the abstract

associations (the reasons for this are described in Table 1). For this reason, these endpoints are generally appropriate in early phase studies in which investigators seek preliminary data to rapidly assess safety and efficacy. These endpoints may also be appropriate in some advanced phase trials of therapies for rare diseases or in low risk populations in which it may be too expensive or impractical to recruit sufficient patients to achieve adequate statistical power. In our study, we did not attempt to assess whether or not articles that used these endpoints did so appropriately because of the difficulty of making this determination. However, by only including advanced phase studies, we excluded many studies in which the use of these endpoints is likely to be appropriate.[6]

Another notable finding of our study is that authors commonly report trial results exclusively in relative terms despite a growing recognition that relative risk reporting can distort the magnitude of the effects of an intervention.[13,14,18,19] Given these concerns, readers are likely to be better served by uniform inclusion of absolute risk reporting.

We also found that trials receiving exclusive commercial funding were more likely than those receiving non-commercial funding to use surrogate endpoints and to use disease-specific mortality as an endpoint. We speculate that commercially funded trials—compared with non-commercially funded trials—tend to select endpoints such as surrogate markers and disease-specific mortality that are perceived to have a higher likelihood of generating statistically significant results. As noted above, however, it is possible that some of the commercially funded trials in our analysis that used these endpoints did so appropriately.

Our analysis also showed that trials using surrogate endpoints were more likely than trials using clinical endpoints to generate positive results. This is notable because several recent studies have shown that interventions that have apparently beneficial effects on surrogate markers of disease often do not also have beneficial effects on clinical endpoints[2]. For this reason, positive trials that use surrogate endpoints may inappropriately promote the use of medications that do not have important clinical benefits.[24] These finding support recent attempts by the FDA to avoid granting drug approvals based on the results of trials using surrogate endpoints.[25] In addition, our finding that surrogate endpoints are more common among commercially funded trials may help explain why commercially funded research is more likely to generate positive results than non-commercially funded research[4,26–28]

We also found that trials using mortality endpoints were less likely than those not using mortality endpoints to generative positive results, while those using composite endpoints were less likely to generate positive results. The former finding was expected since developing a medication that reduces mortality is presumably more difficult than developing a medication that has beneficial effects on a surrogate endpoint or a more subjective clinical endpoint such as symptom control. However, the latter was unexpected since composite endpoints are frequently used to increase the power of a study to detect significant findings.[7,10] One possible explanation is that a disproportionately high percentage—81%—of trials using a composite endpoint included a mortality endpoint as part of the composite. Since mortality endpoints were associated with negative/neutral results, the low rate of positive findings among trials using composite endpoints may result from the fact that mortality endpoints were used disproportionately. Alternatively,

and especially since this finding was of borderline significance, the association may represent a chance occurrence.

We did not detect other associations between trial funding and endpoint selection or relative risk reporting or between endpoint selection and trial outcomes, however our power to detect some of these associations was limited.

Finally, we found that 81% of trials using a composite endpoint included mortality as part of the composite while 28% included a surrogate marker. These findings suggest that composite endpoints commonly consist of components of varying clinical importance, which may mislead readers.[7,10,29] For example, if an intervention reduces the composite rate of death, an endpoint with high clinical importance, and coronary revascularization, an endpoint with lower importance, it may not be clear whether the positive result is driven by a reduction in death, revascularization, or both.

There are several limitations to our study. First, our study only focused on medication trials. Second, we only included studies published in the most prominent general medicine and internal medicine journals. Third, our analysis only examined primary trial endpoints. Fourth, we only examined relative risk reporting in trial abstracts. Fifth, the author who extracted the data was not blinded to each study's funding source. Finally, the determination of surrogate vs. clinical endpoints, although made using predefined objective criteria, was only made by one author. In this regard, we are reassured by the high concordance between the clinical vs. surrogate assessments made by the abstracting author and the author whose quality control check was blinded.

Overall, our study shows that the use of surrogate and composite endpoints and endpoints involving disease-specific mortality is common. In addition, articles frequently report results in relative numbers. These findings highlight the need for educational efforts to ensure that readers understand the complexities of these endpoints and of relative risk reporting. A number of recent reports describing the appropriate interpretation of surrogate and composite endpoints,[22,30,31] and of relative risk reporting[32] could serve as guides for these efforts. In addition, Institutional Scientific Review Committees and regulatory agencies (e.g. the FDA) must closely examine the endpoints used in clinical trials and discourage the inappropriate use of surrogate and composite endpoints, and endpoints involving disease-specific mortality. Finally, medical journals may consider instituting editorial policies mandating the reporting of results in absolute numbers.

Contributors: *We have listed everyone who made substantial contributions to this analysis as authors of the article.*

Funders: *This study was unfunded (i.e. it received no internal or external funding).*

Prior Presentations: *This work was presented as a poster at the 2011 Society of General Internal Medicine National Meeting in Phoenix, Arizona.*

Conflict of Interest: *None disclosed.*

Corresponding Author: *Michael Hochman, MD; The Robert Wood Johnson Clinical Scholars® Program, University of California, Los Angeles, and the U.S. Department of Veterans Affairs, 911 Broxton Ave., 3rd Floor, Los Angeles, CA 90024, USA (e-mail: Meh1979@gmail.com).*

REFERENCES

1. **Kraemer HC, Frank E.** Evaluation of comparative treatment trials: assessing clinical benefits and risks for patients, rather than statistical effects on measures. JAMA. 2010;304(6):683–4.
2. **Krumholz HM, Lee TH.** Redefining quality–implications of recent clinical trials. N Engl J Med. 2008;358(24):2537–9.
3. **Bero L, Oostvogel F, Bacchetti P, Lee K.** Factors associated with findings of published trials of drug-drug comparisons: why some statins appear more efficacious than others. PLoS Med. 2007;4(6):e184.
4. **Ridker PM, Torres J.** Reported outcomes in major cardiovascular clinical trials funded by for-profit and not-for-profit organizations: 2000–2005. JAMA. 2006;295(19):2270–4.
5. **la Cour JL, Brok J, Gøtzsche PC.** Inconsistent reporting of surrogate outcomes in randomised clinical trials: cohort study. BMJ. 2010;341:c3653.
6. **Fleming TR, DeMets DL.** Surrogate end points in clinical trials: are we being misled? Ann Intern Med. 1996;125(7):605–13.
7. **Freemantle N, Calvert M, Wood J, Eastaugh J, Griffin C.** Composite outcomes in randomized trials: greater precision but with greater uncertainty? JAMA. 2003;289(19):2554–9.
8. **Yeh RW, Go AS.** Rethinking the epidemiology of acute myocardial infarction: challenges and opportunities. Arch Intern Med. 2010;170(9):759–64.
9. **Ferreira-González I, Busse JW, Heels-Ansdell D, Montori VM, Akl EA, Bryant DM, Alonso-Coello P, Alonso J, Worster A, Upadhye S, Jaeschke R, Schünemann HJ, Permanyer-Miralda G, Pacheco-Huergo V, Domingo-Salvany A, Wu P, Mills EJ, Guyatt GH.** Problems with use of composite end points in cardiovascular trials: systematic review of randomised controlled trials. BMJ. 2007;334(7597):786.
10. **Freemantle N, Calvert MJ.** Interpreting composite outcomes in trials. BMJ. 2010;341:c3529.
11. **Cordoba G, Schwartz L, Woloshin S, Bae H, Gøtzsche PC.** Definition, reporting, and interpretation of composite outcomes in clinical trials: systematic review. BMJ. 2010;341:c3920.
12. **Black WC, Haggstrom DA, Welch HG.** All-cause mortality in randomized trials of cancer screening. J Natl Cancer Inst. 2002;94(3):167–73.
13. **Laupacis A, Sackett DL, Roberts RS.** An assessment of clinically useful measures of the consequences of treatment. N Engl J Med. 1988;318:1728–33.
14. **Nuovo J, Melnikow J, Chang D.** Reporting number needed to treat and absolute risk reduction in randomized controlled trials. JAMA. 2002;287(21):2813–4.
15. **Marshall KG.** Prevent: how much harm? How much benefit? 1. Influence of reporting methods on perception of benefits. CMAJ. 1996;154(10):1493–9.
16. **Lim E, Brown A, Helmy A, Mussa S, Altman DG.** Composite outcomes in cardiovascular research: a survey of randomized trials. Ann Intern Med. 2008;149(9):612–7.
17. **Kip KE, Hollabaugh K, Marroquin OC, Williams DO.** The problem with composite end points in cardiovascular studies: the story of major cardiac events and percutaneous coronary intervention. J Am Coll Cardiol. 2008;51(7):701–7.
18. **Dryver E, Hux JE.** Reporting of numerical and statistical differences in abstracts: improving but not optimal. J Gen Intern Med. 2002;17(3):203–6.
19. **Schwartz LM, Woloshin S, Dvorin EL, Welch HG.** Ratio measures in leading medical journals: structured review of accessibility of underlying absolute risks. BMJ. 2006;333(7581):1248.
20. **Hochman M, McCormick D.** Characteristics of published comparative effectiveness studies of medications. JAMA. 2010;303(10):951–8.
21. Science journal citation reports ISI: Summer 2008. Thomson Reuters Web site. http://scientific.thomson.com/products/jcr/. Accessed July 5, 2011.
22. **Bucher HC, Guyatt GH, Cook DJ, Holbrook A, McAlister FA.** How to use an article measuring the effect of an intervention on surrogate end points. JAMA. 1999;282(8):771–8.
23. **Temple RJ.** A regulatory authority's opinion about surrogate endpoints. In: **Nimmo WS, Tucker GT, eds.** Clinical Measurement in Drug Evaluation. New York: Wiley; 1995.
24. **Grady D, Redberg RF.** Less is more: how less health care can result in better health. Arch Intern Med. 2010;170(9):749–50.
25. **Rosen C.** Revisiting the rosiglitazone story – lessons learned. NEJM. 2010;363(9):803–5.

26. **Als-Nielsen B, Chen W, Gluud C, Kjaergard LL.** Association of funding and conclusions in randomized drug trials: a reflection of treatment effect or adverse events? JAMA. 2003;290(7):921–8.
27. **Kjaergard LL, Als-Nielsen B.** Association between competing interests and authors' conclusions: epidemiological study of randomised clinical studies published in the BMJ. BMJ. 2002;325(7358):249–52.
28. **Bekelman JE, Li Y, Gross CP.** Scope and impact of financial conflicts of interest in biomedical research: a systematic review. JAMA. 2003;289 (4):454–65.
29. **Tomlinson G, Detsky AS.** Composite endpoints in randomized trials: there is no free lunch. JAMA. 2010;303(3):267–8.
30. **Guyatt G, Rennie D, Meade MO, Cook DJ.** Can one be confident that the component endpoints share similar relative risk reductions? Users' guides to the medical literature. Chapter 10.4. Composite Endpoints.
31. **Montori VM, Permanyer-Miralda G, Ferreira-González I, Busse JW, Pacheco-Huergo V, Bryant D, Alonso J, Akl EA, Domingo-Salvany A, Mills E, Wu P, Schünemann HJ, Jaeschke R, Guyatt GH.** Validity of composite endpoints in clinical trials. BMJ. 2005;330(7491):594–6.
32. **Guyatt GH, Sackett DL, Cook DJ.** Users' guides to the medical literature. II. How to use an article about therapy or prevention. A. Are the results of the study valid? JAMA. 1993;270(21):2598–601.