

End points in clinical trials: are they moving the goalposts?

D Y Leung, J K French

In selecting and defining composite end points in clinical trials, are we trading off clinical significance for statistical significance?

Heart 2006;92:870–872. doi: 10.1136/hrt.2005.082024

Integral to the assessment of the efficacy and safety of an experimental treatment or strategy in clinical trials is the nature and the definition of the end points. Over the last two decades clinical trials in treatment of coronary heart disease have evolved in their end point definitions. In the landmark ISIS-2 trial,¹ which enrolled over 17 000 patients with suspected acute myocardial infarction (MI) presenting within 24 h of symptom onset, total mortality was the outcome which determined the superiority of aspirin and/or streptokinase over control treatments. Such trials require tens of thousands of patients to have sufficient power to determine the effects of investigational treatment(s). While mortality is the most clinically relevant and important end point, in other clinical scenarios such as the performance of percutaneous coronary intervention (PCI) this is fortunately a relatively rare event.

COMPOSITE END POINTS

Composite end points are often used in such clinical trials. As these composite end points usually consist of events which are to a greater or lesser degree associated with increased mortality, they are thus seen as “surrogates” for mortality. As surrogate end points occur more frequently than death, it is possible to determine potential treatment benefits in a trial while recruiting a smaller number of patients. For example, recurrent MI is frequently used as a surrogate end point and a component of a composite end point because of its association with increased mortality.² The limitation of this approach is that the association between the surrogate and the “hard clinical end point” such as mortality may not be particularly strong. Furthermore, composite end points often include both efficacy and safety factors such as recurrent MI, further revascularisation procedures, stroke, and bleeding. The direction of the treatment effects with respect to efficacy and safety, however, may not be the same. A particular pharmacological and/or mechanical therapeutic strategy, for example, may have divergent effects on any individual factor in a composite end point.

When the end points in clinical trials are defined, various assessments of consequent risks are often made.³ Risk is defined as the probability or likelihood of encountering an event in a particular group of persons. While

absolute risk is easily understood as the likelihood of an event occurring, the concept of relative risk is often used in clinical trials where a treatment is compared with placebo or when two treatments are compared. Relative risk is the ratio of the absolute risks of events occurring in the two groups and the relative risk reduction (or excess) is the difference between the relative risk and unity. However, neither the absolute nor the relative risk measures the proportion of an event, usually death, attributable to the exposure to a particular risk factor. Quantification of this attributable risk is of importance as it allows interpretation of the trial data and the assessment of the potential impact of the trial findings on clinical practice, facilitating targetting of intervention strategies.

It is apparent that the magnitude of these risks is determined as much by the intrinsic (real) risks as it is by the definition of what constitutes an event. Death is a dichotomous variable with no ambiguity in its definition. Even death may be subdivided into cardiac and non-cardiac causes, a distinction that may be problematic. Also not so straightforward is use of combined end points which consist of events defined as a dichotomous variable based on continuous (parametric) measurements. A prime example in clinical trials in the setting of PCI is the definition of periprocedural MI.

EFFICACY AND SAFETY

Efficacy and safety considerations are the main focus in clinical trials in the setting of PCI with novel device and/or adjuvant pharmacological therapies. Elevation of creatine kinase (CK)-MB values occur in up to 30% of patients undergoing percutaneous intervention; elevations of troponin I or T values are more frequent whereas elevations of the total enzymatic concentration of CK are less frequent.² An increase in the risk of adverse outcomes correlates with increasing values of periprocedural elevations of CK-MB.⁴ Furthermore, an increase in the relative risk associated with increasing values of periprocedural CK-MB elevations has been shown to be similar to the risks associated with spontaneous events.⁵ Bleeding complications are considered important as they potentially lead to morbidity and even mortality, and may compromise an otherwise angiographically successful intervention. Therefore, composite end points commonly adopted in clinical trials in the setting of urgent and/or elective PCI such as REPLACE-2 consist

See end of article for authors' affiliations

Correspondence to:
Professor John K French,
Department of Cardiology,
Liverpool Hospital,
Elizabeth Street, Liverpool,
NSW 2170, Australia;
j.french@unsw.edu.au

Abbreviations: CABG, coronary artery bypass graft surgery; CK, creatine kinase; MI, myocardial infarction; PCI, percutaneous coronary intervention

of total mortality, periprocedural MI, and repeat revascularisation.⁶

MYOCARDIAL INFARCTION AS AN END POINT

Two questions arise: (1) How much elevation above the upper limit of the reference range of a cardiac marker of myocardial necrosis such as CK-MB (and troponin I or T) constitutes an infarct? (2) How valid is this composite end point a surrogate for late mortality?

To address these questions, in this issue of *Heart* Chew and colleagues examine the data from REPLACE-2.⁷ They evaluated the relationship between the risks of 12-month mortality attributable to the two components of the composite end point—those of peri-procedural MI and bleeding. They also measured the risks attributable to each of these components of the end points towards 12-month mortality, when these end points were defined using different cut-off values. Logistic regression models were used to allow for the confounding effects of baseline clinical parameters that are known to have impact on mortality. Limiting the definition of MI to events > 48 h after the coronary intervention excluded periprocedural MI and led to a relatively high relative risk of 12 month mortality (odds ratio 13.5). As MI thus defined was infrequent, the attributable risk was only 5.4%. The development of Q waves is normally fairly specific for MI though not very sensitive for small amounts of myocardial necrosis. With increasingly sensitive threshold definitions of MI, the relative risk of 12 month mortality decreased while the attributable risks increased. The cut-off values of CK or CK-MB value of $\geq 3 \times$ upper limits of the reference range appeared to be associated with the most optimal compromise between relative risks (odds ratio 3.5) and attributable risks (13.2%). Troponin-based definitions of periprocedural MI are in evolution. As periprocedural troponin elevations are somewhat more frequent (and more sensitive) than CK-MB elevations, attributable risks of troponin elevations are likely to be higher and the relative risks are likely to be lower.

A similar example where an end point is defined as a dichotomous variable based on continuous measurements with arbitrary cut-offs is seen with elevation of cardiac markers in the setting of coronary artery bypass graft operations (CABG). Similar to the situation with MI in the setting of PCI, elevated blood values of cardiac markers/enzymes after CABG are associated with increased late mortality or cardiac events.⁸ Furthermore, increasing concentrations of CK-MB correlated with six months mortality after CABG in a roughly linear manner.⁹ While a CK-MB elevation five times the upper limit of the reference range appears to be the most appropriate level for the definition of MI after CABG, a lower level of elevation of CK-MB is not entirely benign.⁹

BLEEDING COMPLICATIONS

In the article by Chew and colleagues,⁷ examination of different definitions of bleeding complications also yielded similar results. More stringent criteria of bleeding complications like TIMI major bleeding were associated with a high relative risk of 12 month mortality (odds ratio 6.1) but a low attributable risk of 3.5%. More liberal definitions of bleeding likewise led to lower relative risks and higher attributable risks. Combining the optimal cut-offs of the different components of the composite end points led to an overall attributable risk of 25.8% before adjustment for baseline clinical variables and of 27% after adjustment. In this analysis by Chew and colleagues, 514 patients (9.1%) and 12 events (1.3%) were included on the basis of total CK data alone. This less sensitive MI marker may have had a minor impact on the results. Furthermore, no data were available on 315 patients (5.2%), a relatively high number for only a

12-month follow up period in a randomised clinical trial with a low overall mortality rate. Survivors and non-survivors at 12 months differed significantly in baseline clinical characteristics, many of which were also defined by arbitrary cut-off values. For example, the odd ratios and attributable risks of many of the clinical risk factors (renal impairment, heart failure, diabetes and anaemia) for late mortality were as high as, if not higher than, periprocedural MI or bleeding. It will be an interesting exercise to assess the impact of different definitions of these parameters on the calculated relative risks and attributable risks for late mortality.

CHANGING THE DEFINITIONS

Thus, it can be seen that the goalposts can be shifted by simply changing the definitions. The importance of these definitions of “softer” end points in clinical trials in the setting of PCI and how they impact on the overall results is illustrated by periprocedural MI in the article by Chew and colleagues.⁷ The basis of the MI may result in different weightings in terms of attributable risks and relative risks. In the RITA-3 study where an early invasive strategy was compared with a conservative management strategy after non-ST elevation MI,¹⁰ redefining MI as any elevation of cardiac markers above the upper limit of reference range (new American College of Cardiology/European Society of Cardiology definition of MI) would result in the number of MIs in the invasive versus the conservative arm increasing from 45 v 56 ($p = \text{NS}$) to 84 v 129 ($p = 0.002$).²

The trade-off between relative and attributable risks is very similar to that between sensitivity and specificity. Changing the cut-off values would usually result in opposite changes in sensitivity and specificity without changing the overall accuracy of the tests or measurements in question: we are just operating on different parts of the receiver operative characteristics curve. While it serves useful purposes in the context of defining end points in clinical trials and achieving statistical significance, changing the goalposts in the definition of peri-procedural MI or bleeding does not necessarily make such events more or less useful in the overall risk assessment of a patient.

Do the findings of the Chew article resolve the dilemma of the goal post (or definition) of peri-procedural MI? While we have learned that defining periprocedural MI by a CK-MB value of ≥ 3 times upper limit of the reference range is associated with an acceptable compromise between relative and attributable risks of late mortality, such definitions would still account for only about 13% of the 12-month mortality. Combining periprocedural MI, urgent revascularisation and protocol defined minor bleeding can only account for a quarter of the 12-month mortality. A major assumption of attributable risk analysis is that the late mortality is causally linked to exposure of the particular risk factor in question. This is of particular interest in assessing the potential costs and benefits associated with, and its impact on avoidance of, the late event. We caution that this assumption has limitations in the setting of MI or bleeding with PCI.

Ultimately, in selecting and defining composite end points in clinical trials, we have to ask ourselves whether we are sacrificing clinical significance for statistical significance. Should we consider cardiac markers as continuous variables with their full implications rather than as discrete all-or-none variables in data analysis?

Authors' affiliations

D Y Leung, J K French, Cardiology Department and South West Sydney Clinical School (UNSW), Liverpool Hospital, Elizabeth Street, Liverpool, NSW 2170, Australia

REFERENCES

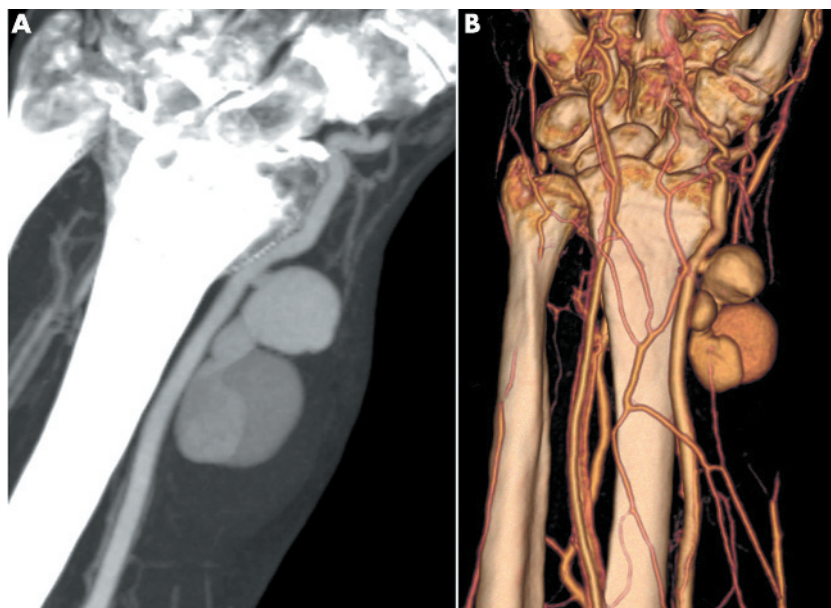
- 1 ISIS-2 (Second International Study of Infarct Survival) Collaborative Group. Randomized trial of intravenous streptokinase, oral aspirin, both, or neither among 17,187 cases of suspected acute myocardial infarction: ISIS-2. *Lancet* 1988;ii:349-60.
- 2 French JK, White HD. Clinical implications of the new definition of myocardial infarction. *Heart* 2004;90:99-106.
- 3 Sedgwick JEC. Absolute, attributable, and relative risk in the management of coronary heart disease. *Heart* 2001;85:491-2.
- 4 Roe MT, Mahaffey KW, Kilaru R, et al. Creatine kinase-MB elevation after percutaneous coronary intervention predicts adverse outcomes in patients with acute coronary syndromes. *Eur Heart J* 2004;25:313-21.
- 5 Akkerhuis KM, Alexander JH, Tardiff BE, et al. Are spontaneous and percutaneous coronary intervention-related events different? *Circulation* 2002;105:554-6.
- 6 Lincoff AM, Kleiman NS, Kereiakes DJ, et al. Long-term efficacy of bivalirudin and provisional glycoprotein IIb/IIIa blockade vs heparin and planned glycoprotein IIb/IIIa blockade during percutaneous coronary revascularization: REPLACE-2 randomized trial. *JAMA* 2004;292:696-703.
- 7 Chew DP, Bhatt DL, Lincoff AM, et al. Clinical end point definitions after percutaneous coronary intervention and their relationship to late mortality: an assessment by attributable risk. *Heart* 2006;92:945-50.
- 8 Costa MA, Carere RG, Lichtenstein SV, et al. Incidence, predictors, and significance of abnormal cardiac enzyme rise in patients treated with bypass surgery in the arterial revascularization therapies study (ARTS). *Circulation* 2001;104:2689-93.
- 9 Klatte K, Chaitman BR, Theroux P, et al. Increased mortality after coronary artery bypass graft surgery is associated with increased levels of postoperative creatine kinase-myocardial band isoenzyme release: results from the GUARDIAN trial. *J Am Coll Cardiol* 2001;38:1070-7.
- 10 Fox KA, Poole-Wilson PA, Henderson RA, et al. Interventional versus conservative treatment for patients with unstable angina or non-ST-elevation myocardial infarction: the British Heart Foundation RITA 3 randomised trial. Randomized intervention trial of unstable angina. *Lancet* 2002;360:743-51.

IMAGES IN CARDIOLOGY

doi: 10.1136/hrt.2005.074658

Radial artery pseudoaneurysm post-cardiac catheterisation: imaging with multidetector row CT peripheral angiography

A 60 year old woman underwent cardiac catheterisation via a right radial artery puncture. Over the subsequent three days she developed a pulsatile swelling over the right wrist with no circulatory or neurological compromise to the hand. An initial duplex ultrasound examination was performed which demonstrated an approximately 3 cm complex, loculated and thrombus-filled pseudoaneurysm with a small distal radial artery communication; prolonged compression of the pseudoaneurysm, with the ultrasound probe under direct vision, failed to obliterate inflow. Multidetector row computed tomographic peripheral angiography was performed (16-slice Lightspeed scanner, GE Healthcare Technologies, Waukesha, Wisconsin, USA) for surgical planning, with contrast bolus-tracking (iopromide) at the brachial artery and scanning from elbow to hand using 0.625 mm collimation. This confirmed the presence, anatomy and location of a 2.5 cm × 4.5 cm radial artery pseudoaneurysm with 4 mm diameter communication. Ulnar artery and palmar arch integrity were also confirmed. Thick "maximum intensity projection" (left panel), three dimensional volume rendered (right panel) anteroposterior view images and a three dimensional model were produced, which facilitated surgical planning for this patient. Successful treatment was achieved through tying off the radial artery.



N E Manghat
F M Ellwood
C A Roobottom
docnatman@msn.com