

A collection of misleading surrogate end points

Staffan Svensson + Healthy Skepticism members

Angered Health Care Centre and Dept of Clinical Pharmacology, University of Gothenburg, Sweden

1 Introduction

In randomised controlled trials (RCT) of prophylactic medical interventions, outcomes may either be the disease itself (eg stroke) or some substitute measure thought to be associated with the disease (eg high cholesterol). Such intermediate measures are known as surrogate end points¹. Here is a definition of a surrogate end point:

A laboratory measurement or a physical sign used as a substitute for a clinically meaningful end point that measures directly how a patient feels, functions or survives [17] (quoted in [18]).

2 Benefits of surrogate end points

Clinically meaningful end points (“hard outcomes”) are often very time consuming and expensive to study. In a stroke prevention RCT, for example, researchers may have to recruit 10,000 participants and wait 5 years for the results. In comparison, surrogate end points are much faster and easier to study: for an RCT on a cholesterol lowering drug it may suffice with a few hundred patients for a few months. Surrogate end points are therefore much cheaper than hard outcomes.

3 Problems with surrogate end points

On the other hand, surrogate end points may be misleading when they do not translate into clinically important outcomes, or when the clinical outcome is opposite to what was expected. The latter was the case with type-I antiarrhythmics in patients who had heart rhythm disturbances after myocardial infarction. Among these patients, it had been shown that the antiarrhythmic drugs encainide and flecainide decreased electrocardiographic (ECG) instances of arrhythmia (the surrogate end point), which was one of the reasons for their use. When tested in an RCT, however, patients who took encainide and flecainide turned out to be more than two times more likely to die from cardiac arrest or other causes than those randomised to placebo [1]. In effect, the drugs improved ECG looks but killed people.

3.1 Use in clinical practice

Because of this and many other examples of failed surrogate end points, it is not advisable to start therapy based on such results only. This is, nevertheless, often done in clinical practice [19]. Part of the problem is that the surrogate end point is not seldom confused with the disease itself, rather than being more properly identified as a risk factor for it. This is further complicated by the fact that such mixing of the concepts is at times warranted – some surrogate end points may indeed be both part of the disease and a risk factor for it. Blood glucose, for example, is a surrogate end point as regards late complications of diabetes (eg myocardial infarction), but at the same time high blood glucose in itself may cause a number of symptoms such as tiredness and increased thirst.

3.2 Use in advertisements

Where hard outcome data is not available, pharmaceutical companies often use surrogate end points in their marketing. An example is Figure 1, where an antidiabetic drug (sitagliptin) is recommended on the grounds that it, among other things, “enhances incretins” and “enhances physiological glucose control”. Such enhancements may of course be of benefit, or they may not: no-one knows as there is as yet no data on the effect, if any, of sitagliptin on mortality or diabetic complications [20].

¹Synonymous terms are “surrogate outcome measures” and “surrogate variables”.



Figure 1: An ad for sitagliptin (Januvia)

4 A collection of failed surrogate end points

For reasons outlined above, it would be useful to have a collection of examples to illustrate the dangers of relying too much on surrogate end points. A request for input was therefore made to members of the Healthy Skepticism mail list. There was some debate about the purpose and scope of such a collection: should surrogate end points with no proven positive clinical outcomes qualify, or should the list be restricted to cases with proven negative outcomes? In the end, the latter option was chosen as it was felt to be more attention grabbing (and thus better for teaching), and as it made for a much shorter list. The result is shown in Table 1.

4.1 Choice of examples

Many examples were suggested, 16 of which are included in the present version of the table. Inclusion was based on recommendations and assessment of references. No formal inclusion criteria were used. The table necessarily simplifies complex issues and many of the more recent examples are very controversial as to what effect they should have on medical practice. If the table is to be used in teaching or debates, the presenter should be aware of such issues.

4.2 Future development + contributors

This is a work in progress which needs to be extended, modified and updated. Hopefully, those who have already contributed will continue to do so, as will others who take an interest in the matter. So far, the following people have offered assistance: Peter Davoren, Mark McConnell, Ralph Faggotter, Joseph Ross, Joel Lexchin, Philip Clarke, Tom Perry, Tim McCulloch, Jerome Hoffman, David Egilman, Jim Dickinson, Dinesh Mehta, Adriane Fugh-Berman and Marilyn Mann. Thanks!

References

- [1] Echt DS, Liebson PR, Mitchell LB, Peters RW, Obias-Manno D, Barker AH, et al. Mortality and morbidity in patients receiving encainide, flecainide, or placebo. The Cardiac Arrhythmia Suppression Trial. *N Engl J Med.* 1991 Mar;324(12):781–788.
- [2] ALLHAT Collaborative Research Group. Major cardiovascular events in hypertensive patients randomized to doxazosin vs chlorthalidone: the antihypertensive and lipid-lowering treatment to prevent heart attack trial (ALLHAT). *JAMA.* 2000 Apr;283(15):1967–1975.
- [3] Rossouw JE, Anderson GL, Prentice RL, LaCroix AZ, Kooperberg C, Stefanick ML, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women’s Health Initiative randomized controlled trial. *JAMA.* 2002 Jul;288(3):321–333.
- [4] Barter PJ, Caulfield M, Eriksson M, Grundy SM, Kastelein JJP, Komajda M, et al. Effects of torcetrapib in patients at high risk for coronary events. *N Engl J Med.* 2007 Nov;357(21):2109–2122.
- [5] Sackner-Bernstein JD, Kowalski M, Fox M, Aaronson K. Short-term risk of death after treatment with nesiritide for decompensated heart failure: a pooled analysis of randomized controlled trials. *JAMA.* 2005 Apr;293(15):1900–1905.
- [6] Hagenauer D, Welch V, Shea B, Tugwell P, Adachi JD, Wells G. Fluoride for the treatment of postmenopausal osteoporotic fractures: a meta-analysis. *Osteoporos Int.* 2000;11(9):727–738.
- [7] Nissen SE, Wolski K. Effect of rosiglitazone on the risk of myocardial infarction and death from cardiovascular causes. *N Engl J Med.* 2007 Jun;356(24):2457–2471.
- [8] Gerstein HC, Miller ME, Byington RP, Goff DC, Bigger JT, Buse JB, et al. Effects of intensive glucose lowering in type 2 diabetes. *N Engl J Med.* 2008 Jun;358(24):2545–2559.
- [9] Packer M, Carver JR, Rodeheffer RJ, Ivanhoe RJ, DiBianco R, Zeldis SM, et al. Effect of oral milrinone on mortality in severe chronic heart failure. The PROMISE Study Research Group. *N Engl J Med.* 1991 Nov;325(21):1468–1475.
- [10] Hampton JR, van Veldhuisen DJ, Kleber FX, Cowley AJ, Ardia A, Block P, et al. Randomised study of effect of ibopamine on survival in patients with advanced severe heart failure. Second Prospective Randomised Study of Ibopamine on Mortality and Efficacy (PRIME II) Investigators. *Lancet.* 1997 Apr;349(9057):971–977.
- [11] Califf RM, Adams KF, McKenna WJ, Gheorghide M, Uretsky BF, McNulty SE, et al. A randomized controlled trial of epoprostenol therapy for severe congestive heart failure: The Flolan International Randomized Survival Trial (FIRST). *Am Heart J.* 1997 Jul;134(1):44–54.
- [12] University Group Diabetes Program. The University Group Diabetes Program. A study of the effects of hypoglycemic agents on vascular complications in patients with adult-onset diabetes. V. Evaluation of phenformin therapy. *Diabetes.* 1975;24 Suppl 1:65–184.
- [13] Phrommintikul A, Haas SJ, Elsik M, Krum H. Mortality and target haemoglobin concentrations in anemic patients with chronic kidney disease treated with erythropoietin: a meta-analysis. *Lancet.* 2007 Feb;369(9559):381–388.
- [14] Fergusson DA, Hébert PC, Mazer CD, Fremes S, MacAdams C, Murkin JM, et al. A comparison of aprotinin and lysine analogues in high-risk cardiac surgery. *N Engl J Med.* 2008 May;358(22):2319–2331.
- [15] Devereaux PJ, Yang H, Yusuf S, Guyatt G, Leslie K, Villar JC, et al. Effects of extended-release metoprolol succinate in patients undergoing non-cardiac surgery (POISE trial): a randomised controlled trial. *Lancet.* 2008 May;371(9627):1839–1847.
- [16] Oliver MF, et al. A co-operative trial in the primary prevention of ischaemic heart disease using clofibrate. *Br Heart J.* 1978 Oct;40(10):1069–1118.
- [17] Temple RJ. A regulatory authority’s opinion about surrogate endpoints. In: Nimmo W, Tucker GT, editors. *Clinical measurement in drug evaluation.* John Wiley & Sons Inc; 1995. p. 57.
- [18] Bucher HC, Guyatt GH, Cook DJ, Holbrook A, McAlister FA. Users’ guides to the medical literature: XIX. *JAMA.* 1999 Aug;282(8):771–778.
- [19] Lexchin J. Nesiritide and Lessons Unlearned. *Israeli Journal of Emergency Medicine.* 2005 Sep;5(6):69–72.
- [20] Richter B, Bandeira-Echtler E, Bergerhoff K, Lerch CL. Dipeptidyl peptidase-4 (DPP-4) inhibitors for type 2 diabetes mellitus. *Cochrane Database Syst Rev.* 2008;(2):CD006739.

Drug(s)	Surrogate end point	Hard end point	Reference
Encainide, flecainide	↓ Premature depolarizations	↑ Cardiac death & arrest	[1]
Doxazosine	↓ Blood pressure	↑ Heart failure	[2]
Estrogen, progestin	↓ Cholesterol	↑ Stroke, dementia, breast cancer	[3]
Torcetrapib	↓ Cholesterol	↑ Mortality	[4]
Nesiritide	↓ Wedge pressure, dyspnoea	↑ Mortality	[5]
Fluoride	↑ Bone mineral density	↑ Nonvertebral fractures	[6]
Rosiglitazone	↓ S-HbA _{1c}	↑ Myocardial infarction	[7]
Intensive treatment in Diabetes type 2	↓ S-HbA _{1c} < 6%	↑ Mortality	[8]
Milrinone	↑ Cardiac contractility	↑ Mortality	[9]
Ibopamine	↑ Cardiac contractility	↑ Mortality	[10]
Epoprostenol	↑ Cardiac contractility	↑ Mortality	[11]
Tolbutamide, phenformin	↓ Blood glucose	↑ Mortality	[12]
Erythropoietin	↑ Hemoglobin in chronic renal failure	↑ Mortality	[13]
Aprotinin	↓ Blood loss in cardiac surgery	↑ Mortality	[14]
Metoprolol	↓ Perioperative ischaemia	↑ Mortality	[15]
Clofibrate	↓ Cholesterol	↑ Mortality	[16]

Table 1: The table