

Short communication

Foundation clinical trials

David Cameron

NCRN Co-ordinating Centre, MacMillan House, Fairburn House, 71–75 Clarendon Road, Leeds LS2 9LN, UK

Corresponding author: David Cameron, d.cameron@ncrn.org.uk

Published: 18 December 2009

This article is online at

<http://breast-cancer-research.com/supplements/11/S3/S23>

Breast Cancer Research 2009, **11(Suppl 3):S23**

(doi:10.1186/bcr2442)

© 2009 BioMed Central Ltd

The present talk will address the following questions, posed as part of the 2009 Breast Cancer Controversies Meeting: Are clinical trials still essential to develop the new generation of biological agents? Are the traditional large clinical trials appropriate in developing drugs that may have a target in a minority of breast cancers? Are clinical trials compatible with the aim of individualising treatments?

The following answers to these questions are perhaps more conservative than might have been expected: perhaps being conservative in a controversies meeting might in itself be seen as controversial.

The past 50 years of clinical medicine have seen a shift in the evidence required to justify a treatment from being based on experience, gleaned from intelligent analysis of anecdotes and both retrospective and prospective clinical series, to that where the aim is to have randomised clinical trials underpinning the rationale for a particular therapeutic strategy. It is therefore with trepidation that one even asks the question of whether clinical trials are still essential, since it would seem to challenge the work of many eminent physicians and statisticians. But what is surely open to challenge is whether in an era of much greater understanding of the biology of a disease, and with the advent of agents that are only expected to be active in subsets of patients, the traditional inclusive, phase III design is still the best model to provide a robust estimate of the effect of a novel intervention in the relevant population of patients. Clinical trials are simply well-designed clinical experiments to test a particular hypothesis: and most anticipate a future in which the interaction between a treatment and the biology of the disease is pivotal to the hypothesis under scrutiny.

Furthermore, many founder clinical trials have changed practice and outcomes for women with breast cancer. The analysis of improvements in the population of British Columbia breast cancer [1] or in French patients presenting with metastatic disease [2], as well as indirect comparisons

between adjuvant chemotherapy trials (Barros C, personal communication), all confirm that a series of well-designed phase III trials in patient populations not selected for on the basis of tumour biology (and only sometimes on the basis of risk of relapse) have revolutionised the outcomes for women with breast cancer.

It is evident, however, that none of our current armamentarium of systemic treatments works in every patient with breast cancer. Even hormonal therapy, long since demonstrated by the Oxford meta-analyses of many large trials as saving the lives of around 1 in 10 patients diagnosed with breast cancer, does not work in all patients. An intriguing thought experiment is therefore to consider how we might have re-designed those founder clinical trials with our current knowledge.

Data published by Allred and colleagues demonstrate using immunohistochemistry that patients whose breast cancers have no, or very low levels of, oestrogen receptor (ER) have a poorer outcome [3]. These data led to the US Food and Drug Administration definition of ER-positive breast cancer being anyone whose cancer has an Allred category score ≥ 3 (equivalent to at least 1% of cells staining moderately, or at least 10% of cells with any degree of staining). These data, together with Oxford overview data demonstrating no evidence of benefit for the use of adjuvant tamoxifen in women with cancers deemed to be ER-negative by a different, biochemical assay, have led some to believe that all patients with ER-positive breast cancers, as defined by the US Food and Drug Administration, should have adjuvant hormonal therapy. In contrast, there are many data confirming there to be subgroups of patients who have very endocrine-sensitive cancers, who in an age of targeted therapies would seem to be the ideal subpopulation in which one would test the benefit of adjuvant hormonal therapy.

Clues to the diagnostic for this sensitive subgroup can be found in many studies. The likelihood of a breast cancer responding to an endocrine agent, whether in the primary

ER = oestrogen receptor; ITT = intent-to-treat; PgR = progesterone receptor.

cancer or in metastatic disease, is greater in those patients whose tumours have much higher levels of ER expression. Furthermore, retrospective analyses of the Arimidex, Tamoxifen Alone or in Combination trial in fact suggested that those women whose cancers were ER-positive but progesterone receptor (PgR)-negative had greater additional benefits from the use of 5 years' anastrozole in place of tamoxifen, than their counterparts whose tumours had significant levels of PgR. Other analyses in that same study, however, as well as the Breast International Group 1-98 and Tamoxifen and Exemestane Adjuvant Multicentre trials, do not confirm this differential benefit – and rather suggest that ER-positive cancers with either coexpression of HER2 or lack of expression of PgR do worse with any endocrine agent, and that the relative level of additional benefit for the use of an aromatase inhibitor may be similar across all types of ER-positive breast cancers.

So how would we *now* design an adjuvant trial with a control, no-hormonal therapy arm? If we target the trial at those patients with very hormone-sensitive disease (ER strongly positive, PgR-positive and/or HER2-negative), we would enrich the population for those patients with better outcomes, greater relative benefit from the therapy and longer time to recurrence. The trial would certainly be positive, but might take many years to conclude benefit, and would provide data to suggest that only these very endocrine-sensitive patients should be thus treated. Amongst the patients with weakly endocrine-sensitive tumours, however, there is not only better outcome on adjuvant tamoxifen than patients with ER-negative disease, but also evidence of additional benefit from an aromatase inhibitor. It is therefore highly likely that patients with these tumours also benefit from adjuvant hormonal therapy, and indeed – given their overall higher risk of relapse, and their earlier relapse, particularly amongst those patients with greater competing risks of death – their benefits may be similar to the patients with more sensitive tumours. Perhaps we should therefore design our trial to include such patients.

But as one broadens the inclusion criteria, one runs into another problem of design of targeted therapies – how sure can we be that our pathologists could accurately identify the right target group at the time of patient inclusion? This raises two key questions. How can we allow for the fact that we cannot be sure a diagnostic threshold for entry to a trial of a biological agent is optimal? Also, how can we accommodate the likelihood that science will progress during the life of the trial to the extent that we will have a better diagnostic – or at least a better, optimised threshold for that entry diagnostic?

Let me illustrate with a simple example. Trastuzumab was developed as an effective therapy for patients whose tumours overexpress HER2. The efficacy of trastuzumab is certain: its target population is now reasonably, but not perfectly, defined. The pivotal registration trial H0648 allowed inclusion

of patients with either 2+ or 3+ staining by immunohistochemistry in a single central laboratory [4]. Of course, we now know that this is not the precise group of patients whom it is appropriate to treat with the drug, although we are still a little unclear as to the best diagnostic, and have still to deal with inevitable interlaboratory variations in results. Furthermore, this key phase III trial in metastatic breast cancer also allowed patients to be treated with either of two chemotherapy regimens, depending on prior chemotherapy exposure. The overall trial was clearly positive, but it was really only the data from the 31% (145/469) of patients treated with paclitaxel who had 3+ tumours that provided the basis for subsequent licensing and clinical practise. Indeed, these data were not included in the primary manuscript, only in a later article [5]. These data were therefore in reality only hypothesis-generating, and the confidence intervals were inevitably broader than those in the primary paper.

Therefore, although the subsequent randomised trial of docetaxel with or without trastuzumab subsequently confirmed that adding trastuzumab to a taxane in women with HER2 3+ tumours was beneficial, there was a not inconsiderable risk that the benefit observed in the appropriate taxol-treated group of the original phase III trial might have been much less than what was reported – which could have led to a considerable delay in access to a drug that revolutionised therapy and outcomes for women with an aggressive form of breast cancer. Next time, the clinical community might not be so lucky in an unplanned subgroup analysis of the *post-hoc*-defined appropriate subgroup of patients.

So how do we balance the desire to select only the appropriate target group of cancers with the need to be able to later refine the diagnostic test and accommodate interlaboratory variations in assays? Perhaps the answer is that we need to rethink the statistical designs. The concept that the only valid conclusion from a trial comes from the primary analysis of the intent-to-treat (ITT) population may have served us well when *post-hoc* elimination of patients from the analysis population was seen to risk introducing bias. We risk, however, either increasing the chance of a false-negative trial by sticking to the primary ITT analysis, simply because we did not allow for scientific progress during the life of the trial (as would have happened had we not known how to test, even crudely, for trastuzumab sensitivity in the 1990s), or underpowering the key analyses because the true sensitive population has only poor overlap with those patients actually enrolled.

Could we not consider trials with sufficient power to allow two primary analyses? One analysis based on the original trial population, and a second using a different diagnostic to be applied post randomisation but just before the primary analysis? In parallel with the parent, phase III practice-changing trial, therefore, a programme of smaller, targeted phase II and/or quality assurance studies designed to refine

the diagnostic test for sensitivity could be conducted. This would perhaps be a better strategy than conducting a large and expensive phase III trial with a higher risk of being negative if the entry diagnostic was wrong, or delaying the drug development until such point as the diagnostic was better defined, costing not only money but also patient lives?

So I believe we still need phase III trials to develop new agents in an era of stratified medicine.

How does this approach deliver individualised therapeutic strategies? In simple terms, no prior trial can answer the question as to the best therapy for a particular patient, since each tumour/patient combination is essentially unique. Short of studies that can sequentially test different therapies in one patient [6], deductions need to be derived from studies in selected populations that most closely match the patient in the clinic. By designing future trials in selected patient groups with spare power to render additional analyses more robust, however, one might be able to allow data from an earlier era, with different biomarker availability, to be re-analysed in the light of new developments. The increasingly routine practise of storing biological samples from patients in trials with full consent for later analyses is an important way of securing the ability to perform such analyses.

Clinical trials are with us for the foreseeable future. Their design, however, will need to accommodate analyses of interaction effects between biological variables and the intervention under study, and will need to have the appropriate statistical power to do this in the light of new knowledge appearing during the lifetime of the trial.

Competing interests

The author declares that they have no competing interests.

Acknowledgement

This article has been published as part of *Breast Cancer Research* Volume 11 Suppl 3 2009: Controversies in Breast Cancer 2009. The full contents of the supplement are available online at <http://breast-cancer-research.com/supplements/11/S3>.

References

1. Olivetto IA, Bajdik CD, Plenderleith IH, Coppin CM, Gelmon KA, Jackson SM, Ragaz J, Wilson KS, Worth A: **Adjuvant systemic therapy and survival after breast cancer [see comment]**. *N Engl J Med* 1994, **330**:805-810.
2. Andre F, Slimane K, Bachelot T, Dunant T, Namer M, Barrelier A, Kabbaj O, Spano J-P, Marsiglia H, Rouzier R, Delalogue S, Spielman M: **Breast cancer with synchronous metastases: trends in survival during a 14 year period**. *J Clin Oncol* 2004, **22**:3302-3308.
3. Allred DC, Harvey JM, Berardo MD, Clark GM: **Prognostic and predictive factors in breast cancer by immunohistochemical analysis**. *Mod Pathol* 1998, **11**:155-168.
4. Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, Bajamonde A, Fleming T, Eiermann W, Wolter J, Pegram M, Baselga J, Norton L: **Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2 [see comment]**. *N Engl J Med* 2001, **344**:783-792.
5. Smith IE: **Efficacy and safety of Herceptin in women with metastatic breast cancer: results from pivotal clinical studies**. *Anti-Cancer Drugs* 2001, **12**(Suppl 4):S3-S10.

6. Thomlinson RH: **Measurement and management of carcinoma of the breast**. *Clin Radiol* 1982, **33**:481-493.