Proliferation and Apoptosis as Markers of Benefit in Neoadjuvant Endocrine Therapy of Breast Cancer

Mitch Dowsett,1 Ian E. Smith,2 Steve R. Ebbs,3 J. Michael Dixon,4 Anthony Skene,5 Clive Griffith,6 Irene Boeddinghaus,1,2 Janine Salter,1 Simone Detre,1 Margaret Hills,1 Susan Ashley,2 Stephen Francis,7 Geraldine Walsh,2 and Roger A’Hern1

Abstract The study of changes in proliferation as a marker of treatment benefit during presurgical endocrine treatment of breast cancer has become increasingly popular, particularly using the nuclear marker Ki67, and holds the potential for prioritizing new treatments for full clinical development. There are weakly significant relationships between Ki67 change and clinical response that differ according to data handling. In the neoadjuvant Immediate Preoperative Anastrozole, Tamoxifen, or Combined with Tamoxifen trial, suppression of Ki67 at both 2 and 12 weeks was greater with the aromatase inhibitor anastrozole than with either tamoxifen or the combination of anastrozole and tamoxifen. We report here that absolute values of Ki67 after 2 weeks were also significantly lower with anastrozole than with tamoxifen and the combination. This indicates that it may be possible to make such comparisons using surgical samples only. We argue that these changes in proliferation and concurrent changes in apoptosis may be expected to be more predictive of adjuvant benefit from endocrine therapy than clinical response.

At present, the assessment of new treatments in early breast cancer requires the conduct of very large randomized adjuvant clinical trials involving thousands of patients, with follow-up extending over several years before first results emerge. Reliable intermediate end points that allow more rapid evaluation would be of enormous value. The neoadjuvant (preoperative) setting in which medical therapies may be evaluated before surgery is being increasingly exploited with this in mind.

The Immediate Preoperative Anastrozole, Tamoxifen, or Combined with Tamoxifen (IMPACT) trial was a randomized neoadjuvant comparison of the aromatase inhibitor anastrozole versus tamoxifen versus the combination in 330 postmenopausal women with estrogen receptor–positive primary breast cancer (1,2). A major goal of the trial was to determine whether either clinical or biological response predicted the outcome of the much larger (9,336 patients) and longer adjuvant comparison of the same three treatments in the Arimidex, Tamoxifen, Alone or in Combination (ATAC) trial (3). Overall, there was no significant difference in the clinical response rates between the arms although patients who were not deemed eligible for breast-conserving surgery before treatment had a greater chance of having such an operation after treatment if they received anastrozole.

Core biopsies were taken before and after 2 weeks of treatment and samples of the excision specimen were used to assess any change in proliferation as measured by the nuclear antigen Ki67. This analysis, which was the prespecified primary biomarker end point of the trial, revealed a highly significant greater reduction in Ki67 in patients treated with anastrozole than was seen in either of the other arms (2). Importantly, the result was parallel to, and could be considered predictive of, the greater benefit found with anastrozole in the ATAC adjuvant trial of these three treatments (3). This was despite the absence of an overall significant relationship between percentage decrease in Ki67 at either 2 or 12 weeks and clinical response when the latter was expressed as a conventional categorical division of responders versus nonresponders (2).

This article reports on further exploratory analysis of these data to determine if closer relationships with clinical response might exist with different statistical handling of the Ki67 data and with the changes found in apoptosis, the other primary determinant of overall tumor dynamics.

Methodology

The trial design and the methods for obtaining tissue and its staining for Ki67 and apoptosis have previously been described (2). In summary, this was a randomized, double-blind, double-dummy, multicenter trial in which patients with primary breast cancer were randomized 1:1:1 to receive a preoperative daily
dose of anastrozole (1 mg) plus tamoxifen placebo, or tamoxifen (20 mg) plus anastrozole placebo, or anastrozole (1 mg) plus tamoxifen (20 mg) for 12 weeks before surgery. Eligible patients were postmenopausal women with previously untreated, core needle biopsy-proven, invasive, estrogen receptor–positive breast cancer. Tumors were operable or locally advanced but potentially operable (after medical downstaging) and without evidence of metastatic spread. Any women receiving hormone replacement therapy stopped this before the trial. To be evaluable for the biomarker end points, patients had to have ceased such therapy at least 4 weeks before the start of treatment.

Biopsies were stained for Ki67 using the MIB-1 antibody (DAKO, Glostrup, Denmark) and for apoptosis using the terminal deoxynucleotidyl transferase biotin-dUTP nick end labeling technique.

Objective clinical response was calculated as the percentage of patients with a clinical complete response or partial response at 3 months. Complete response was defined as the clinical disappearance of tumor maintained until the 3-month point, and partial response was defined according to WHO criteria as a ≥50% decrease from baseline in the product of two perpendicular diameters maintained until 3 months. Minor response was defined as a decrease of 25% to 50% in the area of the tumor from baseline, and no change was defined as a decrease of <25% or an increase of <25% in the area of the tumor from baseline (both of these groups were included in the category of stable disease). Progressive disease was defined as an increase of ≥25% in the area of the tumor from baseline.

Core-cut biopsies were taken before starting therapy and at 2 weeks (nonobligatory). Patients not proceeding to surgery for whatever reason were invited to have a further core biopsy at 12 weeks. Core biopsies and the excision biopsy were fixed in 10% neutral-buffered formalin for 24 to 28 hours before processing and embedding at local pathology centers in paraffin wax blocks.

The study was done in accordance with the ethical principles of the Declaration of Helsinki. The method was approved by an institutional review board and patients gave their written informed consent to participate. The statistical analyses have been previously described and were based largely on ANOVAs with a two-sided 5% significance level for within and between treatment comparisons (4).

Results

Figure 1 shows the individual on-treatment/pretreatment ratios of Ki67 at 2 and 12 weeks on a logarithmic scale according to response expressed as five categories. There were few patients in the categories of complete clinical response and progressive disease; in the other three categories, there was a wide distribution of Ki67 suppression. Spearman rank correlation showed that this relationship was statistically significant at 2 weeks but not at 12 weeks. None of the individual treatment groups showed a significant relationship. There was a weakly significant relationship between percent tumor shrinkage and change in Ki67 in the overall comparison (Spearman $r = 0.181$, $P = 0.03$) but not for the individual treatment arms.

**Fig. 1.** Change in Ki67 expressed as on-treatment/pretreatment ratio at 2 weeks (A) and 12 weeks (B) for individual patients according to response category. PD, progressive disease; NC, no change; MR, minimal response; PR, partial response; CR, complete response.
Figure 2 shows the values of Ki67 at 2 weeks in the same five categories. Again, a wide distribution is apparent in each category, and although there were trends for a lower Ki67 to be associated with greater response, this was not significant at either 2 or 12 weeks.

Comparison of the on-treatment values among the three arms did, however, reveal significantly lower values of Ki67 with anastrozole than with either tamoxifen at both 2 and 12 weeks \((P = 0.013\) and \(0.0006\), respectively) or the combination \((P = 0.039\) and \(0.0003\), respectively; Fig. 3). There were no significant differences between tamoxifen and the combination at either time point. The study was not designed to assess differences between anastrozole and the combination therapy; \(P\) values for these comparisons are shown in Fig. 3 but it should be recognized that these are exploratory analyses.

The correlation between Ki67 and apoptosis before treatment and at 2 weeks is shown for all patients in Fig. 4 on logarithmic scales to allow the lower values to be visually separated. As previously reported, there is a positive correlation between Ki67 and apoptosis at baseline. It can be seen here that this positive relationship is maintained but is a little weaker after 2 weeks of treatment. Whereas the decrease in Ki67 between the two time points is obvious by comparison of Fig. 4A and B, there was no major change in apoptosis. These overall data do not allow the previously reported significant decrease in apoptosis in the anastrozole-treated patients after 2 weeks to be appreciated; this is therefore shown separately (Fig. 5). Surprisingly, the 2-week change in apoptosis was positively correlated with the change in tumor size (i.e., the greater the decrease in apoptosis, the greater the tumor shrinkage; \(r = 0.267\), \(P < 0.002\)). However, among the individual groups, this was significant only for the anastrozole-treated patients \((r = 0.365\), \(P = 0.009\)).

**Discussion**

The expanding number of new agents with a potential for therapeutic benefit in breast cancer (and other cancers) requires an improved approach to clinical testing and development. Continued application of the current paradigm of phase I-III...
testing, with the final stage taking many years, will lead to an accumulation of untested or partially tested compounds and will not allow recent advances in molecular understanding of breast cancer to be exploited. Neoadjuvant and short-term presurgical treatments have begun to be used as test beds for new compounds that allow their pharmacologic and therapeutic efficacy to be assessed. For studies in this setting to be of greatest value, reliable biomarkers of clinical efficacy must be identified and validated.

The greater reduction in Ki67 levels by anastrozole than either tamoxifen or combination after 2 weeks in the IMPACT study was therefore particularly instructive in its prediction of the better recurrence-free survival with anastrozole than either tamoxifen or the combination in the ATAC trial (3). Although IMPACT reported after ATAC (5, 6), it had been designed beforehand to test the ability of the endocrine neoadjuvant scenario to predict clinical benefit in an analogous adjuvant context; it had been initiated years before the ATAC trial result and remained double-blind with an established statistical plan that was unaffected by the ATAC result. The result after only 2 weeks of treatment had major implications because it suggests that such comparisons may be made with validity during the approximate 2-week gap between diagnosis and surgery. Treatment during this period requires established safety of an agent but not proven efficacy.

Although it is highly unlikely that such testing could ever replace conventional phase III testing, in particular because of the need for long-term safety data, the use of this approach to prioritize treatments for full clinical development is attractive. The ATAC/IMPACT trials are a useful illustration of how this could work. Had anastrozole and tamoxifen been under consideration for full clinical development at the same time, the 2-week time point of IMPACT would have allowed the selection of anastrozole over tamoxifen as the more effective agent. Additionally, had IMPACT reported before ATAC, it seems unlikely that the combination arm would have been included in ATAC.
We have noted previously that although the changes in Ki67 in IMPACT predicted the improved recurrence-free survival with anastrozole above tamoxifen and the combination in ATAC, there was no overall improvement in clinical response with anastrozole and change in Ki67 was not a strong predictor of clinical response in the IMPACT trial itself (2). However, the detailed analyses shown above indicate that exploration of the data does reveal significant associations of Ki67 with response.

Many centers have an interest in applying the presurgical model for assessment of new agents but find it difficult to obtain a pretreatment sample for research. Some centers have found it possible to request prediagnostic consent for extra biopsies for research. These are available to replace the diagnostic biopsy if the latter is found to be insufficient for diagnosis. It would be advantageous, however, if the on-treatment surgical biopsy contained the important biological information that allowed significant distinction between more and less efficacious treatments to be made. The data presented (Fig. 3) indicate that this may indeed be possible: the absolute level of Ki67 at 2 weeks showed the same differences between anastrozole, tamoxifen, and the combination as the percentage decrease in Ki67 (2). The statistical significance was, however, weaker and this indicates that a larger number of patients would be required to achieve the same statistical power using such an approach. Another disadvantage of using only surgical samples for comparison is that pretreatment determinants of response cannot be assessed (4).

A complication in the use of proliferation indices as a single marker of response/benefit is the undoubted contribution of apoptosis to tumor growth dynamics. We and others have previously reported the significant relationship which exists between Ki67 and apoptosis in breast carcinomas (2, 7). In this report, we have shown that the relationship is maintained but is weaker after 2 weeks of treatment. It was a surprise in these presurgical studies to find that apoptosis was in fact significantly reduced by aromatase inhibition, a consistent finding between this study with anastrozole and our previous smaller study with vorozole (8). The significant relationship between the reduction in apoptosis and tumor shrinkage in the anastrozole arm is superficially particularly surprising. However, it is possible that the reduction in cell death is a reaction to the profound reduction in proliferation, and the relationship with tumor shrinkage may reflect this.

These data suggest that shrinkage of the tumors in response to endocrine therapy occurs as a result of markedly reduced proliferation alongside the largely maintained rate of apoptosis (Figs. 3–5). Given that tumors with a higher proliferation rate pretreatment are associated with a generally higher apoptosis rate (Fig. 4A), on average a decrease from a high Ki67 level (e.g., 40%) to a lower value (e.g., 10%) would be associated in the on-treatment sample with a higher apoptosis level than a tumor that has decreased to the same low Ki67 level from a moderate level (e.g., 20%). This implies that the tumor growth dynamics of two tumors with similar Ki67 levels on treatment may be quite different depending on their initial proliferation/apoptosis balance.

A particularly striking result is the very wide variability in the decrease in Ki67 that is present in the different response categories; although there is a significant relationship with response, this is very modest in strength. Imprecision in the measurement of clinical response is well known and is likely to contribute to the weakness of this relationship but is very unlikely to be the major contributor. Full consideration of the contribution made by reduced proliferation to clinical response in fact suggests that a close relationship should not be expected. Reduced proliferation may be considered of benefit to a patient and, in a wider context, be described as a response; however, for an objective clinical response to be recorded in a fast-growing tumor, a major reduction in proliferation to reverse that growth may be required. In contrast, any reduction in proliferation would be expected to contribute to an improved recurrence-free survival in the adjuvant setting. These relationships are illustrated in Fig. 6.

Ki67 is a convenient and readily measured marker of proliferation but it is possible that enhanced relationships with response and more precise indices of benefit from treatment may be identified by different approaches. To assess this, we have published in abstract form gene expression microarray analyses of 17,500 genes measured before and after 2 weeks of either letrozole or anastrozole in 35 postmenopausal primary breast cancer patients (9). Over 2,000 genes decreased or increased significantly (false discovery rate of 0.01) over this period. This approach would potentially allow the identification of both a global index of dependence on estrogen and a proliferation index that is composed of multiple genes. This approach also has the advantage of reducing the impact of inaccuracies on the measurement of any one analyte. In addition, the performance of these studies on homogenized core biopsies results in the data being a reflection of gene expression in millions of cells rather than the 1,000 counted in our analyses of Ki67, although it does have the disadvantage of including any stromal cells in the readout.

This type of approach should also allow the identification of gene changes that may facilitate the rational combination of endocrine and other therapy. Whereas such combination treatment is generally considered in relation to the pretreatment phenotype of a tumor, it is possible that the endocrine treatment–induced phenotype may be more relevant. Thus, short-term presurgical therapy may reveal markers of sensitivity or resistance to particular therapies that would not otherwise been apparent.

**Open Discussion**

**Dr. Carlos Arteaga:** The lesson I learned from the IMPACT trial is that it may have allowed us to predict one “loser” arm that may not have been required in the ATAC trial. Eliminating one arm, in this case the combination arm, would have spared the enrollment of 3,000 patients and potentially a year for the completion of the study. So, are there any big randomized studies out there that need their own IMPACT counterpart?

**Dr. Dowsett:** Matt Ellis’ 0.24 study [J Clin Oncol 2001;19:3808–16] predicted BIG 1-98 but there were only two arms there, letrozole and tamoxifen. I think that is the only other study.

**Dr. Aman Buzdar:** In IMPACT, if you would have taken the clinical response data on face value, you would have thought that the combination would do fantastic. From hindsight, you would have picked the wrong therapy.
**Dr. Arteaga:** I don’t think this type of study is going to tell me much about clinical response. My thinking is rather: Can I use this to select the loser combination? Can I use it as a prioritization system for clinical trial design?

**Dr. James Ingle:** In Matt Ellis’ study, you would have predicted the loser in terms of clinical response. I agree with Dr. Buzdar regarding IMPACT, you would have looked at the clinical response and concluded that the combination might be of value. Now, the issue is, how do you validate the Ki67?

**Dr. Dowsett:** I think it depends on what you are validating it for. I wouldn’t want to give the impression that I am suggesting this is ever going to be a tool for predicting an individual patient’s response. I think it is very much a useful tool for doing analyses of new agents, for working out the mechanisms, and for identifying those patients who are not going to benefit from those treatments.

**Dr. Ingle:** But you would like to see a validation of your Ki67?

**Dr. Dowsett:** I think this is as close to a validation as one can get at the moment. It is actually related to the adjuvant outcome. It is not related to the clinical response.

**Dr. Steven Come:** Doesn’t that trouble you, though, that it didn’t hold up? It correlated with response to tamoxifen in 2 weeks but not at 12 weeks. Isn’t the early decrease in Ki67 in 80% to 85% of the patients reflecting the fact that they are estrogen receptor positive and they respond to taking estrogen away? Shouldn’t the focus be, instead, on global estrogen sensitivity, maybe screening for an up-regulation of resistance mechanisms that might separate responders and nonresponders?

**Dr. Dowsett:** We are now investigating this global index of estrogen dependence, which is a better marker than Ki67. Once we’ve got that, we will work on its pretreatment prediction. The drug mechanisms will determine whether or not the index of estrogen deprivation can be affected. This is a tool to work out where the resistance mechanisms occur.

**Dr. Myles Brown:** In estrogen receptor–positive, well-differentiated tumors, the presurgical response in proliferation and apoptosis may be particularly misleading because only a minority of the cells are actually the clonogenic cells driving the tumor growth. It raises the possibility that with Ki67, you are looking at growth and survival of cells that are irrelevant to the ultimate behavior of the tumor. You may be better off with something like genotyping or estrogen dependence as a marker because that may be a better reflection of what the clonogenic cell is.

---

**Fig. 6.** Diagrammatic representation of how a decrease in Ki67 would be expected to be associated with reduced tumor growth, yet be associated with either progression (A) or regression (B) according to initial tumor growth dynamics. In both cases, the slowing of growth would be expected to improve outcome as reflected in recurrence-free survival (RFS) although the faster-growing tumor would have a poorer prognosis.

---

45x304]Dr. Arteaga: I don’t think this type of study is going to tell me much about clinical response. My thinking is rather: Can I use this to select the loser combination? Can I use it as a prioritization system for clinical trial design?

Dr. James Ingle: In Matt Ellis’ study, you would have predicted the loser in terms of clinical response. I agree with Dr. Buzdar regarding IMPACT, you would have looked at the clinical response and concluded that the combination might be of value. Now, the issue is, how do you validate the Ki67?

Dr. Dowsett: I think it depends on what you are validating it for. I wouldn’t want to give the impression that I am suggesting this is ever going to be a tool for predicting an individual patient’s response. I think it is very much a useful tool for doing analyses of new agents, for working out the mechanisms, and for identifying those patients who are not going to benefit from those treatments.

Dr. Ingle: But you would like to see a validation of your Ki67?

Dr. Dowsett: I think this is as close to a validation as one can get at the moment. It is actually related to the adjuvant outcome. It is not related to the clinical response.

Dr. Steven Come: Doesn’t that trouble you, though, that it didn’t hold up? It correlated with response to tamoxifen in 2 weeks but not at 12 weeks. Isn’t the early decrease in Ki67 in 80% to 85% of the patients reflecting the fact that they are estrogen receptor positive and they respond to taking estrogen away? Shouldn’t the focus be, instead, on global estrogen sensitivity, maybe screening for an up-regulation of resistance mechanisms that might separate responders and nonresponders?

Dr. Dowsett: We are now investigating this global index of estrogen dependence, which is a better marker than Ki67. Once we’ve got that, we will work on its pretreatment prediction. The drug mechanisms will determine whether or not the index of estrogen deprivation can be affected. This is a tool to work out where the resistance mechanisms occur.

Dr. Myles Brown: In estrogen receptor–positive, well-differentiated tumors, the presurgical response in proliferation and apoptosis may be particularly misleading because only a minority of the cells are actually the clonogenic cells driving the tumor growth. It raises the possibility that with Ki67, you are looking at growth and survival of cells that are irrelevant to the ultimate behavior of the tumor. You may be better off with something like genotyping or estrogen dependence as a marker because that may be a better reflection of what the clonogenic cell is.
like, rather than the survival and growth of the nonclonogenic cells.

**Dr. Dowsett:** In terms of cytotoxic chemotherapy, I think that the rationale is very tight. In terms of endocrine therapy, it looks as though the mechanism of action is largely antiproliferative. So why not measure the mechanism?

**Dr. Buzdar:** On the clinical side, last year I was convinced that maybe there was something to Ki67. Checking it after a couple of weeks on therapy might help us to identify response versus nonresponse. Today, either I am totally confused or there isn’t any real correlation with outcomes.

**Dr. Dowsett:** The correlation with clinical outcomes is very modest. But I presented the arguments as to why you would expect it to be poor, looking at apoptosis; looking at the fact that if you have a tumor growing, you can slow the growth. We have statistics that are very strong. What we need is actually a 2-week preoperative study in perhaps 3,000 patients to look at the changes in some of the other markers and see how that relates to long-term outcome.

**Dr. Ingle:** Why is it 2 weeks? We’ve been looking at early response and late response genes. The late response genes are up by 24 hours. Why aren’t you looking at 2 days?

**Dr. Dowsett:** Two weeks has been the average time between diagnosis and surgery in the United Kingdom and a few other countries. We are trying to take advantage of that normal time frame without manipulating it.

**Dr. Ingle:** I’d feel better if you convinced us that 2 weeks was a good time to look. It may be that it’s 2 or 3 days where you are going to get some value.

**Dr. Dowsett:** Tamoxifen has a 10-day half-life; anastrozole and letrozole have a half-life of 2 days. So they reach steady state at 10 days.

**Dr. Brown:** So 2 weeks might be too short. You want to make sure that you have equal dose response in terms of total saturation. I don’t know if you get to saturation with tamoxifen in that time period in postmenopausal women.

**Dr. Ingle:** Whatever rationale you use, that should determine when you sample rather than a convenience sample. So much effort has gone into this research that it would make sense to see some studies done. When is the proper time to check? Is it a function of what drug you are looking at?

**Dr. Arteaga:** That’s a problem because here you don’t have a control for each patient. You don’t know the optimal timing for the individual patient. We could do it with xenografts to determine, for example, whether the data we get at 5, 10, or 15 days are the same or not.

### Appendix A. The IMPACT Trialists Group

W.H. Allum, S. Ashley, A. Bradley, I. Boeddinghaus, D. Brett, G. Buí, J. Diggins, J. Holborn, A. Ring, N. Sacks, C. Shannon, I. Smith, G. Walsh (Royal Marsden Hospital, London, United Kingdom); S. Detre, M. Dowsett, M. Hills, J. Salter (Royal Marsden Laboratory, London, United Kingdom); S. Ebbs, J. Kember, C. Chu (Mayday University Hospital, Croydon, United Kingdom); I. Batty, K. Kazim, A. Skene (Royal Bournemouth Hospital, Bournemouth, United Kingdom); J.M. Dixon, J. Murray, L. Renshaw (Western General Hospital, Edinburgh, United Kingdom); F. McNeill, K. Rooko (Essex County Hospital, Essex, United Kingdom); C. Griffith, J. Bevington (Royal Victoria Infirmary, Newcastle, United Kingdom); A. Evans, M. Pigdley (Poole General Hospital, Poole, United Kingdom); J-U. Blohmer, W. Lichtenegger, Universitätsklinikum Charité, Berlin, Germany; P. Sauven, K. Rooko (Chelmsford and Essex Centre, Chelmsford, United Kingdom); C. Holcombe, K. Makinson (Royal Liverpool University Hospital, Liverpool, United Kingdom); L. Barr, N.J. Bundred, T. Pritchard (University Hospital of South Manchester, Manchester, United Kingdom); N. Harbeck (Frauenklinik der TU München, München, Germany); J. Clarke, J. Mansi (St. George’s Hospital, London, United Kingdom); H. Stehle (Marienhospital, Stuttgart, Germany); T. Reimer (Universitäts-Frauenklinik, Rostock, Germany, K. Brunnett, Zentrum für Senologie und Plastische Chirurgie, Osnabrück, Germany); M. Lanskjow, J. Hepper (St. James’s University Hospital, Leeds, United Kingdom); D. Dubois, H. Stansby (Portsmouth Oncology Centre, Portsmouth, United Kingdom); and Z. Rayter (Bristol Royal Infirmary, Bristol, United Kingdom).

### References


Proliferation and Apoptosis as Markers of Benefit in Neoadjuvant Endocrine Therapy of Breast Cancer


Updated version
Access the most recent version of this article at:
http://clincancerres.aacrjournals.org/content/12/3/1024s

Cited articles
This article cites 7 articles, 2 of which you can access for free at:
http://clincancerres.aacrjournals.org/content/12/3/1024s.full.html#ref-list-1

Citing articles
This article has been cited by 6 HighWire-hosted articles. Access the articles at:
/content/12/3/1024s.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.