Are surrogate endpoints useful in clinical studies?

Surrogate endpoints are often used as measures of therapeutic effect in clinical studies, for example, blood pressure as a risk indicator for cardiovascular disease. Sometimes the presumed association is uncertain. What do we need to be aware of when we interpret the results of such studies, and what is the usefulness of surrogate endpoints for patients?

A clinical endpoint describes the therapeutic effect that is most beneficial for the patient. When undergoing cancer treatment, most patients find that new treatment is meaningful if it prolongs their lives and/or improves the quality of their remaining life. Survival and health-related quality of life are therefore sound clinical endpoints. The word «surrogate» means «instead of» or «a substitute for something better». Surrogate endpoints are frequently used to replace clinical endpoints in studies, to save time and money. Since they can be measured earlier in the course of illness and thereby result in lower attrition, more patients will achieve the endpoint. Studies of this type therefore require a smaller sample size to be able to demonstrate a significant difference between the patient groups.

When survival is the endpoint of a study, median survival should be reached (one-half of the patients are dead) before it is possible to draw a definitive conclusion. This applies particularly to palliative studies. For patients with cancer diagnoses with a good prognosis, it may take a very long time before median survival is reached. Therefore the choice is often to report three or five-year survival (1). For rare diseases and for treatment aimed at small patient groups, it is difficult to include a sufficient number of patients to be able to demonstrate a significant difference between the treatments. Long-lasting trial periods are costly and may make it less attractive for the industry to invest in such research.

Health-related quality of life comprises physical, mental and social dimensions and must be reported by the patients themselves (2). Lack of measuring instruments was previously used as an argument against using quality of life as an endpoint, but today there are validated methods of measurement (3). Despite the fact that researchers are continually encouraged to use quality of life as an endpoint in clinical studies, it is still infrequently used (4). We believe this may be due to a lack of knowledge on how, in practical terms, such studies can be conducted.

How to recognise a surrogate endpoint?

It is not always easy to extract from an article whether a surrogate endpoint has been used as a measurement of efficacy. It can be helpful to analyse whether the endpoint is meaningful for the patient. The patient does not necessarily observe that his/her blood pressure is lower or that a tumour is growing when it is only visible on x-ray images. Other surrogate endpoints, for example for survival in cancer treatment, are tumour response, local control and progression-free and symptom-free survival. The Karnofsky Performance Scale and recording of adverse effects (toxicity) of the treatment (5) have often been used as surrogate endpoints for health-related quality of life. However, these measure only one dimension of quality of life, and are not usually patient-reported. Many authors through the years have claimed that tumour response or progression-free survival is synonymous with improved quality of life, without measuring quality of life (6). The effect of progression-free survival on the patients' quality of life has yet to be determined (7).

«Always maintain a critical eye when you assess the results of studies in which surrogate endpoints have been used»

Is the surrogate endpoint valid?

In studies using a surrogate endpoint, it is important to evaluate whether this can really replace the clinical endpoint. For example, has it been demonstrated that tumour response leads to improved survival for the patient group and treatment in question. An optimal surrogate endpoint requires that there is a simple causal relationship between disease, treatment and clinical endpoint (Fig. 1) (8). In reality it is rarely that simple, and there may be many reasons why a surrogate endpoint fails to give an accurate picture of the clinical endpoint (Fig. 2) (8). In 1997 the U.S. Food and Drug Administration (FDA) accepted surrogate endpoints as being sufficient for approval of drugs (9). Trials had to demonstrate that the endpoint was valid, and further studies – with a clinical endpoint – had to be conducted in order to demonstrate that the treatment was clinically beneficial for the patients. There are two main methods that researchers can use to evaluate the validity of a surrogate endpoint (10). One is considered adequate to evaluate the validity of a single study, and is based on statistical criteria formulated by Prentice as early as in 1989 (11). It has been widely used (12), but many believe the criteria are too stringent. The meta-analytical method, currently considered to be the gold standard, evaluates the validity of a surrogate endpoint based on pooled results from previous studies (13). It calculates the correlation between the surrogate endpoint and the clinical endpoint, and the correlation between the efficacy of the treatment on the surrogate endpoint and on the clinical endpoint (10).

There are many examples of a surrogate endpoint being valid for one treatment or diagnosis, but not necessarily valid in another context. A study based on data from ten chemotherapy studies in advanced colon cancer demonstrated high correlation between progression-free survival and survival, i.e. progression-free survival was a valid surrogate endpoint for survival (14). In another study of the same patient group, but with a different treatment, progression-free survival was significantly prolonged without any difference in survival being demonstrated (15). Bevacizumab in the treatment of metastatic breast cancer was approved based on a prolonged median progression-free survival of 5.9 months (9). However, further studies showed a lack of correlation between progression-free survival and total survival, and the FDA withdrew its approval in 2011. In addition, no improvement in quality of life was demonstrated, and several adverse effects were reported (16).

Are surrogate endpoints useful?

Surrogate endpoints in clinical studies are sometimes justified; in other cases their relation to clinical endpoints and validity

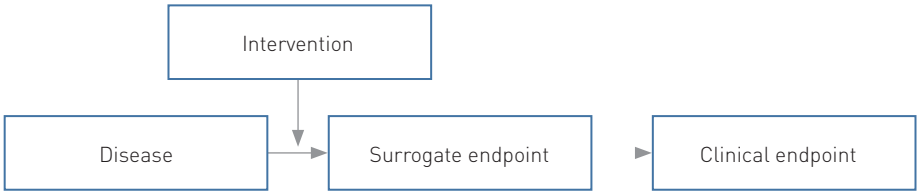

Figure 1 *An optimal surrogate endpoint requires a simple causal relationship between disease, intervention and clinical endpoint, and that all the treatment effect is captured by the surrogate endpoint (8)*

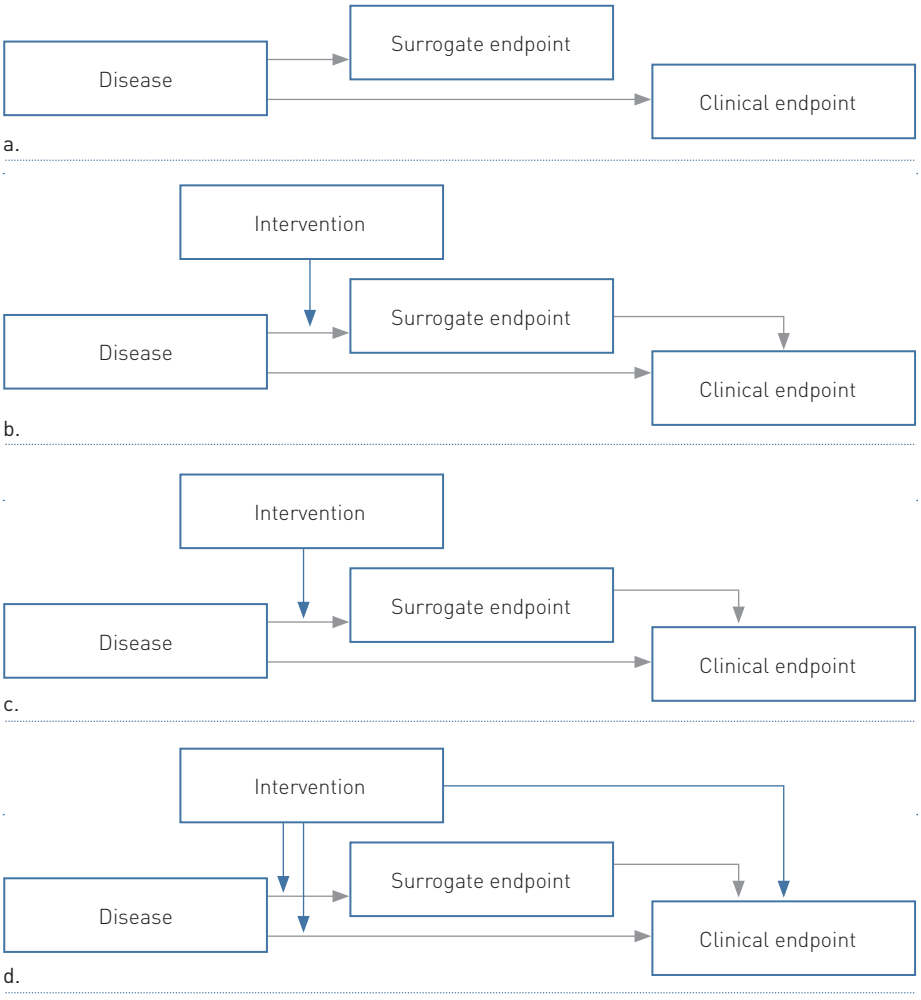

Figure 2 *Possible reasons for failure of a surrogate endpoint (8). Although the surrogate endpoint is correlated with the clinical endpoint, it does not necessarily involve the same pathophysiological mechanisms. a) The effect of the disease on the surrogate endpoint is independent of its effect on the clinical endpoint, b) The disease has several mechanisms of actions and the treatment acts only on the pathway that is captured by the surrogate endpoint. The surrogate endpoint can thereby capture the efficacy of the treatment, but its significance for the clinical outcome is uncertain, c) The treatment acts upon a mechanism that is independent of the surrogate endpoint, d) The treatment can have another, unwanted impact on the clinical endpoint than through the surrogate endpoint*

in the actual study are uncertain. Always maintain a critical eye when you assess the results of studies in which surrogate endpoints have been used. Check whether the author has assessed the validity of the surrogate endpoint, and whether it has been substantiated that the particular treatment is beneficial for the patients. Positive results of a study do not necessarily mean that it is clinically beneficial for patients. It is also important to assess whether the results are representative for patients in clinical practice. Finally, such assessments will be important in communication with the patients. Many patients are increasingly well informed about results of clinical research and they need adequate information on what published results mean for them.

The article is based on Cecilie Delphin Amdal's lecture for the PhD in clinical medicine, University of Oslo, 13 September 2013.

Cecilie Delphin Amdal
cecia@ous-hf.no
Kristin Bjordal

Cecilie Delphin Amdal (born 1969) is a specialist in oncology, PhD and senior consultant at the Department of Oncology, Oslo University Hospital. She has conducted research on the assessment of therapeutic efficacy in clinical studies on patients with oesophageal cancer. The author has completed the ICMJE form and declares no conflicts of interest.

Kristin Bjordal (born 1958) is a specialist in oncology, Head of Research Support Services at Oslo University Hospital and professor at the Institute of Clinical Medicine, University of Oslo. She has conducted clinical research on quality of life measures and other patient-reported outcomes (PRO).
The author has completed the ICMJE form and declares no conflicts of interest.

References

1. Fosså A, Kaasa S, Sørbye H et al. Dokumentasjon av behandlingseffekt. I: Dahl O, Lehne G, Baksaas I et al, red. Cytostatikaboken. 7. utg. Bergen: Haukeland universitetssykehus, 2009.
2. Bjordal K, Kaasa S. Psychometric validation of the EORTC Core Quality of Life Questionnaire, 30-item version and a diagnosis-specific module for head and neck cancer patients. Acta Oncol 1992; 31: 311–21.
3. Young T, de Haes H, Curran D et al. Guidelines for assessing quality of life in EORTC clinical trials. Brüssel: EORTC Quality of Life Study Group, 2002. http://groups.eortc.be/qol/sites/default/files/archives/clinical_trials__guidelines_qol.pdf (13.5.2014).
4. Amdal CD, Jacobsen AB, Guren MG et al. Patient-reported outcomes evaluating palliative radiotherapy and chemotherapy in patients with oesophageal cancer: a systematic review. Acta Oncol 2013; 52: 679–90.
5. Trotti A, Colevas AD, Setser A et al. CTCAE v3.0: development of a comprehensive grading system for the adverse effects of cancer treatment. Semin Radiat Oncol 2003; 13: 176–81.
6. Iaffaioli RV, Frasci G, Palmieri G et al. Chlorambucil, vincristine and cytarabine (COA) treatment of low grade lymphomas. Leuk Lymphoma 1995; 17: 147–53.
7. Gutman SI, Piper M, Grant MD, Basch E. Progression-free survival: what does it mean for psychological well-being or quality of life? Rapport nr. 13-EHC074-EF. Rockville, MD: Agency for Healthcare Research and Quality, 2013.
8. Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? Ann Intern Med 1996; 125: 605–13.
9. U.S. Food and Drug Administration. Food and Drug Administration modernization act (FDAMA) of 1997. www.fda.gov/RegulatoryInformation/Legislation/FederalFoodDrugandCosmeticActFDCAct/SignificantAmendmentstotheFDCAct/FDAMA/default.htm (13.5.2014).
10. Fiteni F, Westeel V, Pivot X et al. Endpoints in cancer clinical trials. J Vis Surg 2014; 151: 17–22.
11. Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. Stat Med 1989; 8: 431–40.

>>>

12. D'Amico AV, Chen MH, de Castro M et al. Surrogate endpoints for prostate cancer-specific mortality after radiotherapy and androgen suppression therapy in men with localised or locally advanced prostate cancer: an analysis of two randomised trials. Lancet Oncol 2012; 13: 189–95.
13. Buyse M, Molenberghs G, Burzykowski T et al. The validation of surrogate endpoints in meta-analyses of randomized experiments. Biostatistics 2000; 1: 49–67.
14. Buyse M, Burzykowski T, Carroll K et al. Progression-free survival is a surrogate for survival in advanced colorectal cancer. J Clin Oncol 2007; 25: 5218–24.
15. Van Cutsem E, Peeters M, Siena S et al. Open-label phase III trial of panitumumab plus best supportive care compared with best supportive care alone in patients with chemotherapy-refractory metastatic colorectal cancer. J Clin Oncol 2007; 25: 1658–64.
16. U.S. Food and Drug Administration. FDA Commissioner announces Avastin decision. www.fda.gov/newsevents/newsroom/pressannouncements/ucm280536.htm (7.5.2014).

Received 21 March 2014, first revision submitted 9 May 2014, accepted 12 May 2014. Editor: Hanne Støre Valeur.