

Surrogate Endpoints: How Reliable Are They?

By Kunal Gupta, Jyotsna Gupta, and Sukhdeep Singh

In clinical research, a clinical endpoint is a sign or symptom that constitutes a target outcome of a study. Unfortunately, there may be no convenient clinical endpoint for some studies. For example, the clinical endpoint of hypertension is death over the next 50 years, so blood pressure is used as a "surrogate endpoint." Similarly, the concentration of viral particles is used in HIV/AIDS studies. Bone density is used in osteoporosis studies. A true endpoint must be a medical condition that a patient would want treated and a physician or dentist would agree deserves treatment.

The U.S. National Institutes of Health (NIH) defines surrogate endpoint as "a biomarker intended to substitute for a clinical endpoint."^{1,2} Temple defined it as "a laboratory measurement or a physical sign used as a substitute for a clinically meaningful endpoint that measures directly how a patient feels, functions or survives."³ It may be related to the clinical (true) endpoint, but the relationship between the two may not be direct.⁴

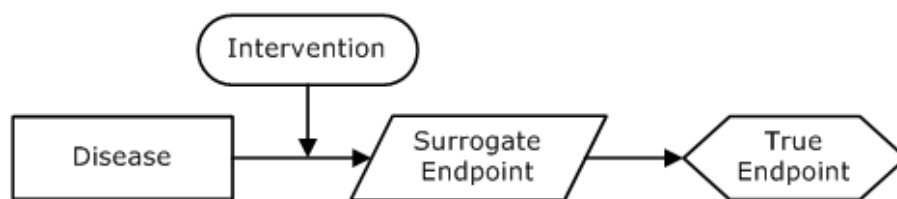
Surrogate endpoints are physiological or biochemical markers that are easier to measure than true endpoints. Surrogate endpoints often occur soon after the intervention in the disease process. Such surrogate endpoints can, therefore, be viewed as intermediate endpoints. Also, surrogate endpoints may not be of direct practical importance to the patient. A given disease process may have multiple markers that could be used as the surrogate endpoint. Or, more than one could be used.

Surrogate endpoints offer three main advantages to clinical studies:

- The study becomes simpler. Since surrogates are usually measures of symptoms or laboratory biomarkers, they make it easier to quantify comparisons.
- The study becomes shorter. It generally takes less time to see the effect of an intervention on a surrogate than on the final clinical outcome, especially if the surrogate marks an intermediate point in the disease process.
- The study becomes less expensive. Since the study duration is shorter, the cost decreases. Measurement of the surrogate may be less costly than measurement of the true outcome. In addition, waiting for a clinical outcome may involve more medical care for sicker patients.

For a surrogate endpoint to be an effective substitute for the true endpoint, the intervention's effect on the surrogate must reliably predict the intervention's effect on the true endpoint. An ideal surrogate endpoint is one in which all mechanisms of action to the true endpoint are mediated through the surrogate endpoint (Figure 1). But in practice, this ideal situation is often not the case.

Figure 1. Surrogate Endpoint



For a surrogate endpoint to be accepted as a valid substitute for the true endpoint, three conditions must be met:

- Informativeness. There must be evidence that the surrogate predicts the true endpoint.⁵
- Specificity. The intervention's effect on the true endpoint must be mediated through the surrogate endpoint.⁵
- Completeness. A study involving the surrogate endpoint must also capture all the information on adverse effects associated with the intervention.

Table 1 presents examples of the surrogate and true endpoints of some disease processes:

Table 1. Examples of Surrogate and True Endpoints

Medical Disease	True Endpoint	Surrogate Endpoint
Heart Disease	Survival Rate	Cholesterol Level
HIV Infection	Survival Rate	CD4 Counts
Hyperthyroidism	Weight Loss, Bulging Eyeballs, Tremors	Serum T3 Level
Lung Cancer	Survival Rate	Tumor Shrinkage

Dental Disease	True Endpoint	Surrogate Endpoint
Dental Caries	Pain, Sensitivity, Food Lodgment	DMFT Index, Streptococcus mutans counts, Saliva pH
Gingivitis	Bleeding Gums, Halitosis, Gingival Swelling	Gingival Index, Plaque Index
Malocclusion	Esthetics, Function	Angles Class I Molar Relation
Periodontitis	Tooth Loss	Pocket Depth, Grade of Mobility

Surrogate Endpoints Can be Deceptive

Surrogate endpoints have a history of not being related to the desired clinical outcome.⁶ Some studies using surrogate endpoints have shown benefit from the study drug, but other studies, looking at true endpoints, have subsequently not shown a benefit and may even have shown a harm.⁷ The use of surrogate endpoints can be deceiving:

- The FDA conditionally approved flosequinan, a drug for congestive heart failure, because it could improve exercise tolerance in patients who did not respond to or could not tolerate a full regimen of other agents, including diuretics and angiotensin-converting enzyme inhibitors. Exercise tolerance was measured using treadmill tests. Later, the PROFILE study provided significant evidence that flosequinan increased mortality, leading the manufacturer to withdraw the product from the market.⁸
- The drug Avastin was introduced in 2004 to treat colorectal cancers. In 2006, the FDA also approved it to treat lung cancers. In 2008, the FDA also approved it for treatment of advanced breast cancer. These approvals were based on surrogate endpoints like decrease in the tumor size. However, in longer-term studies of breast cancer, the treatment was found not to increase lifespan, the true endpoint. Further, in some breast cancer patients, the drug had significant side effects, such as heart

failure and gastrointestinal perforation. Nevertheless, the FDA allowed the drug to stay on the market for treatment of advanced breast cancer because it temporarily stopped the symptoms (other than tumor size) of the disease from worsening.

- Vytorin (a combination of Zocor and Zetia) was tested in a clinical study. Although Vytorin reduced blood levels of LDL (bad) cholesterol and C-reactive protein, it did not reduce arterial plaque buildup nor increase survival rate. The reasons for this disappointing result are unclear.
- FDA approved three drugs (encainide, flecainide and moricizine) for use in life-threatening or severely symptomatic arrhythmias, based on a surrogate endpoint of EKG results. Later, the Cardiac Arrhythmia Suppression Trial (CAST) evaluated how the three drugs affected survival of patients who had had a myocardial infarction and had at least 10 premature ventricular beats per hour. Unfortunately, mortality was higher in the treatment arms than in the placebo arm.⁹

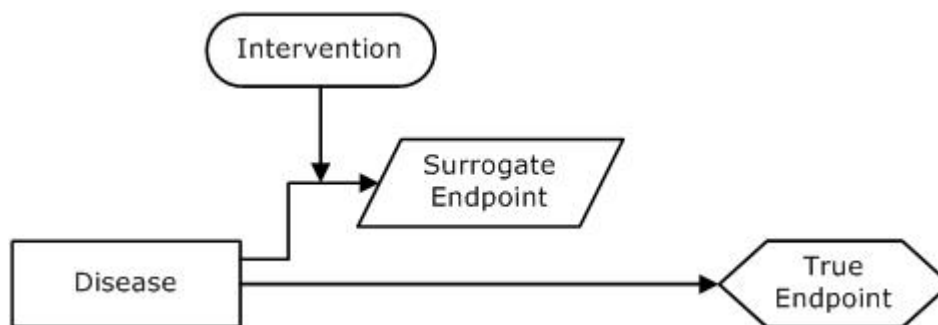
The scientists who wrote the protocols for these studies presumably had reason to believe their surrogate endpoints were sufficient. Clearly, they were not. Therefore, serious consideration should be given to long-term follow-up studies of mortality and other serious adverse effects for any study that uses surrogate endpoints.

How Surrogate Endpoints Fail

Surrogate endpoints can fail for several physiological reasons.

First, the surrogate endpoint may not be on the causal pathway of the disease process (Figure 2). Such situations arise when the surrogate endpoint lies on an entirely different pathway than the true outcome, and the pathways are not related to each other. In such situations, an intervention can affect the surrogate endpoint, but not the true endpoint. For example, the surrogate may measure a symptom that is not related to the final outcome of a disease process.

Figure 2. Surrogate Endpoint not on Pathway

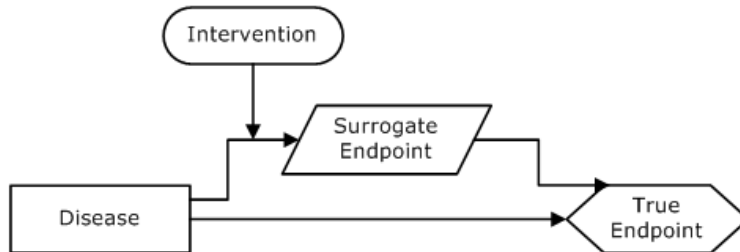


An example of this situation would be assessing the molar relationship based on Angle's Malocclusion Classification as a measurement of effectiveness of an orthodontic treatment for malocclusion (defective bite). In this example, molar relationship is not a true endpoint because it is not an outcome of interest to the patient. A true endpoint for malocclusion would be esthetics or function, i.e., a nice smile or easy chewing. Patients with Angle's Class I (best) molar relationship may or may not have good esthetics and function. Molar relationship does not necessarily cause bad esthetics or function, so molar relationship is a weak surrogate endpoint.

Second, the surrogate may mediate one pathway, but not all pathways (Figure 3). In such a situation, the intervention may show effect on the surrogate endpoint, but not on an unrelated pathway to the true endpoint. In some cases, the intervention may have opposite effects on the two pathways, leading to exactly wrong conclusions.⁴

An example of this situation would be evaluation of the level of cytokines as an indicator of

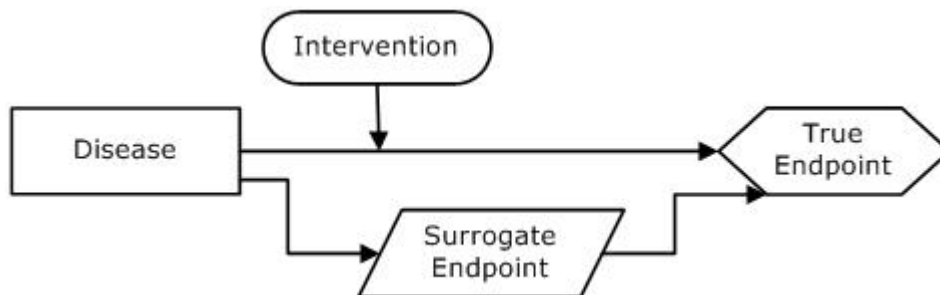
Figure 3. Surrogate Endpoint not on all Pathways



inflammation, after use of a mouth rinse in a patient suffering from periodontitis (disease of the periodontium, the tissue supporting the teeth). In this example, evaluation of cytokine level is not a true endpoint because the patient cannot perceive changes in cytokine levels. A true endpoint for a patient suffering from periodontitis is tooth loss or increase in tooth mobility. Even if a mouth rinse is helpful in reducing the cytokine levels, a patient can still end up with tooth loss or increased tooth mobility because the disease may be progressing through other pathways involving prostaglandins, collagenases, etc.

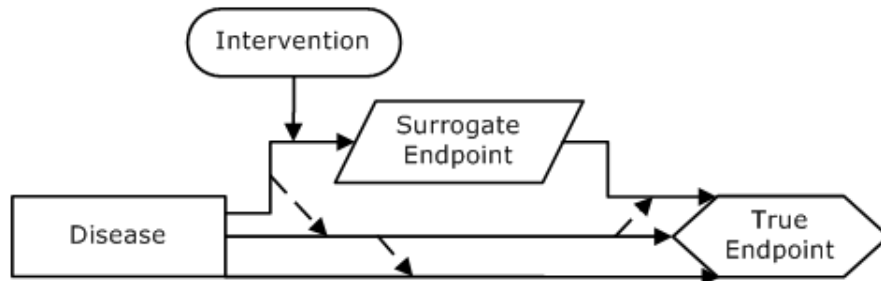
Third, the surrogate endpoint may not be on the pathway of the intervention's effect (Figure 4). In this situation, the surrogate endpoint is completely oblivious to the intervention. An example of this situation would be a clinical trial in which a mouth rinse is used to treat gingivitis (disease of the gums, which causes bleeding gums, swollen gums, and persistent halitosis). A clinical study might use gingival index, a measurement primarily based on gum bleeding, redness and swelling, as the surrogate endpoint. Although the mouth rinse may affect a factor that contributes to halitosis, it may have no effect on the surrogate endpoint of gingival index.

Figure 4. Surrogate Endpoint not on Intervention's Pathway



Fourth, in the extreme case, the surrogate endpoint is on only one pathway of a multifactorial disease that operates with complex interactions between the pathways (Figure 5). A given pathway may be significant under some poorly understood conditions, but not

Figure 5. Surrogate Endpoint in Multifactorial Disease



others. In this situation, the same study may yield different results at different times because the underlying mechanisms are poorly understood.

An example of this situation is dental caries. Dental caries is a multifactorial disease, the main causative factors being diet, size and shape of the tooth, and salivary factors like pH, viscosity, buffering capacity, and bacterial counts. None of these factors have a definitive affect on the true endpoints: pain, sensitivity, holes and food lodgment. A clinical study might test an intervention for reducing dental caries by measuring changes in the salivary counts of *Streptococcus mutans* and *Lactobacillus*. In this example, *Streptococcus mutans* and *Lactobacillus* are not true endpoints because the patient cannot perceive changes in their counts. Elevated bacterial levels may contribute to formation of dental caries, but only in the presence of other factors, so it is a weak surrogate endpoint.

Conclusion

In some cases, surrogate endpoints are excellent substitutes for true endpoints. In others, they are not. Investigators should choose them carefully after ascertaining the mechanisms of action of the disease process. Until a surrogate endpoint is proven to be reliable, follow-up studies with true endpoints are well-advised.

References

1. Controlled Clinical Trials 2001; 22: 485–502.
2. Cohn JN "Introduction to Surrogate Markers". Surrogate Endpoints in Clinical Trials: Are We Being Misled? Circulation (American Heart Association) 2004; 109: 20.
3. Temple RJ. A regulatory authority's opinion about surrogate endpoints. In: Clinical measurement in drug evaluation. Nimmo WS, Tucker GT, editors. New York: Wiley. 1995
4. Psaty BM, Weiss NS, Furberg CD, *et al.* "Surrogate endpoints, health outcomes, and the drug approval process for the treatment of risk factors for cardiovascular disease". JAMA 1999; 282: 786–790
5. Hujoel PP. Endpoints in periodontal trials: the need for an evidence-based research approach. *Periodontology 2000*; 36: 196–204.
6. D'Agostino RB. The slippery slope of surrogate outcomes. *Curr Control Trials Cardiol* 2000. 2: 76-78.

7. Fleming TR, DeMets DL. Surrogate Endpoints in Clinical Trial-Are we being Misled? *Annals of Internal Medicine*, 1996. 125:605-613.
8. Packer M, Rouleau J, Swedberg K, Pitt B, Fisher L, Klepper M, et al. Effect of Flosequinan on survival in chronic heart failure: preliminary results of the PROFILE study [Abstract]. *Circulation*. 1993; 88(Suppl 1):I-301.
9. Effect of the antiarrhythmic agent moricizine on survival after myocardial infarction. The Cardiac Arrhythmia Suppression Trial II Investigators. *N Engl J Med*. 1992; 327:227-33.

Authors

Kunal Gupta, MDS, is a Reader in the Department of Pedodontics & Preventive Dentistry, Institute of Dental Studies and Technologies, Modinagar. Contact him at kunalgupt@yahoo.com.

Jyotsna Gupta, Msc(Biotechnology), Cert.(CRCDM), is a Research Scientist "A" Grade in the CTRI Project, National Institute of Medical Statistics, New Delhi.

Sukhdeep Singh, MDS, is a Professor and Head of the Department of Pedodontics & Preventive Dentistry, Institute of Dental Studies and Technologies, Modinagar.