Prognostic and predictive role of ESR1 status for postmenopausal patients with endocrine-responsive early breast cancer in the Danish cohort of the BIG 1-98 trial


1Danish Breast Cancer Cooperative Group Statistical Center; Department of Oncology, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark; 2International Breast Cancer Study Group Statistical Center, Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, Boston, USA; 3Dako A/S, Glostrup, Denmark; 4Department of Biostatistics, Harvard School of Public Health; Department of Medicine, Harvard Medical School, Boston, USA; 5Department of Breast Cancer Research, Institute of Cancer Biology, Danish Cancer Society, Copenhagen, Denmark; 6International Breast Cancer Study Group Statistical Center, Frontier Science and Technology Research Foundation, Boston, USA; 7Department of Pathology, Herlev Hospital, Copenhagen University Hospital, Herlev, Denmark; 8Division of Pathology and Laboratory Medicine, International Breast Cancer Study Group Pathology Review Office, European Institute of Oncology, University of Milan, Milan, Italy

Received 1 June 2011; revised 10 August 2011; accepted 19 August 2011

Background: Estrogen Receptor 1 (ESR1) aberrations may be associated with expression of estrogen receptor (ER) or progesterone receptor (PgR), human epidermal growth factor receptor-2 (HER2) or Ki-67 labeling index and prognosis.

Patients and methods: ESR1 was assessed in 1129 (81%) of 1396 postmenopausal Danish women with early breast cancer randomly assigned to receive 5 years of letrozole, tamoxifen or a sequence of these agents in the Breast International Group 1-98 trial and who had ER ‡ 1% after central review.

Results: By FISH, 13.6% of patients had an ESR1-to-Centromere-6 (CEN-6) ratio ‡ 2 (amplified), and 4.2% had ESR1-to-CEN-6 ratio <0.8 (deleted). Deletion of ESR1 was associated with significantly lower levels of ER (P < 0.0001) and PgR (P = 0.02) and more frequent HER2 amplification. ESR1 deletion or amplification was associated with higher-Ki-67 than ESR1-normal tumors. Overall, there was no evidence of heterogeneity of disease-free survival (DFS) or in treatment effect according to ESR1 status. However, significant differences in DFS were observed for subsets based on a combination of ESR1 and HER2 status (P = 0.02).

Conclusions: ESR1 aberrations were associated with HER2 status, Ki-67 labeling index and ER and PgR levels. When combined with HER2, ESR1 may be prognostic but should not be used for endocrine treatment selection in postmenopausal women with endocrine-responsive early breast cancer.

Key words: aromatase inhibitor, breast cancer, ERBB2, ESR1, HER2, tamoxifen

introduction

Breast cancer patients whose tumors do not contain or express estrogen receptor (ER) alpha are unlikely to benefit from endocrine therapy [1]. Assessment of ER status is therefore widely recommended in breast cancers [2, 3]. A significant proportion of patients still relapse despite the appropriate use of endocrine therapy guided by ER, and there is a need to develop additional biomarkers.

*Correspondence to: Dr B. Ejlertsen, DBCG, Building 2501 Rigshospitalet, 9 Blegdamsvej, DK-2100 Copenhagen, Denmark. Tel. +45-3545-5988; Fax: +45 3545 4992; E-mail: ejlertsen@rh.dk

© The Author 2011. Published by Oxford University Press on behalf of the European Society for Medical Oncology. All rights reserved. For permissions, please email: journals.permissions@oup.com

Gene expression studies have identified molecular subtypes dissimilated by ER and human epidermal growth factor receptor-2 (HER2) status and in particular identified molecular subtypes within ER-positive cancers (mainly luminal A and B) [4, 5]. The prognostic implication of proliferation and its relation to classical prognostic factors has been demonstrated clearly in a meta-analysis of gene expression profiles in breast cancer [6]. The Ki-67 labeling index is an established indicator of proliferation [7, 8] and also appears to differentiate luminal A from luminal B subtypes in ER-positive and HER2-negative breast cancer [9].
The implementation of gene expression profiles in clinical practice is complicated by the requirement for either freshly frozen tumor tissue or otherwise complicated extraction processes. Recently, studies using FISH have suggested that copy number changes of Estrogen Receptor 1 (ESR1), the gene encoding the ER, quite frequently are present in breast cancers and may hold prognostic information [10–12].

In the Breast International Group (BIG) 1-98 study, we compared 5 years of tamoxifen and letrozole with sequences of 2 years of one of these agents followed by 3 years of the other. Initial results showed that letrozole given alone as compared with tamoxifen alone reduced the risk of recurrences in particular at distant sites [13]. A later protocol-specified analysis showed that sequential treatment with letrozole as compared with letrozole monotherapy did not improve disease-free survival (DFS) [14].

The purpose of this report is to assess associations between aberrations of ESR1, alone and combined with HER2 status and Ki-67 labeling index, and to evaluate the predictive value of ESR1 for initial adjuvant treatment in postmenopausal women with ER-positive breast cancer enrolled in the BIG 1-98 trial.

patients and methods

The BIG 1-98 patient population was defined as postmenopausal women with early invasive breast cancer whose tumors were assessed by local pathologists as either ER positive, progesterone receptor (PgR) positive or both. Histological type according to the World Health Organization and histological grade (ductal or lobular carcinomas) according to Elston and Ellis were recorded locally. From March 1988 to March 2000, patients were randomly assigned to 5 years of monotherapy with tamoxifen or letrozole and from April 1999 to May 2003, to 5 years of monotherapy or the sequential administration of one drug for 2 years followed by the other for 3 years [13, 14].

The trial enrolled 1402 Danish patients and 1396 of these were included in the intention-to-treat population. A negative sentinel node biopsy or axillary clearance (level I and part of level II) in combination with breast-conserving surgery or mastectomy was required. Radiotherapy was mandatory to the breast following lumpectomy (48 Gy) and the chest wall following mastectomy (48 Gy) if the tumor was >5 cm node or positive and against regional nodes (48 Gy) in node-positive disease, all in 2-Gy fractions and 5 fractions per week. Adjuvant chemotherapy was not recommended at Danish centers. The Danish Medicines Agency and the Danish National Committee on Biomedical Research Ethics approved in 1997 the double-blinded BIG 1-98 trial (KF 02-178/97) and the Ethical Committee of the Capital Region approved the current biomarker study before its activation (KF 12-142/04).

central assessment of ER, PgR, HER2 and Ki-67

The IBCSG Central Pathology Laboratory carried out central review on whole tissue sections of paraaffin-embedded primary tumor specimens for ER (clone 1D5; Dako, Glostrup, Denmark) and PgR (clone 1A6; Dako) by immunohistochemistry (IHC) [15] and for HER2 by IHC (HercepTest kit; Dako) and FISH using PathVysion HER-2 DNA Probe Kit (Abbott Molecular-Vysis, Chicago, IL) [16]. Tumors were considered to express ER or PgR if they showed at least 1% of immunoreactive cells. Tumors were scored as HER2 amplified by FISH by a HER2-to-Centromere-17 ratio >2.2 in one case with nonassessable FISH results if IHC was 3+. Ki-67 labeling index was assessed by IHC using the Mib-1 mAb (1:200 dilution; Dako) and categorized as high (≥14%) or low [17]. All pathology central review was carried out without knowledge of other characteristics, treatment assignment or outcomes.

assessment of ESR1

Tissue microarrays (TMAs) were constructed from formalin-fixed and paraffin-embedded tumor blocks by means of a TMA maker (AH diagnostics, Aarhus, Denmark). A target area was identified in the donor block on hematoxylin-stained sections and two 2-mm tissue cores were transferred to the recipient TMA block [18].

ESR1 copy number was assessed on TMAs using a FISH probe (Dako) covering the entire ESR1 gene at 6q25. At least 60 gene signals were scored. The centromere of chromosome 6 signals were scored in the same nuclei, and the gene-to-centromere ratio was calculated as previously described [11]. In brief, tumors were scored as ESR1 deleted, normal or amplified according to a predefined ratio of <0.8, 0.8–1.9 and ≥2.0 and aberrations (deletions and amplifications) were classified as abnormal. ESR1 was scored without knowledge of patient characteristics, treatment assignment or outcomes.

statistical methods

The protocol-specified primary end point was DFS, which was defined as the time from randomization to the earliest time of invasive local, regional or distant recurrence; a new invasive breast cancer in the contralateral breast; any second (non-breast) malignancy or death without prior cancer event [13]. Patients were grouped according to their ESR1 status: normal, deleted or amplified. The association of ESR1 status with patient and tumor characteristics was assessed using Fisher’s exact test for categorical information and Wilcoxon rank-sum tests for continuous information such as level of ER, PgR and Ki-67 expression. Log-rank tests, stratified by randomization option (two-arm and four-arm), were used to compare DFS among the ESR1 status groups. Kaplan–Meier estimates of DFS were calculated. Multivariable Cox proportional hazards modeling, stratified by randomization option and by therapy assignment, was used to adjust for tumor size, tumor grade, nodal status, PgR status, HER2 status and Ki-67 level. Hazard ratios (HRs) and 95% confidence intervals (CIs) were estimated using the Cox model, and Wald statistics were used to test for interactions. Analyses were conducted to assess DFS outcome according to ESR1 status alone and ESR1 status together with HER2 status. Of the 413 patients assigned to tamoxifen, 152 (37%) selectively crossed over to letrozole after the BIG 1-98 results were reported in 2005. For analyses comparing treatments, the follow-up of these 152 patients was censored at the time of selective crossover [14, 19]. The BIG 1-98 trial is registered on the clinical trials site of the USA National Cancer Institute website http://www.clinicaltrials.gov/ct/show/NCT00004205.

results

Archival tissue from the primary tumor was collected from 1323 (95%) of the 1396 Danish participants enrolled in the intent-to-treat population of the BIG 1-98 trial. The analytic cohort consisted of 1129 (81%) patients whose tumors were assessable for ESR1 and were confirmed to be ER positive (≥1%) by the central pathology assessment. No significant difference was shown in DFS for the 1129 patients in the analytic cohort compared with the 267 excluded patients (P = 0.52).

ESR1 status according to other patient and tumor characteristics

The clinical and tumor characteristics of the analytic cohort according to ESR1 status are shown in Table 1. Among these
1129 patients, 47 (4.2%) had an ESR1 deletion, 154 (13.6%) an ESR1 amplification and 928 (82.2%) an ESR1-normal tumor. Patients with ESR1 aberrations (deletions and amplifications) had a similar age distribution to those with ESR1-normal tumors.

Amplification and deletion of ESR1 were associated with a higher tumor grade ($P = 0.0003$) but not with larger tumor size ($P = 0.69$). More patients with amplification of ESR1 (42.2%) were lymph node positive than patients with ESR1-normal (34.4%) or -deleted (25.5%) tumors, but this did not reach statistical significance ($P = 0.07$).

Although the median ER expression was 90% regardless of ESR1 status, the 25th and 75th percentiles of ER were shifted lower for ESR1-deleted tumors ($P < 0.0001$; Figure 1A).

Deletion of ESR1 was also associated with lower PgR expression levels as compared with ESR1-normal and -amplified tumors ($P < 0.02$; Figure 1B). High Ki-67 ($\geq 14\%$) was observed significantly more often in patients with ESR1 aberrations than in those with normal ESR1 ($P < 0.0001$; Fisher’s exact test; Figure 1C). More HER2-amplified tumors (19.1%) were seen among patients with ESR1-deleted tumors compared with ESR1-amplified (7.1%) and -normal (8.3%) tumors ($P = 0.04$; Figure 1D).

In this ER-expressing population, an interesting pattern emerged when three subtypes were defined using HER2 and Ki-67 according to the 2009 St Gallen recommendations [2]. The ESR1-normal group had a higher percent of patients with Ki-67-low and HER2-normal tumors than the other groups, while the ESR1-amplified group had a higher percent of Ki-67-high and HER2-normal tumors, and the ESR1-deleted group had a higher percent of HER2-amplified tumors ($P < 0.0001$; Table 1).

### Prognostic value of ESR1 status

ESR1 status alone was not a prognostic factor for DFS in univariable analysis (log-rank $P = 0.38$; Figure 2A). In multivariable Cox proportional hazards modeling adjusting for tumor size, tumor grade, nodal status, PgR status, HER2 status and Ki-67, ESR1 status was not associated with DFS ($P = 0.83$). When compared with the 928 ESR1-normal patients in multivariable analysis, DFS was not significantly different for the 47 patients with deletion of ESR1 (HR = 1.0; 95% CI 0.98–1.02) or for the 154 patients with amplification of ESR1 (HR = 1.05; 95% CI 0.77–1.43; Figure 3).

As expected, the small group of patients with a HER2-amplified tumor had a worse DFS outcome compared with those having HER2-normal disease (HR = 1.79; 95% CI 1.27–2.51; $P = 0.0008$). Therefore, to explore the possible prognostic impact of combining information about HER2 and ESR1 status, specifically to assess ESR1 in HER2-normal tumors, patients were categorized in four subsets (HER2 amplified, HER2 and ESR1 normal, HER2 normal/ESR1 deleted, HER2 normal/ESR1 amplified). As reflected in Kaplan–Meier curves (Figure 2B), the univariable log-rank test showed an association of this categorization with DFS ($P = 0.001$). For patients with HER2-normal tumors, compared with the ESR1-normal group, we observed an increased risk of a DFS event in the ESR1–amplification group (HR = 1.37; 95% CI 1.00–1.88; $P = 0.05$) but not in ESR1 deletion (HR = 0.81; 95% CI 0.40–1.63; $P = 0.55$).

After adjustment for tumor features in a multivariable model, subsets according to HER2 and ESR1 remained significantly associated with DFS ($P = 0.02$). When analyses were restricted to HER2-normal cases, DFS was not significantly different for the 38 patients with deletion of ESR1 (HR = 0.76; 95% CI 0.37–1.54; $P = 0.44$) or the 143 patients with amplification of ESR1 (HR = 1.24; 95% CI 0.89–1.71; $P = 0.20$) when compared with the 850 ESR1-normal patients (Figure 3). We found no statistical evidence of heterogeneity in DFS when further subdividing patients with ESR1- and HER2-normal tumors according to Ki-67 (Figure 3).

### Predictive value of ESR1 status

Of the 1129 patients in the analytic cohort, 670 were randomly assigned to receive 5 years of letrozole or tamoxifen and were included in analyses assessing the role of ESR1 status to predict
relative treatment effects of the monotherapy regimens. In multivariable analyses censoring follow-up for selective crossover patients and adjusting for prognostic factors, the estimated reduction in the risk of a DFS event was seen for letrozole compared with tamoxifen (HR = 0.82; 95% CI 0.61–1.09). There was no evidence of heterogeneity in the

Figure 1. Box plots illustrating the distribution of ER, PgR and Ki-67 labeling index expression levels according to \( ESR1 \) status. Boxes indicate 25th, 50th (median) and 75th percentiles. \( P \) values were derived from Wilcoxon rank-sum test for ER percent \( (P < 0.0001) \), PgR percent \( (P = 0.02) \) and Ki-67 labeling index percent \( (P < 0.0001) \) and from Fisher’s exact tests for HER2 \( (P = 0.04) \). Missing values were omitted. ER, estrogen receptor; PgR, progesterone receptor; HER2, human epidermal growth factor receptor-2.

Figure 2. Kaplan–Meier estimates of disease-free survival according to \( ESR1 \) status (A) and according to \( ESR1 \) status and HER2 status categories (B). HER2, human epidermal growth factor receptor-2.
Others have reported much lower frequencies using 14% in both the current study and our preceding pilot study. Amplification ranging from 20% to 23%, as compared with Deletion of number may assist delineating the phenotype of breast cancers. Tamoxifen regardless of interaction), indicating similar benefit of letrozole versus outcomes for all three treatments (data not shown).

Array-comparative hybridization [20, 21], supplemented with quantitative PCR analysis (1%) compared with the result obtained with FISH (23%) [12]. Likewise, a recent comparative study showed a lower frequency of amplifications when comparing multiplex ligation-dependent probe amplification (2%) with FISH (12.5%) [25]. Our study is based on selected samples. In a direct comparison, the frequency of amplification was considerably lower when using real-time quantitative PCR analysis (1%) with FISH (12.5%) [25].

Our study is based on a cohort of patients with HER2-normal/Ki-67-low. HER2-normal/Ki-67-high were treated and monitored according to strict guidelines of the protocol. In addition, this report involves central ESR1 FISH as well as central ER, PgR, HER2, FISH and Ki-67 on >1100 pathological specimens. FISH is considered the most accurate method for detection of amplification, especially at low levels, and is compared with other methodologies less affected by a high content of normal tissue within the tumor [26]. We used a highly standardized FISH assay, developed for HER2 and approved by the Food and Drug Administration, and applied these methods for determination of ESR1 precisely as recommended by the manufacturer. The central biomarker evaluations were carried out blinded to treatment and outcome. The ESR1 FISH was carried out in a laboratory separated from the central laboratory that carried out the ER, PgR, HER2 and Ki-67 analysis, and the laboratory data were transferred directly to the IBCSG statistical office.

In our pilot study, we found that HER2-amplified tumors were present in 19% of patients with HER2-normal/Ki-67-low tumors. HER2-amplified tumors were present in this study (19%) leaving HER2-high tumors to be a better predictor of outcome than HER2-normal tumors. HER2-amplified tumors were associated with a decreased DFS, and in the univariable analysis of the current study, we found similar results in patients with HER2-high tumors, as opposed to the current study, Moelans et al. and our prior study demonstrating an association between high grade while others found no association with grade [11, 12, 25].

Figure 3. Cox proportional hazards model for disease-free survival comparing ESR1-deleted and ESR1-amplified cohorts (aberrant cohorts) versus the ESR1-normal cohort overall and for subgroups defined by HER2 status and Ki-67 label index category. Hazard ratio values <1.0 indicate a better outcome for the aberrant (either deleted or amplified) compared with ESR1 normal. HER2, human epidermal growth factor receptor-2.

treatment effect according to ESR1 status (P = 0.82 for interaction), indicating similar benefit of letrozole versus tamoxifen regardless of ESR1 status (data not shown).

A total of 691 patients on the four-arm randomization option who were not assigned to tamoxifen were available to assess differences in treatment effect for letrozole alone compared with either sequence according to ESR1 status. There was no evidence of heterogeneity of treatment effect according to ESR1 status (P = 0.75 for interaction), indicating similar outcomes for all three treatments (data not shown).

Discussion

In this analysis of Danish participants from the BIG 1-98 trial, we showed that ESR1 was amplified in 13.6% and deleted in 4.2% of the tumors. We have previously published the results from a pilot study, where we included a cohort of patients with identical characteristics. Using the same highly standardized FISH assay, similar percentages of ESR1 amplifications (14.2%) and deletions (4.4%) were detected in the pilot study [11]. More tumors with deleted ESR1 were HER2 positive, more ESR1-amplified tumors were Ki-67 high and more ESR1-normal tumors were Ki-67 low, suggesting that ESR1 copy number may assist delineating the phenotype of breast cancers. Deletion of ESR1 was associated with lower ER and PgR levels. Two recent studies have reported a frequency of ESR1 amplification ranging from 20% to 23%, as compared with 14% in both the current study and our preceding pilot study [10–12]. Others have reported much lower frequencies using array-comparative hybridization [20, 21], supplemented with FISH [22, 23] or chromogenic in situ hybridization [24] on selected samples. In a direct comparison, the frequency of amplification was considerably lower when using real-time quantitative PCR analysis (1%) compared with the result obtained with FISH (23%) [12]. Likewise, a recent comparative study showed a lower frequency of amplifications when comparing multiplex ligation-dependent probe amplification analysis (2%) with FISH (12.5%) [25]. Our study is based on patients who participated in the BIG 1-98 trial and therefore were treated and monitored according to strict guidelines of the protocol. In addition, this report involves central ESR1 FISH as well as central ER, PgR, HER2, FISH and Ki-67 on >1100 pathological specimens. FISH is considered the most accurate method for detection of amplification, especially at low levels, and is compared with other methodologies less affected by a high content of normal tissue within the tumor [26]. We used a highly standardized FISH assay, developed for HER2 and approved by the Food and Drug Administration, and applied these methods for determination of ESR1 precisely as recommended by the manufacturer. The central biomarker evaluations were carried out blinded to treatment and outcome. The ESR1 FISH was carried out in a laboratory separated from the central laboratory that carried out the ER, PgR, HER2 and Ki-67 analysis, and the laboratory data were transferred directly to the IBCSG statistical office.

In our pilot study, we found that ESR1 amplification was associated with a decreased DFS, and in the univariable analysis of the current study, we found similar results in patients with ER-positive and HER2-normal breast cancer. However, when adjusting for tumor size, positive lymph nodes, malignancy grade, PgR status and Ki-67, there was no statistical evidence in support of an association between ESR1 status and the risk of recurrence. In contrast, Holst et al. [10] found ESR1 amplification to be associated with good prognosis in tamoxifen-treated patients. This may in part be explained by an association between amplification of ESR1 and low malignancy grade in the Holst study, as opposed to the current study, Moelans et al. and our prior study demonstrating an association with high grade while others found no association with grade [11, 12, 25].

Amplification and overexpression of HER2 are associated with a high risk of recurrence, even with tamoxifen or aromatase inhibitors [16, 27, 28]. Among the 47 patients with ESR1-deleted tumors, 9 were HER2 amplified (19%) leaving too few events, even in this large study, to explore a possible association between HER2 status and DFS in the ESR1-deleted subset. Adjuvant chemotherapy and trastuzumab was not available for the Danish participants in the BIG 1-98 trial. These
treatments have been recommended for patients with HER2-positive disease since 2005 [2] and could potentially have changed the outcome for such patients. In a retrospective and exploratory analysis, we found support for heterogeneity in DFS when combining HER2 and ESR1 status. Patients with ESR1-amplified and HER2-negative tumors seem to have an intermediary prognosis compared with patients with ESR1- and HER2-normal tumors and patients with HER2-amplified tumors. We found no evidence supporting a differential benefit from letrozole over tamoxifen according to amplification or deletion of ESR1.

This study has some potential limitations. First, the BIG 1-98 trial did not include a control group of patients not receiving endocrine therapy, and a clearer prognostic and predictive value of ESR1 status may emerge in comparison with an untreated control group. Secondly, even within this large sample, only 154 tumors were ESR1 amplified and 47 were ESR1 deleted, and our conclusions must be considered with caution in view of the limited number of events. Thirdly, we defined cut-offs and scoring algorithms in advance and have omitted any optimization of the methodology behind the ESR1 FISH test.

In summary, in a subset of the BIG 1-98 study population, we have confirmed the frequencies of ESR1 aberrations established in a preceding pilot study. Differences in ER and PgR levels, HER2 status and Ki-67 levels were observed according to ESR1 status. ESR1 status was, however, not an independent prognostic marker and did not seem to be a selection criterion for treatment with letrozole versus tamoxifen in early breast cancer. Exploratory analysis suggested that ESR1 might be informative in patients with HER2-normal disease.

acknowledgements
We thank the patients, nurses, data managers, and physicians who contributed to the BIG 1-98 clinical trial; and the International Breast Cancer Study Group and Danish Cancer Cooperative Group for collaboration on this project.

funding
Danish Ministry of Health (2006-12103-272); The Danish Council for Strategic Research (2101-07-0022), 'Fonden til fremme af klinik eksperimentel cancerforskning specielt vedrørende cancer mammae'; United States National Institutes of Health (CA-75362).

disclosure
KVN and SM are Dako A/S employees and BE has a close relative employed by Dako A/S. No other authors have conflicts of interest.

references


