Cancer Chemother Pharmacol (1981) 7: 1–4

Cancer Chemotherapy and Pharmacology

Editorial Commentary

Predictors of Response and Their Clinical Evaluation

Stephen K. Carter

Director, Northern California Cancer Program, 1801 Page Mill Road, Bldg. B, Palo Alto, CA 94303, USA

The clinical evaluation of predictors of response will require a specific strategy which is only now in evolution. The experience that has been gained with the estrogen receptor (ER) assay and the stem cell assay will, hopefully, lay the groundwork for a more expeditious and efficient clinical evaluation of future potential predictors.

The evaluation of a predictive assay goes through two broad stages of study. The first stage involves retrospective correlation and the second involves prospective correlation. The initial clinical study of a potential predictive assay takes place within a single institution where patients are treated empirically and their response, or lack thereof, is correlated with the prediction made by the assay in question. In a small number of patients, the true-positive and true-negative rates are calculated to see whether larger-scale studies are indicated. This initial study could be described as the phase I evaluation. In this study, a preliminary estimate of cost-effectiveness is made and an initial determination of a cut-off point for positivity can be made.

If this initial study is encouraging, then a larger-scale retrospective evaluation (phase II) is the next step. This may involve data from many institutions or just wider-scale investigation within a smaller number. The design of the study is still to treat patients empirically, independently of the assay, and afterwards correlate the effect with the assay results. The end-point is still a calculation of the sensitivity and specificity, as demonstrated by the true- and false-positive and true- and false-negative rates. The economics and the reproducibility of the assay on a large scale, particularly in different institutions, becomes an important part of the evaluation. Another critically important aspect of this study is a final determination of what the cut-off point will be for positivity in the large scale prospective studies to come.

The phase III aspect of the evaluation becomes a prospective study of the value of the assay. In this kind of study, the assay is performed first and therapy is determined by the result. The clinical results of treating by the assay must be shown to be superior to the results of treating *a comparable group of patients* independently of the assays. This comparison will involve the standard efficacy end-points (response, duration of response, and survival) in relationship to the toxicity of therapy. If superiority is found, then the cost of the assay must be factored into the analysis to determine whether the increase in the therapeutic index, demonstrated in a selected clinical research population, will be cost-effective in large-scale general patient use.

The phase IV evaluation of the predictive assay will be a study of the assay in a large-scale population to see whether the results obtained in the clinical research setting can be translated in a cost-effective manner to the general population.

The prospective stage of the clinical evaluation has been the most neglected in terms of clinical trial methodology development. For example, a truly prospective evaluation of the ER assay has still to be performed, and only now are investigators beginning to think about how to prospectively evaluate the stem cell assay.

Estrogen Receptor Considerations

The ER assay for predicting response to hormonal manipulation in advanced breast cancer has been the subject of relatively rapid and reproducible phase I and II retrospective studies. The studies indicated that the true-positive rate was 60%–70% and that the true-negative rate was about 95%. This predictability was the same regardless of the type of hormonal therapy used and regardless of menopausal

0344-5704/81/0007/0001/$ 01.00

status. It is worth noting that even today, after years of usage, the exact cut-off point for an ER-'positive' assay is not consistent between laboratories and is developed by each laboratory through evaluation of the range of its results in a quality control sense. The literature of clinical research is made complicated by the fact that in some studies ≥ 10 fmol cytosol protein per mg is the cut-off for positivity, while in others ≥ 3 or ≥ 5 fmol is used. It must be assumed that there are some patients who would be deemed ER-positive in one study who would be labelled as ER-negative in another study. The use of the terms ER-rich and ER-poor, while closer to the biologic reality, also makes for a relative lack of precision in comparing clinical trial data.

The comparability of clinical trial data is made even more complex by other considerations, including the following: 1) Different laboratories use different laboratory techniques to measure the ER content in tumor; 2) Quality control checks within laboratories indicate a range of results when an identical sample is repetitively assayed. The error rate in calling a patient ER-positive or -negative, based on laboratory variance, has not been extensively reported in the literature but may be in the range of 5%–15%. This fact is largely ignored and its impact on general usage has barely been studied; 3) There is variability in the handling of the tissue samples as they are transported to the laboratory for assay; 4) The assay results on the primary lesion are not perfectly predictable for the assay results on a metastatic lesion from the same patient. When a metastatic lesion cannot be biopsied for ER assay, treatment decisions based on the results from the primary will be incorrect to a certain degree. The extent of this imprecision has also not been extensively studied but may be in the 15% range; 5) When metastatic lesions are repetitively assayed after various therapeutic interventions the results may change. This involves a certain number of ER-positive assays changing to ER-negative. The reverse apparently rarely occurs. Comparison of clinical trial data must differentiate between hormonal therapy that is the initial therapy for metastatic disease and hormonal therapy of metastatic that is second- or third-line.

What is the practical impact of the ER assay in the therapy of clinically evident metastatic breast cancer? It can be broken down as follows for initial treatment: 1) ER-negative patients, who would have been routinely treated with hormonal therapy, will now be given cytotoxic chemotherapy instead, and 2) ER-positive patients, who *might* have been treated with chemotherapy, would now be given hormonal therapy. Since the empiric initial therapy for advanced breast cancer has always been hormonal manipulation, the true impact of #2 is questionable and has never been analyzed in the literature of breast cancer treatment. Patients presenting with highly aggressive visceral disease, e.g., lymphangitic pulmonary metastases and hepatic metastases with liver dysfunction, would always go immediately to chemotherapy regardless of their ER assay, although most would probably be ER-negative.

The traditional triage of women with advanced breast cancer prior to the use of the ER assay is shown in Fig. 1. The triage according to the results of the ER assay, the system in current use, is shown in Fig. 2.

While the ER assay is now commonly used in a prospective manner, the proper trial to absolutely prove its value still remains to be performed. This trial would involve women with first recurrence after mastectomy and would be restricted to those in whom an assay result was available for decision-making. These women would be randomized to be treated in one of two ways: 1) Following the results of the assay, or 2) as if the assay results were not available. The critical final end-point would be the overall survival of the two groups. Subanalysis would involve the response rate to initial therapy, the duration of the initial response and the toxic cost. It would be important to study the response rate to second-line therapy and its duration as well. What is often forgotten in the analysis of advanced breast cancer trials is that the therapy is a sequence of treatments

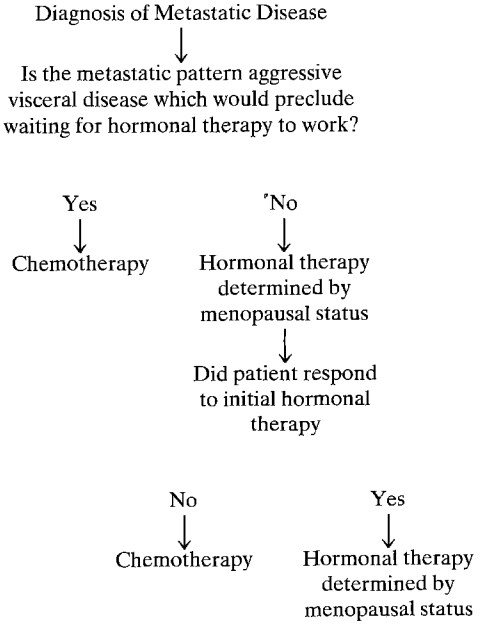


Fig. 1. Traditional triage of women with advanced breast cancer

with hormones and drugs. The success of this therapeutic sequence is measured in terms of the overall survival, the diagnosis of clinically evident metastatic diseases being used as the zero point. Most breast cancer trials test a therapy, or compare two or more therapies, within the heterogeneity of the varying sequences which can be utilized. The ultimate value for the ER assay has to be shown within the context of the entire therapeutic continuum. Within this continuum, does the use of ER assay lengthen survival and/or diminish treatment-related morbidity? Data to answer this question currently do not exist. On the surface, it would seem logical that the ER assay should improve survival and/or diminish treatment-related toxicity. It should be remembered, however, that the major impact of the ER on treatment decision-making may be to highlight a subset of women who should *not* be treated initially with hormones, but with chemotherapy instead. If this is true, it is unlikely that the ER assay will diminish treatment-related side-effects, and it remains to be demonstrated that this will improve survival.

The design of a prospective trial to test the value of the ER assay is shown in Fig. 3. This trial is designed for the first treatment of clinically evident metastatic disease, which is where the ER assay should make its major impact. Only a trial that will encompass all patients initially presenting with metastatic disease will be able to answer the question clearly within the context of the overall treatment continuum. The trial would elucidate a common initial chemotherapy regimen to be used in all patients in whom cytotoxic drugs were to be the therapy. Hormone therapy would be fixed in the same way for each of the menopausal subsets.

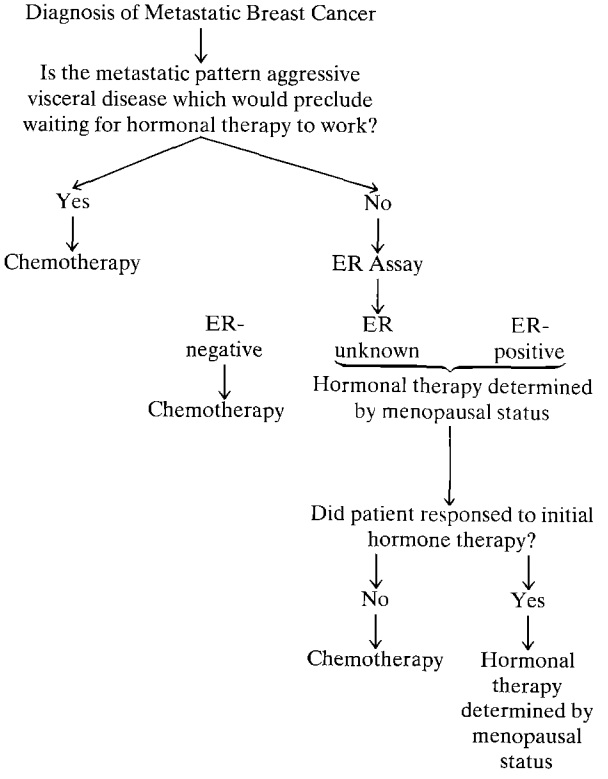
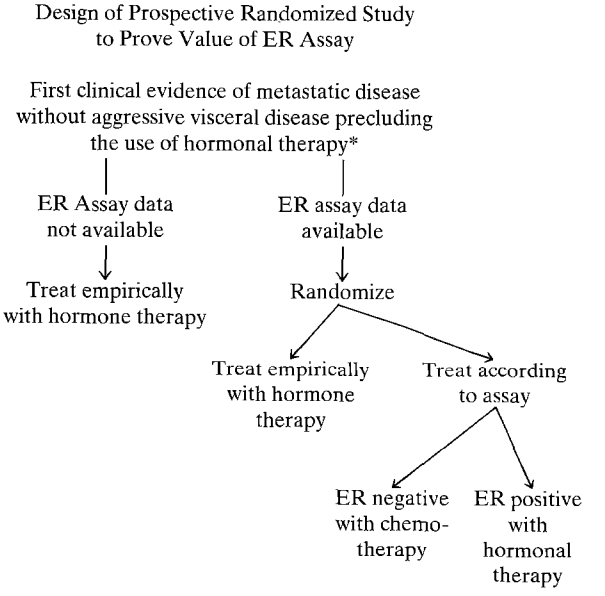


Fig. 2. Triage of women with advanced breast cancer: system in current use as a result of the ER assay

Clonogenic Assay Considerations

The clonogenic assay is a new and exciting development in clinical and basic research. The major clinical question that has to be answered about this approach is how well it can actually predict for response to chemotherapy both with single agents and with combinations. The current clinical data base concerning this question is totally retrospective in nature. A series of assays have been performed and retrospectively compared with what has happened in the clinic. The results have been encouraging, indicating a positive predictive ability of about 65% and a negative predictive ability in excess of 90%. These results, however, must be validated by prospective studies, which will be complex and difficult to perform.

On the surface, a prospective study of the clonogenic assay would appear to be easy to perform.



* Patients with visceral disease treated with chemotherapy

Fig. 3. Design of a prospective trial to test the value of the ER assay

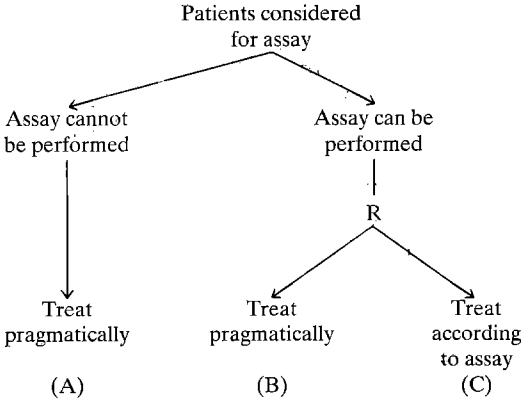
In a series of patients, the assay could be performed and the results analyzed. The patients would then be treated according to the interpretation of the assay and the response, or the lack thereof, correlated. This approach, however, does not get at the critical question of whether therapy, according to the assay, is superior to standard pragmatic therapy based on past clinical responsiveness to the regimens in question. The clonogenic assay, if truly cost-effective, should give superior results as measured by response rate, toxicity, response duration, or survival. Measurement of this will require a randomized study comparing therapy according to assay with pragmatic therapy ignoring the assay.

Again, on the surface, such a trial would be simple. One design would be to treat patients according to the assay, if the assay can be performed. The results would be compared with the pragmatic data obtained in patients who could not have their tumors assayed. This study would suffer from the criticism that those patients whose tumors could not be assayed were a different group, prognostically, from those patients in whom the assay could be successfully interpreted.

An ideal experimental design for a prospective clinical trial of the clonogenic assay is shown in Fig. 4. In this design, patients in whom the assay can be performed are randomized to be treated either empirically (B) or according to the assay (C). Those patients in whom the assay cannot be performed would also be treated empirically (A).

A specified regimen would be used in all patients to be treated empirically. In ovarian cancer it might be something like the three-drug PAC regimen (cisplatin, adriamycin and cyclophosphamide). The end-point of this trial would be to see whether patients treated according to the assay had a higher response rate and a longer survival under treatment. This would be most convincing in a B vs C analysis. It would be both interesting and instructive if B vs A was different and if a dichotomy was found in the comparison of B vs C and C vs A. Within the group treated by the assay would be a subset found sensitive to one or more agents tested (C_1) and a subset found resistant to all agents tested (C_2). This latter subset (C_2) could also be treated with the standard empirical regimen chosen for the trial. Another interesting comparison would be C_2 vs B and C_2 vs A.

Design for a Prospective Clinical Trial of the Clonogenic Assay



Compare B vs C*
Compare A vs C**
Compare A vs B**

* Most meaningful
** Subjected to the vagaries of potential lack of comparability

Fig. 4. An ideal experimental design for a prospective clinical trial of the clonogenic assay

Conclusion

Predictors of response, such as the ER assay and clonogenic assay, require a comprehensive clinical trial strategy. While their potential is exciting, their premature or inappropriate usage will drive up health care costs without the requisite benefit being accomplished. These assays must be demonstrated to be effective in properly designed, executed, and analyzed *prospective* clinical trials before they are recommended to the general population. Unfortunately, both assays have been viewed as being successful on the basis of retrospective studies. A positive retrospective study should only be viewed as a justification for the cost and effort of a large-scale well-designed prospective study. A cost-benefit analysis must be made to elucidate which subsets of patients have their therapy actually changed as a result of the assay and whether this subset actually has a better end-result which justifies the expense to the entire population. Such an analysis may also highlight subsets who do *not* need the assay, since it would not change their therapy. In recent years, cancer treatment has been plagued by premature claims of success for new therapies based on premature analysis of preliminary data. We should not make the same mistake with predictors of response.