

## Short communication

# Clinical outcomes: to be a surrogate or not to be ...?

David Cameron

NCRN Co-ordinating Centre, University of Leeds, Cookridge Hospital, Hospital Lane, Leeds, LS16 6QB, UK

Corresponding author: David Cameron, [d.cameron@ncrn.org.uk](mailto:d.cameron@ncrn.org.uk)

Published: 20 December 2007

This article is online at <http://breast-cancer-research.com/content/9/S2/S26>

© 2007 BioMed Central Ltd

*Breast Cancer Research* 2007, **9(Suppl 2)**:S26 (doi:10.1186/bcr1824)

### Introduction

Clinical trials remain the bedrock of the introduction of new therapies for patients with cancer. Over the past 30 years there have been enormous improvements in outcomes for patients with breast cancer, based largely (but not exclusively) on the widespread implementation of the results of randomized trials. Widespread use of screening mammography, breast conservation, adjuvant hormonal therapy, adjuvant chemotherapy and, most recently, adjuvant trastuzumab have all been based on the results of well designed clinical trials. All of these interventions have been shown to either improve survival or, in the case of breast conservation, maintain survival despite less radical surgery. For most women with early breast cancer, it is the avoidance of the death sentence they feel hangs over them when they are first diagnosed with cancer, that is the most important reason why they undergo these treatments.

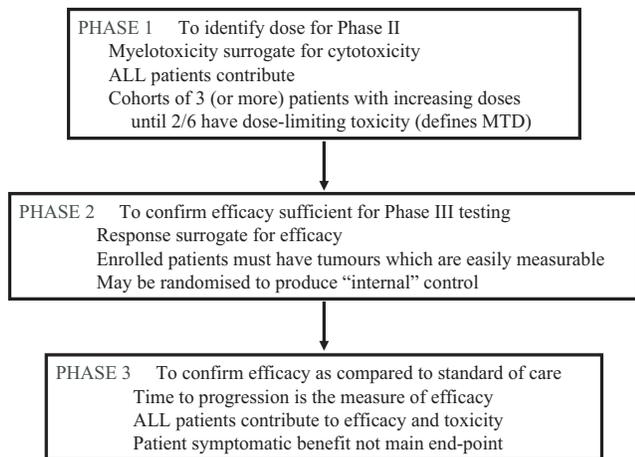
Clinical research in breast cancer remains as active as ever, with newer interventions being tested in ever larger and/or more complex trial designs. Many studies may not be designed to test questions about overall survival, with recent studies also addressing tolerability, issues of limited resources and, increasingly, means to target treatments to the subgroups of patients who really benefit from the specific therapy. The goal of studies that aim to optimize treatment may not be the same for the researcher and the patient. However, most would agree that it would be ideal if we could reduce the diagnosis of breast cancer to one that had the same implications as being diagnosed with a 'touch of blood pressure', namely the concept that although a few patients may still suffer unwanted consequences of the diagnosis, for the vast majority the implication is the necessity to undergo relatively nontoxic treatment that effectively prevents recurrence. To this end, many trials are now designed to address different primary end-points from the traditional one, still much beloved of the US Food and Drug Administration, of overall survival (OS).

### Surrogate end-points: fit for purpose?

Even when improvements in OS are the ultimate aspiration of a study, it is common for reasons of timeliness to make another end-point the primary determinant of success. For adjuvant trials, the use of disease-free survival (DFS) is an accepted surrogate because, to date, mature follow up of both individual trials and their meta-analyses have consistently confirmed that improvements in DFS subsequently translate into firm improvements in OS. For advanced disease studies, time to disease progression (TTP) is frequently used, but in fact this much less commonly precedes clear improvements in OS, although a few notable exceptions exist (use of trastuzumab and in some studies involving taxanes). The reasons for this divergence are not entirely clear, because in both adjuvant and advanced disease studies the possibility exists of a loss of effect of the earlier use of a novel intervention as a consequence of its use after relapse/progression. Furthermore, most successful adjuvant interventions are designed on the basis of a positive improvement in TTP in an advanced disease study, and survival gains in early disease can often be seen despite the lack of gains in OS in a comparable advanced disease study. However, this may not be as big an issue as it first appears because it highlights an important question; when should a trial be designed simply to produce the data required to justify the testing of an intervention in early disease, and when should it be designed to benefit the patient population in whom it is actually being tested?

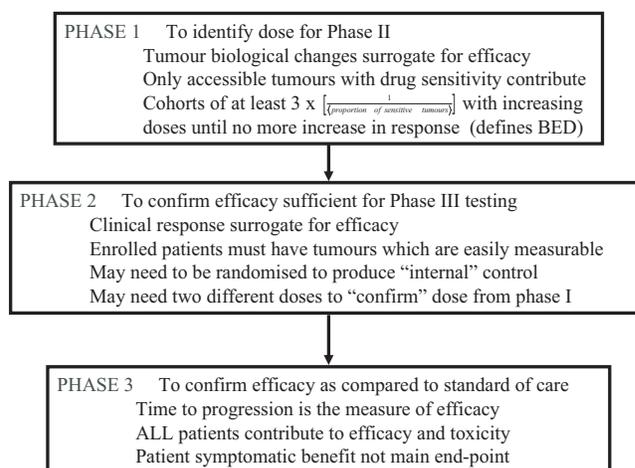
In advanced disease, although an improvement in survival remains the over-arching wish of most patients, when this is not likely there nevertheless remain important improvements that are highly clinically relevant. There are good data that improvements in quality of life and/or reduction in symptoms correlate with tumour response, so that where patients are very symptomatic, demonstration of an increased response rate is a worthwhile gain, provided that this does not come at the cost of major toxicity. In contrast, for many women with more indolent, relatively asymptomatic disease, absence of

Figure 1



Traditional drug development - defining maximum tolerated dose (MTD) using advanced disease as a model.

Figure 2



Biological drug development - defining biologically effective dose (BED) using advanced disease as a model.

progression may be the primary goal. It is interesting to note, therefore, that patients with stable disease for at least 6 months on hormonal therapy often have similar OS to those whose disease actually shrinks on therapy, justifying the use of TTP and/or clinical benefit as the primary end-point for many such studies.

### Conventional drug development model

For cytotoxics, the phase I-II-III development sequence (as shown in Figure 1) was in fact based on the experience that the maximum tolerated dose (MTD) in phase I often turned out to be an effective and tolerated dose in early stage disease. However, for cytotoxics the identification of a dose on the basis that it caused an acceptable level of cytotoxicity in normal tissues, which could then cause a desirable level of cytotoxicity in malignant tissues, is perhaps not a great surprise! All patients contribute to the toxicity end-point in a conventionally designed phase I study, and so it is a relatively efficient way to reach the dose level likely to deliver efficacy if it exists. In contrast (Figure 2), the way in which we might determine a maximum biologically effective dose could be more difficult.

Therefore, in a translational phase I or dose-finding study, in which there is a good surrogate normal tissue with little variation in sensitivity (the biological equivalent of the bone marrow or gastrointestinal mucosa for a classical cytotoxic), this may not be a problem. When the surrogate end-point used is in the tumour, however, we have the problem that we will have to increase the number of patients at each dose level by a factor related to the proportion of patients with sensitive tumours. For example, if only half of the patients have sensitive disease, then in a cohort of three patients (the classic phase I design) there is a one in eight (0.125) chance that no biological effect will be seen at one level simply

because all three tumours were totally resistant; across three active dose levels the chance that we will have one cohort with no responses becomes 0.29, or almost one in three! If one then considers the risk that a biologically maximally effective dose is identified because no responses are seen at a higher dose level, it becomes apparent that this can happen not infrequently. Hence, if only half of the patients have sensitive disease, then one will need at least five patients per dose cohort to have less than a 5% chance of seeing no response at a biologically active dose, a figure that rises as the proportion of responders falls.

### Presurgery systemic therapy

Perhaps the area of clinical research where there is most interest in surrogates is in the use of systemic therapy before surgery. Essentially two models exist: the short 'preoperative' course, designed not to deliver systemic benefit but only short-term biological changes in the breast cancer; and the longer 'neoadjuvant' or 'primary systemic' therapy, in which clinical changes in the primary tumour are the goal, using drugs that deliver systemic benefits. It became clear over many years that when patients are treated with 3 to 6 months of chemotherapy, those patients who have no residual invasive disease in the primary tumour and/or ipsilateral axillary lymph nodes have the best long-term outcome. It was anticipated, therefore, that where additional therapy given before surgery increased the proportion of patients achieving such pathological complete responses (pCR), this would lead to gains in long-term outcome. To date, that has not been confirmed, in particular in the National Surgical Adjuvant Breast and Bowel Project (NSABP) B-18 trial, in which addition of docetaxel before surgery doubled the proportion of patients achieving pCR but made no significant difference to the distant DFS. However, it remains true that on the one

hand the addition of taxanes and/or trastuzumab to neoadjuvant chemotherapy does increase the proportion of patients achieving pCR, and on the other that the addition of those same agents to postoperative adjuvant therapy does lead to improved OS. Therefore, pCR would appear to be a surrogate predictor of a more effective adjuvant therapy; what it does not yet appear to do is to identify the precise patients who will benefit from the therapy!

For patients given short-term exposure to a drug before surgery, there is interest in understanding what biological changes in that context mean for long-term clinical benefit. To date, no study has assessed the prognostic or predictive implications for a specific biological change induced by a short-term exposure to therapy before routinely timed surgery. However, we do have data on the biological changes seen after 2 weeks of exposure to tamoxifen and/or anastrozole, in patients who then continued on therapy, had surgery after a further 10 weeks of therapy and then were encouraged to continue on the same hormonal therapy in the adjuvant setting. In this setting, it was clear that those patients whose tumours had the lowest rate of proliferation at 2 weeks had the lowest rate of recurrence. These data would appear to indicate that where a tumour has a very low proliferation rate (either intrinsically or more probably because hormonal therapy has reduced it), there is a low rate of relapse, at least over the first few years. It does need to be borne in mind that where patients are given adjuvant tamoxifen, at least in the older trials that make up the Oxford Overview, the majority of relapses occur after the 5 years of therapy, so that the low proliferating tumours might just be the ones that take longer to relapse. However, it seems reasonable to take the view that reduction in proliferation after 2 weeks of hormonal therapy is a surrogate for identifying a lower rate of relapse during the first few years after surgery, and there is a high chance that therapies that are better at doing this (as was shown in the above study for the aromatase inhibitor anastrozole) will be better at preventing relapse during the first few years (as ATAC [Arimidex, Tamoxifen Alone or Combination] has shown).

## Conclusion

For most patients, and for most trials in early disease, a treatment that cures more people, or one that cures just as many with fewer unwanted side effects, is the desired goal and no surrogate can really replace this. Use of DFS as the

first end-point, because it has consistently been validated as predating improvements in OS, is perfectly acceptable, as long as studies with novel interventions continue to collect the follow-up data to demonstrate that this linkage applies to newer biological interventions just as it does for conventional treatments.

However, for many studies there is a good clinical justification for using other end-points that meet the needs of the patient population being studied; end-points such as TTP, response rate, and clinical benefit rate are therefore not just to be seen as surrogates but as the most appropriate end-point for that trial.

There remains the question as to when a surrogate is a valid end-point for subsequent improvements in DFS and/or OS. Two obvious candidates in the field of preoperative or neoadjuvant therapy are falls in proliferation in patients treated with primary (short-term or long-term) endocrine therapy, or the proportion of patients achieving pCRs to neoadjuvant chemotherapy. Both appear to be reliable, at least most of the time, but neither has yet been shown to have an ideal level of discrimination between agents that can and cannot lead to changes in ultimate outcomes. This lack of clear linkage could in fact be because some of the agents tested in these studies have themselves not delivered a large enough improvement in efficacy to confirm the link with DFS and/or OS.

Surrogates must be proven to be able to take the place of the real end-point in question, and to date, this linkage is only really established for some treatments. For new biological agents this is even less certain, although the success of trastuzumab suggests that we can have confidence that a model developed for untargeted cytotoxics, and loosely targeted hormonal therapies, may work out for many newer agents. However, in my view, this cannot be taken for granted, and the relevant longer term follow up is necessary in a generation of trials of newer agents with a size such that primary end-points are met within a short time of closure to accrual.

## Acknowledgement

This article has been published as part of Breast Cancer Research Volume 9 Supplement 2, 2007: Controversies in Breast Cancer. The full contents of the supplement are available online at <http://breast-cancer-research.com/supplements/9/S2>.