



UvA-DARE (Digital Academic Repository)

Prognostic factors in breast cancer: one fits all?

Mook, S.

Publication date
2011

[Link to publication](#)

Citation for published version (APA):

Mook, S. (2011). *Prognostic factors in breast cancer: one fits all?* [Thesis, fully internal, Universiteit van Amsterdam].

General rights

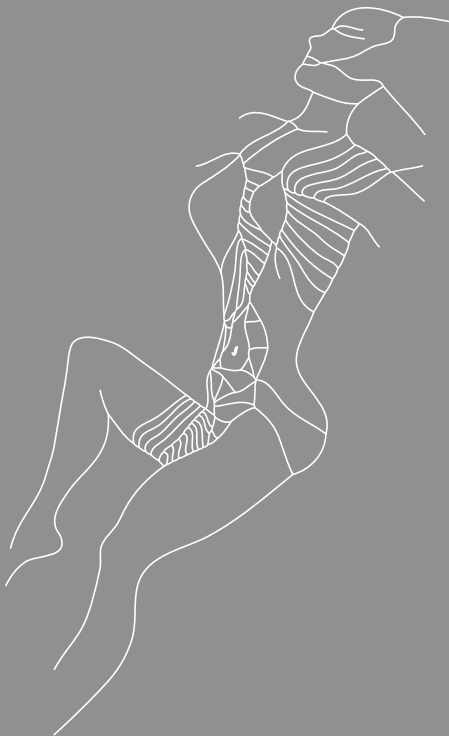
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

Appendix

Gene signature evaluation as a prognostic tool: challenges in the design of the MINDACT trial



Jan Bogaerts
Fatima Cardoso
Marc Buyse
Sofia Braga
Sherene Loi
Jillian A Harrison
Jacques Bines
Stella Mook
Nuria Decker
Peter Ravdin
Patrick Therasse
Emiel Rutgers
Laura J Van 't Veer
Martine Piccart
on behalf of the TRANSBIG consortium

Nat Clin Pract Oncol 2006; 3: 540-551.

Summary

This Review describes the work conducted by the TRANSBIG consortium in the development of the MINDACT (Microarray In Node negative Disease may Avoid ChemoTherapy) trial. The goal of the trial is to provide definitive evidence regarding the clinical relevance of the 70-gene prognosis signature, and to assess the performance of this signature compared with that of traditional prognostic indicators for assigning adjuvant chemotherapy to patients with node-negative breast cancer. We outline the background work and the key questions in node-negative early-stage breast cancer, and then focus on the MINDACT trial design and statistical considerations. The challenges inherent in this trial in terms of logistics, implementation and interpretation of the results are also discussed. We hope that this article will trigger further discussion about the difficulties of setting up and analyzing trials aimed at establishing the worth of new methods for better selection of patients for cancer treatment.

Review criteria

A formal literature search for this review was not performed; this review includes a summary of the authors' own work and knowledge, which covers various fields relating to oncology and molecular biology.

Introduction

In the past 20 years, little progress has been made in identifying new prognostic markers that can assist oncologists in treatment decision-making relating to node-negative early-stage breast cancer. As a result, considerable differences exist worldwide in the selection of women who require adjuvant chemotherapy based on their risk of breast cancer recurrence. The breast cancer TNM (tumor–node–metastases) staging system is based on anatomical extent (*e.g.* the size and lymph-node status) of the tumor, but this classification gives little insight into breast cancer biology. Clinicians have long recognized the heterogeneity of human breast cancers, not only in terms of their diverse natural histories despite identical morphological features, but also in their variation in response to treatment.¹ These differences are also evident in the small (*i.e.* <2 cm), node-negative tumors that would generally be associated with a good clinical outcome. Attempts have been made to identify good and poor prognosis groups based on pathological features such as tumor grade, lymphatic invasion and S-phase fraction,² which might better reflect tumor biology. In recent years, numerous molecular prognostic and predictive markers in oncology have been reported (*Box 1*). These tumor markers have had little impact in routine clinical practice. Studies are often based on small, heterogeneous retrospective series that have not been reported in a rigorous enough fashion to provide sufficient information, particularly with regard to their methodology.³ Many follow-up studies have shown inconsistent data compared with original results, which has been attributed to a lack of statistical power, different patient populations, and technical limitations associated with such studies. There is also a paucity of well-designed, prospective assessments of the clinical value of these tumor markers. As a result, the value of many promising prognostic markers is still uncertain. We have yet to fully translate our increased understanding of breast cancer biology into improved outcome for those with this heterogeneous disease. The prognostic factors accepted by the NIH 2000 Consensus Development Conference on Adjuvant Therapy for Breast Cancer did not include any molecular markers relevant to breast cancer biology apart from the hormone receptors.⁴ The most recent St Gallen consensus panel (2005)⁵ established three risk categories: minimal, intermediate and high. Hormone receptors, tumor size, tumor grade and age remain key discriminating factors, and HER2 status, lymphatic or vascular invasion, or both in the primary tumor are new accepted prognostic factors. In the UK, the Nottingham Prognostic Index is commonly used to predict clinical outcome; this index is based on tumor size, tumor grade and lymph-node status, and has a key role in discriminating node-negative patients for whom chemotherapy should or should not be considered.⁶ All these consensus recommendations, however, have important limitations and, in this era of evidence-based medicine, it is not possible to reliably identify a group of women with excellent long-term clinical outcome. By using gene-expression profiling, the Netherlands Cancer Institute developed a 70-gene prognostic signature for node-negative breast cancer.⁷ The signature was developed

as a dichotomous risk classification for the endpoint of distant metastasis within 5 years. The same group performed a first validation of this gene signature on 295 patients and confirmed that it outperforms all the traditional clinical prognostic factors and clearly separates a group with an excellent prognosis at 10 years from a group with a high risk of recurrence before 5 years.⁸ Furthermore, when compared with current commonly used risk classifications (*i.e.* the St Gallen guidelines and the NIH consensus), the 70-gene signature not only predicted those women who would have needed chemotherapy (as demonstrated by the onset of distant metastases within 5 years), but also women who could have been spared adjuvant chemotherapy, as seen from their excellent long-term outcome. The authors concluded that the 70-gene signature is of great prognostic importance, and could outperform current risk classifications and therefore potentially spare some women from overtreatment. Since the publication of the initial results of the 70-gene prognostic signature, a number of authors have underlined critical issues in gene selection bias, error estimation, fragility of gene signatures and overoptimistic performance estimation in the early validation studies of this signature due to the model that was used. Nevertheless, the potential of this signature to spare 10–20% of women adjuvant chemotherapy without sacrificing long-term outcome and also to identify a group of women who require different targeted treatment strategies to overcome their ‘poor prognosis’ is too compelling to be ignored, especially given the current frequent occurrence of overtreatment (and undertreatment) of women diagnosed with early-stage breast cancer.

Independent multicenter validation of the 70-gene signature

In order to determine the 70-gene signature’s true potential and robustness, a second, independent validation of the initial findings was performed by the TRANSBIG research consortium on data retrospectively collected from five cancer centers in the UK, Sweden and France.^{9,10} This study used frozen archival tumor material from 302 node-negative patients aged 60 years or less, who were diagnosed up to and including 1998, and who, as was the case in the original Dutch cohort, had received only locoregional therapy. The median follow-up was 13.6 years. The frozen samples were sent to Amsterdam for gene-expression profiling using a custom-designed chip (MammaPrint®; Agendia, Amsterdam, The Netherlands) that assesses the mRNA expression of the 70 genes in triplicate. The researchers in Amsterdam were blinded to the clinical outcome data. An independent statistical office in Brussels determined the clinical risk groups and performed the validation analyses. Central pathology review of estrogen receptor (ER) status and grade was carried out in the European Institute of Oncology in Milan for almost 80% of samples. The first results of this validation study confirmed the previous findings, namely that the 70-gene signature was able to delineate a subgroup at significantly higher risk of distant metastases and death, out-performing the Nottingham Prognostic Index, the St Gallen

criteria and the Adjuvant! software,¹¹ which can calculate a 10-year survival probability based on the patient's age, tumor size, grade and ER status.¹² A recent evaluation of the Adjuvant! software found that known clinical prognostic factors were able to predict overall survival (OS), breast-cancer-specific survival, and event-free survival quite accurately in 4,083 patients diagnosed with breast cancer in British Columbia from 1989 to 1993, with the exception of very young patients diagnosed under the age of 35.¹³ This independent validation of the software reinforced its credibility as an accurate clinical tool to evaluate breast cancer prognosis, making the ability of the 70-gene signature to outperform this tool all the more notable.

The 70-gene signature remained a significant prognostic indicator of time to distant metastasis and OS even after adjustment for all clinicopathologic factors known to have prognostic value in this disease. The consortium decided that the low clinical risk group would consist of patients with a 10-year breast cancer survival probability of at least 88% if their tumors were 1% or greater positive for expression of ER using immunohistochemistry, and of at least 92% if they were not. These two cutoffs were chosen to reflect the fact that patients with ER-positive tumors now receive adjuvant endocrine therapy (with an estimated absolute 10-year benefit of about 4% overall), whereas patients in the validation series were all untreated regardless of their ER status. When adjusted for clinical risk based on 10-year survival probability using the Adjuvant! software, the gene-signature adjusted hazard ratios (*Box 1*) were 2.13 (95% CI 1.19–3.82) for time to distant metastasis, 2.66 (95% CI 1.46–4.84) for OS, and 1.36 (95% CI 0.91–2.03) for disease-free survival. Similar hazard ratios were found in Cox multivariate regression analysis. These results indicate that the gene signature adds independent prognostic information to that provided by a risk assessment based solely on clinicopathologic factors. Central pathology review of ER and tumor grade and an independent source verification of all data by external auditors give these findings significant strength. Furthermore, within each gene-signature risk group, the Kaplan–Meier estimates of 10-year OS were almost identical to the two clinical risk groups as assessed by the Adjuvant! software: patients classified as gene-signature low risk had 10-year survival rates of 88% and 89%, respectively, for low and high clinical risk as defined by Adjuvant!, while for patients classified as gene-signature high risk, the 10-year survival rate was 69% for both clinical risk groups.

The external and independent validation therefore confirmed the original findings that the gene signature added significant independent prognostic information to that produced by current clinicopathologic factors, and provided strong support for the initiation of the MINDACT (Microarray In Node negative Disease may Avoid ChemoTherapy) trial.

The MINDACT study

The MINDACT trial is an international prospective, randomized study comparing the 70-gene signature classifier with commonly used clinicopathologic criteria for selecting node-negative breast cancer patients for adjuvant chemotherapy. The trial is intended to address whether a tool such as this can improve on existing methods of risk assessment and treatment decision-making by assisting oncologists to select between node-negative women who need adjuvant chemotherapy and those who do not. We will discuss the challenges that arose in incorporating this question into a suitable design for the prospective clinical trial. Studies similar in purpose to MINDACT might become more frequent in the future as more prognostic and/or predictive signatures require validation before their use in clinical practice.

The need for a randomized trial

Given the available retrospective validation data of the 70-gene prognostic signature, is there any need to perform a large randomized trial? Although the available validation data are compelling, we believe that before being accepted as standard practice, new biological diagnostic tools must go through the same strict validation process as, for example, a new drug or treatment approach. While phase II results may be promising, a new therapy might only become a standard of care after being evaluated in at least one large prospective randomized phase III trial. This is especially true with this technology given its high cost - €2,000 per patient - and the complexity and costs associated with the collection of frozen tumor samples. The 21-gene recurrence score, developed by the National Surgical Adjuvant Breast and Bowel Project and Genomic Health, and based solely on retrospective validations, is marketed under the name Oncotype DX® (Genomic Health, Inc., Redwood City, CA) and has not yet been approved by the FDA. The scientific community shares our belief that a full prospective validation must be performed before this tool can be accepted as standard of care, and such a validation is about to start via a large phase III prospective trial - Tailor x - which is a joint effort between several American groups, funded by the National Cancer Institute. The design of Tailor x is quite similar to that of the MINDACT trial, and collaborations and discussions are underway between the two consortia utilizing the two trials. Only this type of large, prospective, biologically driven phase III trial can provide the necessary level 1 evidence (*Box 1*) for any new biological marker or tool.

The huge research efforts in the development of microarray-based gene signatures are often weakened by restricted numbers of patients. The supervised analysis of expression data of thousands of genes for a limited number of patients has well-known pitfalls.^{14,15} For the external validation data set of the 70-gene signature, while the hazard ratios were smaller than the previously published series, the result was still more powerful than any other available covariate for this data set, providing evidence that the prognosis provided

by the gene signature was robust and the technology reproducible. Nonetheless, as retrospective patient series can also be biased by unknown factors leading to patient selection, a prospective evaluation is vital. Furthermore, the use of the signature to guide chemotherapy decision-making has not been tested by external validation, which dealt only with the prognostic (as opposed to the predictive) value of the signature (*Box 1*).

It can also be argued that the patients selected for participation in a randomized trial will be an investigator-selected subset of the population under consideration. This is indeed an important point, and to evaluate selection bias during accrual an additional step has been incorporated into the MINDACT trial design. Following enrollment of 800 patients (termed the 'pilot' stage; see below), the data will be examined not only for logistical problems, but also for potential bias of investigators and compliance with the randomization.

Some questions have arisen over whether it is too early to initiate this trial, given the current rapid evolution of the technology, the previous methodological criticisms, and the real possibility that the technology may be outdated at the completion of the trial. Undoubtedly, there are and will be many predictive and prognostic signatures derived from high-throughput technology reported in the future. It is clear that gene signatures must be independently and externally validated before they proceed to prospective clinical assessment and widespread use. Furthermore, like all diagnostic approaches, the ultimate diagnostic gene signature may need refining as our knowledge advances. In our opinion, prospective studies are the only way to provide level 1 evidence (*Box 1*) about the clinical relevance of genomic signatures, and therefore the only way to endorse their widespread clinical use, thereby allowing patients to benefit from these advances. Participation in this trial should be strongly encouraged so that this issue can be addressed as soon as possible.

Key Points

- In the past 20 years, little progress has been made regarding new prognostic markers that can assist oncologists in treatment decision-making for node-negative early-stage breast cancer
- The MINDACT trial is a prospective, randomized, controlled trial that is designed to evaluate the use of genomic risk factors in the selection of patients for adjuvant chemotherapy
- The MINDACT trial is a prospective, randomized, controlled trial that is designed to evaluate the use of genomic risk factors in the selection of patients for adjuvant chemotherapy
- The MINDACT trial is a prospective, randomized, controlled trial that is designed to evaluate the use of genomic risk factors in the selection of patients for adjuvant chemotherapy
- The MINDACT trial is a prospective, randomized, controlled trial that is designed to evaluate the use of genomic risk factors in the selection of patients for adjuvant chemotherapy

Box 1. Common statistical terms relevant to the MINDACT trial.

Prognostic

Defines patient outcome based on overall survival or relapse-free survival in a group of patients independent of exposure to therapy

Predictive

Predictive factors should define sensitivity of a tumor to a distinct therapeutic agent

Hazard ratios

In survival analysis, the hazard ratio is an indication of the difference between two survival curves, representing the reduction in the risk of death with treatment compared with control, over the period of follow-up; the hazard ratio is a form of relative risk

Level 1 evidence

Evidence arising from a randomized controlled clinical trial

Specificity

The percentage of patients with a negative test result who were not diagnosed with malignancy

Sensitivity

The number of patients with a true positive test result (positive test result and tumor) divided by the total number of patients diagnosed with malignancy

Study endpoints

The 70-gene signature was developed by supervised microarray analysis, using stored tissue and data collected from women who were not treated with systemic therapy. Around 50% of node-positive and 30% of node-negative breast cancer patients will eventually relapse.¹⁶ Distant metastases from breast cancer represents a virtually incurable disease with almost 100% mortality. For the development of gene-signature studies, therefore, distant metastasis was chosen as the primary endpoint, and two groups of patients were selected: those who had metastasis within 5 years, and those who did not have metastasis within 5 years. This tool is, therefore, expected to be prognostic for distant metastasis-free survival (DMFS). The mechanisms of distant metastasis and locoregional recurrence are probably not the same, and there is, therefore, no guarantee or expectation that the signature will be prognostic for earlier endpoints; indeed, such was the case in the external validation data set.^{17,18}

From the recent update of the Early Breast Cancer Trialists' Collaborative Group metaanalysis,¹⁹ key data for node-negative patients are the following: the overall hazard ratio for breast-cancer-specific mortality for polychemotherapy *versus* no chemotherapy is 0.83 (standard error 0.02), with a trend toward stronger effect for younger patients, ranging from 0.71 (standard error 0.07) for age 40–60 years to 0.91 (standard error 0.04) for age 60–69 years. These hazard ratios do not seem to be affected by nodal status. Chemotherapy has the greatest effect in women with ER-negative tumors; the overall hazard ratio for breast-cancer-specific mortality for poly chemotherapy *versus* no chemotherapy is 0.74. For node-negative women aged less than 50 years, the 10-year OS estimates with and without polychemotherapy are 81% and 76%, respectively. For node-negative women aged 50–69 years, the 10-year OS estimates with and without polychemotherapy are 79% and 76%, respectively. If these hazard ratios are robust across risk groups, one can assume that it is possible to identify groups of relatively good risk, where the toxicity of the treatment administered to the whole group outweighs the absolute improvement in OS (*i.e.* % at 10 years).

Eligibility criteria

The MINDACT trial will enroll node-negative women aged 18–70 years with histologically proven operable invasive breast cancers, who have a negative sentinel node or a negative axillary clearance. Whilst the age of 60 years was used as a cutoff for external validation, recent data from the Netherlands Cancer Institute also supports the signature's value in women up to 70 years of age (L van't Veer, personal communication). The tumor should be stage T1, T2, or operable T3, and unilateral. Ductal carcinoma *in situ* or lobular carcinoma *in situ* can be included provided invasive cancer is present. Surgery options will include breast-conserving surgery or mastectomy combined with either a sentinel-node procedure or full

axillary clearance. No patients treated with previous chemotherapy or radiotherapy will be enrolled.

Determining the control arm

The trial objective is to prove that the 70-gene signature will safely assign fewer node-negative patients to chemotherapy, and is directly related to the control criteria. For these control criteria, two conditions are essential: the criteria should reflect current practice, and they should be applied as homogeneously as possible in the trial. Satisfying both conditions at the same time is a real challenge. Currently, oncologists decide whether to prescribe chemotherapy according to several methods and guidelines, and it is reasonable to assume that many modified versions of these criteria are being applied in practice. Thus, there is no straightforward way of deciding on one set of rules as the standard chemotherapy assignment method. One practical approach could be to allow the participating investigational sites to apply their own predefined set of rules to represent the current standard. Such an approach, however, would lead to considerable variability, and the rate of chemotherapy assignment of such 'standard' criteria would depend on the accrual at each participating site, rather than on the characteristics of the population studied.

As the Adjuvant! software uses information from the San Antonio database, the SEER (Surveillance, Epidemiology and End Results) database, the Overviews of clinical trials, individual clinical trial results, and the literature in general, it is considered appropriate for analysis of available patient prognosis data. It should be noted that the whole risk assignment method is largely prognosis-based, in contrast to other attempts to base the method on predictions of chemotherapy effect. Calculation of prediction has only been applied to some level for ER status, where the method acknowledges the effect of endocrine therapy in ER-positive patients, and the possibly greater effect of chemotherapy in ER-negative patients, leading to a sizable benefit even for good-prognosis ER-negative patients.

Also of note is the fact that the current version of the Adjuvant! software does not include HER2 status, a marker that many believe has considerable data supporting its prognostic value.²⁰ In addition, preliminary data indicate that HER2 status may have important predictive implications particularly related to endocrine therapy.²¹ To tackle this issue, a new version of the Adjuvant! software is being developed that will incorporate HER2 status. An additional consideration for HER2-positive patients is the need to administer adjuvant trastuzumab, which, with its very important efficacy results, could interfere with the detection of any difference between the chemotherapy and the endocrine therapy arms of the trial. How this effect will influence treatment decisions will be evaluated in the pilot phase of the first 800 patients.

Design of the trial

The initial concept of the trial design was straightforward. Each enrolled node-negative patient would be randomized to chemotherapy treatment decision according to either clinicopathologic criteria (control) or gene signature (experimental). The trial would then aim to prove that a lower rate of chemotherapy administration in the experimental arm did not result in inferior efficacy. With this approach, only half of the patients would need to have their microarray analysis performed on a real-time basis.

There were two major and interrelated objections to this design. The first objection was that, while this design tested the two approaches (experimental and control) against each other in an overall fashion, it did not take into account the fact that for more than half of the patients both approaches were in agreement. From a methodological perspective, it is clear that any benefit of either approach would be greatly diluted in such overall comparison, because a majority of patients could achieve the same result regardless of the arm of randomization. Additionally, since it is impossible to 'blind' the investigators to the clinicopathologic prognostic factors, in practice this design would actually compare the combination of the methods with the clinicopathologic risk assessment alone because of selection bias. This issue is related to the second limitation of this approach. Apart from the discussion of defining an appropriate noninferiority threshold (δ), for the time being, let us say we want to reject an inferiority null hypothesis (H_0) on the overall population of $H_0 = \theta_E \div \theta_C \geq 1.25$, where θ_E and θ_C represent the experimental and control hazards, respectively, for some time to event endpoint (DMFS or OS). For any reasonably sized trial, it would be very hard to come up with a credible scenario wherein such a noninferiority test would be likely to fail. In a noninferiority testing situation in which one can assume a hazard ratio between two randomized arms, a one-sided 95% confidence noninferiority test for the above null hypothesis would require about 512 events to perform the analysis with 80% power. We performed simulations for such a trial and primary test, and showed that in a situation in which the clinical criteria would perform very well (identifying 90% of patients who will metastasize), and the gene signature would select patients for chemotherapy at random, the study would still yield powers of up to 50%. For less extreme, but equally unacceptable situations, in which the gene signature should be identified as inferior, the power of 'proving' noninferiority would be even higher.

A trial using an overall (*i.e.* using all patients) noninferiority test would need to have an inferiority threshold hazard ratio of at least 0.90, and probably at least 0.95, to convincingly exclude performances of a gene signature that would still be considered to be very poor. Thus, a noninferiority trial that would reliably exclude the poor performance of a gene signature would need to be huge or of extremely long duration. Even under such conditions, it is likely that a scenario of underperforming gene tests that would yield reasonably nice hazard ratios in an overall comparison would arise. The addition of a large fraction of observations that are the same irrespective of the arm they are randomized

to makes any equivalence test procedure highly suspect. Such observations increase the likelihood of 'random noise', which makes statistically significant rejection of inferiority more likely. Such a situation, in general, is not acceptable for equivalence or noninferiority testing.²² This design, therefore, was not selected, and other possibilities were considered. Two other options that could be considered were the following: to assign chemotherapy according to the 70-gene classifier risk model while assessing the clinical risk, or to assign chemotherapy according to clinical criteria while assessing the 70-gene risk. The first option is a much too big leap forward and the second would not test the 70-gene signature appropriately.

As an outcome of the above considerations, attention and discussion started focusing on discordant patients, and considered such patients as the core group. Using the original approach, this group would have been identifiable only in patients randomized to use the gene signature, as the signature would not have been performed for the other arm. It became clear that one should perform the gene signature on all patients entered into the trial, in order to identify all patients who would be treated differently by the two methods. This group of patients would consist of two distinct subgroups: those who are at low risk according to the gene signature and high risk according to clinicopathologic criteria (stratum A), and those who are at high risk according to the gene signature and low risk according to clinicopathologic criteria (stratum B; *Figure 1*). Having a lower chemotherapy assignment rate with the gene signature is equivalent to saying that stratum A should be larger than stratum B, because more patients would be at low risk according to the gene signature.

Since the main objective of the trial is to put the gene signature to the test, we concluded that its design should randomize patients who have discordant risk assessments to one of the two methods to be used for chemotherapy decision-making. In fact, such a course is equivalent to randomizing such patients to receive chemotherapy or not. For the other patients, who have either both risk assessments as high risk, or both as low risk, a randomization for this type of risk assessment will not have a true value. These patients, therefore, will not be randomized, but will be treated or not with chemotherapy according to the concordant risk assessments, and followed further. All patients with hormone-sensitive disease will also receive endocrine therapy.

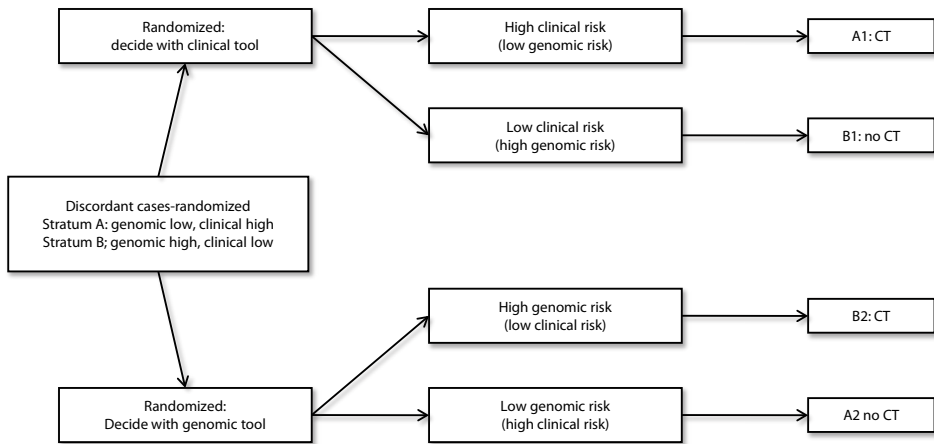


Figure 1. Randomization of discordant cases in the MINDACT trial.

Abbreviation: CT, chemotherapy.

Sample size

At the time of the first discussions regarding the MINDACT trial design, the noninferiority design was proposed and a sample size of 5,000 was envisaged, with microarray analysis performed on half of these patients. As discussions evolved, it was decided that only the discordant cases would be randomized to treatment decision-making using either the genomic tool or the clinicopathologic criteria. In order to be able to formulate and answer questions with regard to the core group of discordant cases, the sample size was increased from 5,000 to 6,000 patients.

Testing

From the available data on the gene signature and Adjuvant! software, there is a strong belief that the gene signature will produce a larger fraction of patients designated as low risk than the clinicopathologic method described above. Assignment fractions according to both methods will be evident from the trial design, and both methods will be available for all patients, so paired data are available for analyzing this endpoint. After the preceding discussions, it would be fair to conduct a noninferiority test on the selected group of patients with discordant risk data. Unfortunately, with an expected proportion of discordant patients of 30–35%, the number of patients needed for the trial would become prohibiting, particularly considering the high costs associated with the microarray technology and the complex trial logistics.

Let us revisit the question of what the real objective of such a trial is. We can first clarify what the objective is not. It is not a trial to find the fractions of high-risk and low-risk patients

according to the gene signature, because these can be found with an observational trial. In this context, an observational trial would involve treating the patients according to the established clinicopathologic guidelines and the patients' genomic risk, but not making a decision based on the treatment outcome. Indeed, for this endpoint, the present trial will function as a very large prospective observational trial. Also, it is not a trial to determine the effect of chemotherapy in specific subgroups. Such an issue may be of interest, but it is not the primary goal of the trial to address it. *A priori* assumption (before having the outcome of the trial) may be that chemotherapy does have an effect (in terms of a hazard ratio) for some of the patients that have a good gene signature, but that for these patients the prognosis is so good that it is not acceptable to treat them all with a toxic treatment. This is the same rationale whereby not all node-negative patients are given chemotherapy. In mathematical terms, the hazard ratio may still be different from 1 (and may be more or less constant for the whole population), but because of the small event rate in an identifiable good-prognosis group, the absolute effect is outweighed by the acute and long-term toxicity of chemotherapy.

If we consider a gene signature (or some other set of criteria) as a diagnostic test to detect those patients who will have recurrent disease that can no longer be treated with curative intent (*i.e.* metastasis), we can discuss its performance in terms of specificity and sensitivity (*Box 1*). We cannot expect the gene signature to be perfect (*i.e.* to have 100% sensitivity and specificity), but we can try to prove that it is good enough in the clinical situation—that is, sufficient to prevent undertreatment of patients, which relates to the sensitivity of the gene signature. To address this requirement, we incorporated the following primary test into the trial design. In the set of patients who have a low-risk gene-signature prognosis and high-risk clinicopathologic criteria, and who will be randomized to use the gene-signature prognosis and thus receive no chemotherapy (group A2 in *Figure 1*), a null hypothesis of a 5-year DMFS of 92% will be tested. With 6,000 patients accrued overall, and based on the available validation data estimates, this set has an expected size of 672 patients. With an accrual of 3 years, and a total duration of 6 years (*i.e.* 3–6 years' follow-up for each patient), a one-sided test at 97.5% confidence level has 80% power to reject this hypothesis if the true 5-year DMFS is 95%.

The major criticism of this primary test is that it is not a test that compares the randomized groups. If the above test is statistically significant, however, and the gene signature does select fewer patients to be treated with chemotherapy while not adversely affecting DMFS, then this can be taken to be equivalent to proving that the signature has a very good sensitivity, as well as a specificity that is better than the clinicopathologic method. In our opinion, therefore, a significant primary test as described above would establish the role of such a signature in chemotherapy treatment decision-making in node-negative breast cancer patients.

The pilot phase

All patients must have an available and high-quality frozen tumor sample to be eligible for the MINDACT trial. The desired number of patients might become impractical because of logistics problems such as transport problems, insufficient material, insufficient quality of RNA, and so on. This issue is currently being assessed in a Pilot Logistics Study, run in seven centers in seven different European countries. Additionally, all these issues and all the assumptions made will be assessed and corrected if necessary during the pilot phase of MINDACT, composed of the first 800 patients.

This pilot phase should ensure that the complex logistical framework put in place for this trial is feasible for the patients, physicians and laboratories involved. Assessment by Agendia of the quality and quantity of RNA samples will be part of this first phase. The second aim of the pilot phase is to ascertain that the patient population recruited upfront for MINDACT is not a biased one. This will be done by checking whether the ratio of low-risk to high-risk patients is as expected. Third, there is a concern that clinicians might not comply with the randomization of the treatment decision, violating the protocol. A patient who is clinically high-risk and has a low genomic risk, who is randomized by use of the genomic tool, should not be given chemotherapy. By contrast, a patient who is clinically low-risk and has a high genomic risk and is randomized to decision with the genomic tool should be given chemotherapy. The compliance with this randomization is to be assessed in the pilot phase. Following the treatment randomization, there will be two additional randomizations: chemotherapy and endocrine therapy. The fourth objective of the pilot phase is to check whether at least 66% of those women assigned to chemotherapy are subsequently randomized in the chemotherapy question. The final aim of the pilot study is to ensure that there is a statistically significant difference between the percentage of patients that have a high clinicopathologic risk and those with a high genomic risk, thus reflecting the expected reduction in chemotherapy administration.

Further randomizations

Since the primary randomization is the most complex and innovative part of this trial, the bulk of this article has focused on this element, but it should be noted that the trial will also have two further randomizations. Patients who are to receive chemotherapy may be randomized to receive either an anthracycline-based regimen or a docetaxel–capecitabine regimen.^{23,24} This randomization (designated R-C) will ask whether a docetaxel–capecitabine regimen can safely replace an anthracycline-based regimen in high-risk node-negative women, with the potential advantage of a reduction in the two long-term toxicities associated with anthracyclines: cardiac toxicity and secondary leukemia. The docetaxel–capecitabine combination is currently being evaluated in the adjuvant setting in the US. As short-term toxicity is of some concern with this regimen, in MINDACT the

first 40–120 patients randomized to this treatment will be closely monitored. Several commonly used anthracycline-based regimens will be accepted within the trial that have adequate anthracycline dose intensity, three or more drug combinations, and six cycles of administration.

Patients eligible for endocrine therapy can participate in the endocrine therapy randomization stage of the trial (designated R-E), which consists of a randomization to 7 years of letrozole, or to 2 years of tamoxifen followed by 5 years of letrozole.^{25–27} There will be stratifications for HER2 status, ER-positive and/or progesterone-negative, and gene-signature risk. These further randomizations will answer clinically relevant questions by taking advantage of the power that the large sample size used in MINDACT offers. As the associated biological material will be collected, there will be also ample opportunity to develop and identify predictive gene signatures, as well as important genes and proteins influencing response to administered agents (*Figure 2*).

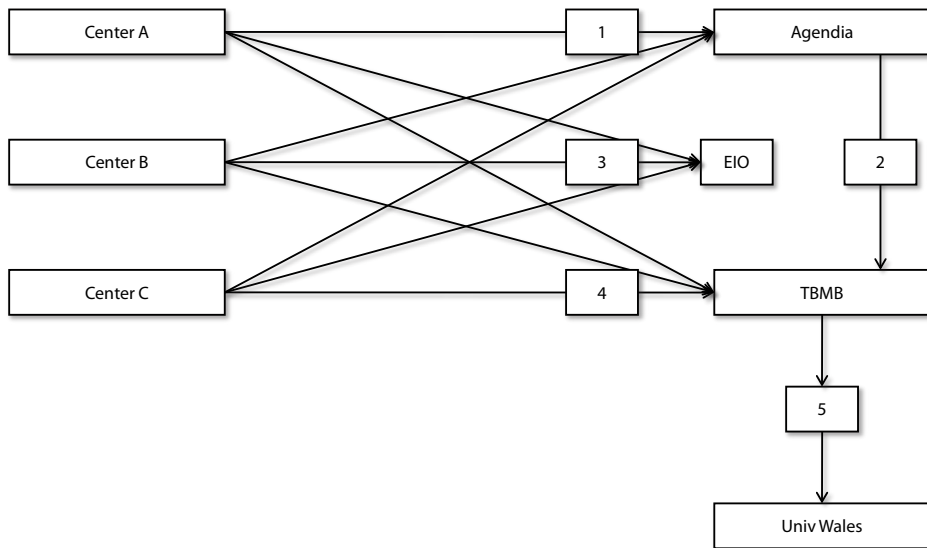


Figure 2. Biological material flowchart in the MINDACT trial. (1) Tissue for RNA extraction. (2) Any remaining tissue and RNA. (3) Paraffin blocks for central pathology review and tumor microarray production. (4) Serum. (5) Tissue and/ or serum for proteomics.

Abbreviations: EIO, European Institute of Oncology; TBMB, TRANSBIG biological materials bank; Univ Wales; University of Wales.

Predictive power of the 70-gene signature

At this time, there are no data on the predictive power of the 70-gene signature. It would be interesting to have predictive power data on anthracycline–taxane-based chemotherapy, tamoxifen, and aromatase inhibitors, because these are the treatments administered in the MINDACT trial. The European Organisation for Research and Treatment of Cancer (EORTC) Breast Cancer Group is now undertaking a project in which the 70-gene signature is being evaluated in some of the patients enrolled in the EORTC 10994/BIG 00-01 ‘p53’ trial.²⁸ The p53 trial randomized patients with locally advanced or large operable tumors to one of two neo adjuvant chemotherapy regimens: six cycles of fluorouracil, epirubicin and cyclophosphamide (epirubicin 100 mg/m²) or three cycles of docetaxel (100 mg/m²) followed by three cycles of docetaxel (75 mg/m²) combined with epirubicin (90 mg/m²). Results are not yet available, but the use of standardized chemotherapy regimens and the availability of enough good-quality frozen material will probably yield data on the predictive power of the signature. These results will help adjustment of estimates for the MINDACT trial, if needed.

The MINDACT trial will provide the setting for prospectively assessing the predictive power of the 70-gene signature and any other signatures currently being developed in response to the chemotherapy regimens and endocrine therapy used.

Logistics

The logistics of MINDACT have been one of the most challenging and expensive parts of the trial. All RNA extraction, quality control and microarray analysis for samples in this trial will be performed at Agendia in Amsterdam. Indeed, at this stage this technology is probably too immature for even the RNA extraction to be performed in external laboratories; operator and technical variability is well known to influence the results of microarray experiments.

Upon diagnosis of a clinically node-negative invasive breast cancer, patients are eligible to sign the first informed consent prior to surgery to allow for a sample of frozen tumor tissue obtained during surgery to be sent to Agendia for RNA extraction (screening informed consent). At this time, only RNA extraction and quality control check will be done. Once the local pathology report confirms node negativity, the genomic risk assessment will be performed. This process will hopefully avoid much unnecessary hybridization and hence reduce the cost. All tumor specimens will be couriered to Agendia by a specifically contracted courier agent specialized in global express transport and storage at –80 °C. If a patient is ineligible for the trial, her tumor material will be returned, stored in the TRANSBIG biological materials bank, or destroyed.

Simultaneously, the investigator responsible for patient care will assess the clinicopathologic risk using the Adjuvant! software embedded in a web-based platform designed specifically for the MINDACT randomization. In Informed consent form 1, the

patient will consent to have her risk assessed by the genomic and clinicopathologic methods and to enter randomization for treatment (R-T) if she belongs to the discordant group (Figure 3). If the patient is assigned to receive chemotherapy, she will be proposed to enter the chemotherapy randomization (R-C) and to sign the Informed consent form 2. After chemotherapy (if applicable) and radiotherapy, the endocrine therapy randomization (R-E) will be proposed to endocrine responsive patients in the Informed consent form 3. Paraffin tumor blocks will be sent every 6 months to the European Institute of Oncology in Milan, for Construction of tissue arrays and for central pathology review to be performed. Proteomic analysis of the tumor and serum samples will also be performed in the MINDACT trial, in collaboration with the University of Wales in Aberystwyth. Additionally, since one of the aims of MINDACT is to create a biological materials bank, frozen tumor samples (as well as whole genome microarray data and paraffin-embedded tissue) will be collected for all patients. The TRANSBIG biological materials bank will be located in Brussels under the guardianship of TRANSBIG, and hence this trial will have great potential for the identification and validation of additional gene signatures with prognostic and predictive value in early breast cancer, as well as other markers and technologies. Figures 2 and 3 summarize some of the logistics involved in this trial.

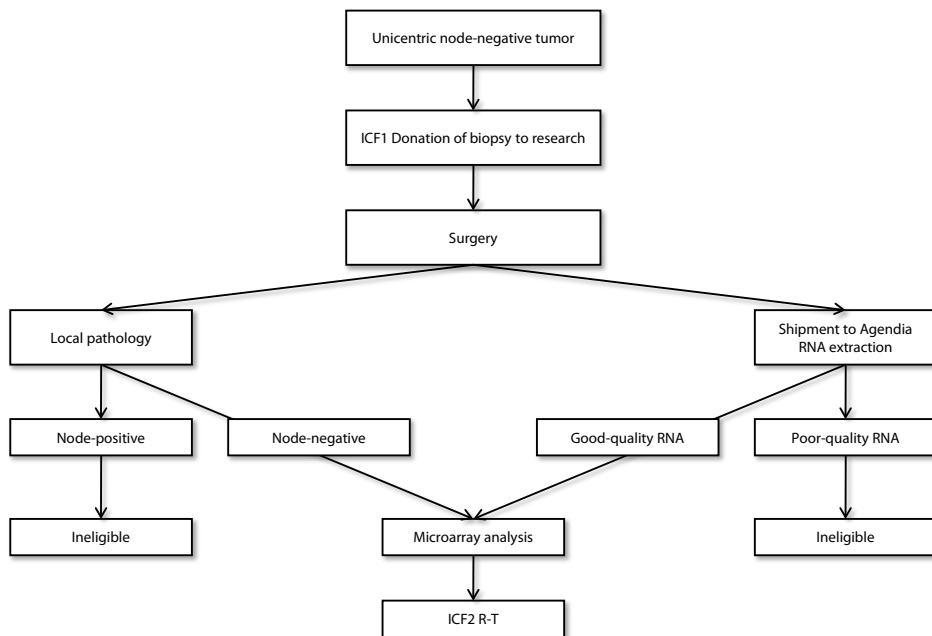


Figure 3. Logistics of MINDACT—tumor biopsy collection, shipment, RNA extraction and eligibility check.

Abbreviations: ICF1, informed consent form 1; ICF2, informed consent form 2; R-T, randomization to treatment.

Conclusions

We have summarized the history of the development of the MINDACT trial, and the reasoning that led to its final design. The goal of the trial is to prospectively validate the 70-gene prognosis signature providing the level 1 evidence of its utility that is needed before its wide clinical application. This study aims to give us a definitive answer regarding the clinical relevance of the 70-gene signature, its performance compared with traditional prognostic factors and, as a secondary aim, its ability to predict response to commonly prescribed adjuvant treatments. Using this new tool, MINDACT aims to better define patient prognosis and therefore to better select patients who need adjuvant chemotherapy; by doing so, it is expected that 10–20% of women who would normally receive adjuvant chemotherapy based on their clinicopathologic factors will be spared the inconvenience and morbidity of this therapy, without having any negative impact on their survival. The primary objective of the trial is to confirm that the number of patients that can be safely spared adjuvant chemotherapy is significantly increased when the decision is based on the 70-gene signature rather than on clinicopathologic methods, while the critical group of patients who have a high risk of recurrence according to the clinicopathologic criteria but a low risk according to the 70-gene prognosis signature is not undertreated. The sparing of patients from chemotherapy has immense benefits for individual patients and society. We strongly believe that clinical trials that aim to individualize cancer therapy are the way forward.

References

1. Van de Vijver MJ, He YD, Van 't Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 2002; 347: 1999-2009.
2. Hutchins L, Green S, Ravdin P, et al. CMF versus CAF with and without Tamoxifen in high-risk node-negative breast cancer patients and a natural history follow-up study in low-risk node-negative patients: First results of Intergroup trial INT 0102 [abstract #2]. *Proc Am Soc Clin Oncol* 1998; 17: a1
3. McShane LM, Altman DG, Sauerbrei W, et al. REporting recommendations for tumor MARKer prognostic studies (REMARK). *Nat Clin Pract Oncol* 2005; 2: 416-422.
4. National Institutes of Health Consensus Development Conference Statement: Adjuvant Therapy for Breast Cancer, November 1-3, 2000. *J Natl Cancer Inst* 2001; 93: 979-989.
5. Goldhirsch A, Glick JH, Gelber RD, et al. Meeting highlights: International expert consensus on the primary therapy of early breast cancer 2005. *Ann Oncol* 2005; 16: 1569-1583.
6. D'Eredita' G, Giardina C, Martellotta M, Natale T, Ferrarese F. Prognostic factors in breast cancer: the predictive value of the Nottingham Prognostic Index in patients with a long-term follow-up that were treated in a single institution. *Eur J Cancer* 2001; 37: 591-596.
7. Van't Veer LJ, Dai H, Van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002; 415: 530-536.
8. Van de Vijver MJ, He YD, Van 't Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 2002; 347: 1999-2009.
9. Piccart MJ, Loi S, Van 't Veer LJ, et al. Multi-center external validation study of the Amsterdam 70-gene prognostic signature in node negative untreated breast cancer: are the results still outperforming the clinical-pathological criteria? [abstract] *Breast Cancer Res Treat* 2004; 88 (Suppl 1): a538.
10. Cardoso F et al. Multi-centre independent validation of a gene prognostic signature for patients with node-negative breast cancer (BC)—final results [abstract]. *Breast Cancer Res Treat* 2005; 95 (Suppl 1): aS301.
11. Adjuvant! Online [<http://www.adjuvantonline.com>].
12. Ravdin PM, Siminoff LA, Davis GJ, et al. Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer. *J Clin Oncol* 2001; 19: 980-991.
13. Olivetto IA, Bajdik CD, Ravdin PM, et al. Population-based validation of the prognostic model ADJUVANT! for early breast cancer. *J Clin Oncol* 2005; 23: 2716-2725.
14. Simon R, Radmacher MD, Dobbin K, McShane LM. Pitfalls in the Use of DNA Microarray Data for Diagnostic and Prognostic Classification. *J Natl Cancer Inst* 2003; 95: 14-18.
15. Dudoit S et al. Comparison of discrimination methods for the classification of tumors using gene expression data. *JASA* 2002; 97: 77-87.
16. De Vita Jr VT et al. (Eds; 2006) *Cancer. Principles and Practice of Oncology*. Philadelphia: Lippincott Williams & Wilkins.
17. Cardoso F et al. Multi-centre independent validation of a gene prognostic signature for patients with node-negative breast cancer (BC)—final results [abstract]. *Breast Cancer Res Treat* 2005; 95 (Suppl 1): aS301.
18. Buyse M, Loi S, Van 't Veer L, et al. Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. *J Natl Cancer Inst* 2006; 98: 1183-1192.

19. Early Breast Cancer Trialists' Collaborative Group. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet* 2005; 365: 1687-1717.
20. Ross JS, Fletcher JA, Linette GP, et al. The Her-2/neu gene and protein in breast cancer 2003: biomarker and target of therapy. *Oncologist* 2003; 8: 307-325.
21. Smith IE, Dowsett M, Ebbs SR, et al. Neoadjuvant treatment of postmenopausal breast cancer with anastrozole, tamoxifen, or both in combination: the Immediate Preoperative Anastrozole, Tamoxifen, or Combined with Tamoxifen (IMPACT) multicenter double-blind randomized trial. *J Clin Oncol* 2005; 23: 5108-5116.
22. ICH Harmonised Tripartite Guideline: statistical principles for clinical trials. International Conference on Harmonisation E9 Expert Working Group. *Stat Med* 1999; 18: 1905-1942.
23. O'shaughnessy J, Miles D, Vukelja S, et al. Superior survival with capecitabine plus docetaxel combination therapy in anthracycline-pretreated patients with advanced breast cancer: phase III trial results. *J Clin Oncol* 2002; 20: 2812-2823.
24. Jones S et al. Three year results of a prospective randomized trial of adjuvant chemotherapy for patients (pts) with stage I-III operable, invasive breast cancer comparing 4 courses of doxorubicin/ cyclophosphamide (AC) to 4 courses of docetaxel/ cyclophosphamide (TC) [abstract]. *Proc Am Soc Clin Oncol* 2003; a59.
25. Goss PE, Ingle JN, Martino S, et al. A randomized trial of letrozole in postmenopausal women after five years of tamoxifen therapy for early-stage breast cancer. *N Engl J Med* 2003; 349: 1793-1802.
26. Coombes RC, Hall E, Gibson LJ, et al. A randomized trial of exemestane after two to three years of tamoxifen therapy in postmenopausal women with primary breast cancer. *N Engl J Med* 2004; 350: 1081-1092.
27. Thurlimann BJ et al. BIG 1-98: randomized double-blind phase III study to evaluate letrozole (L) vs. tamoxifen (T) as adjuvant endocrine therapy for postmenopausal women with receptor-positive breast cancer [abstract #511]. *Proc Am Soc Clin Oncol* 2005; 23: 16S.
28. Bonnefoi H, ebold-Berger S, Therasse P, et al. Locally advanced/inflammatory breast cancers treated with intensive epirubicin-based neoadjuvant chemotherapy: are there molecular markers in the primary tumour that predict for 5-year clinical outcome? *Ann Oncol* 2003; 14: 406-413.