

This Provisional PDF corresponds to the article as it appeared upon acceptance. Fully formatted PDF and full text (HTML) versions will be made available soon.

Triple-negative, basal-like, and quintuple-negative breast cancers: better prediction model for survival

BMC Cancer 2010, **10**:507 doi:10.1186/1471-2407-10-507

Yoon-La Choi (ylachoi@skku.edu)
Eunsel Oh (eoh001@snu.ac.kr)
Sarah Park (parksr@catholic.ac.kr)
Yeonju Kim (yeonju@ncc.re.kr)
Yeon-Hee Park (yeonh.park@samsung.com)
Kyoung Song (songseoul@hotmail.com)
Eun Yoon Cho (eunyoonyoon.cho@samsung.com)
Yun-Chul Hong (ychong1@snu.ac.kr)
Jong Sun Choi (jscmd@naver.com)
Jeong Eon Lee (jeongeon.lee@samsung.com)
Jung-Han Kim (jinnee.kim@samsung.com)
Seok Jin Nam (seokjin.nam@samsung.com)
Young-Hyuk Lim (yh00.im@samsung.com)
Jung-Hyun Yang (drjh.yang@samsung.com)
Young Kee Shin (ykeeshin@snu.ac.kr)

ISSN 1471-2407

Article type Research article

Submission date 2 September 2009

Acceptance date 23 September 2010

Publication date 23 September 2010

Article URL <http://www.biomedcentral.com/1471-2407/10/507>

Like all articles in BMC journals, this peer-reviewed article was published immediately upon acceptance. It can be downloaded, printed and distributed freely for any purposes (see copyright notice below).

Articles in BMC journals are listed in PubMed and archived at PubMed Central.

For information about publishing your research in BMC journals or any BioMed Central journal, go to

© 2010 Choi *et al.*, licensee BioMed Central Ltd.

This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

<http://www.biomedcentral.com/info/authors/>

© 2010 Choi *et al.*, licensee BioMed Central Ltd.

This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Triple-negative, basal-like, and quintuple-negative breast cancers: better prediction model for survival

Yoon-La Choi^{1,2,*}, Ensel Oh^{3,4,*}, Sarah Park⁵, Yeonju Kim⁶, Yeon-Hee Park⁷, Kyoung Song³, Eun Yoon Cho², Yun-Chul Hong⁸, Jong Sun Choi⁹, Jung Eun Lee¹⁰, Jung Han Kim¹⁰, Seok Jin Nam¹⁰, Young-Hyuk Lim⁷, Jung-Hyun Yang¹⁰, Young Kee Shin^{3,4,§}

¹Laboratory of Cancer Genomics and Molecular Pathology, Samsung Medical Center, Seoul, Korea.

²Department of Pathology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea.

³Research Institute of Pharmaceutical Science, Department of Pharmacy, Seoul National University, Seoul, Korea.

⁴Interdisciplinary Program of Bioinformatics, Department of Natural Science, Seoul National University, Seoul, Korea.

⁵Division of Medical Oncology, Department of Internal Medicine, Seoul St. Mary's hospital, Catholic University, Seoul, Korea.

⁶Cancer Early Detection Branch, National Cancer Control Research Institute, National Cancer Center, Goyang, Korea.

⁷Department of Internal Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea.

⁸Department of Preventive Medicine, Seoul National University College of Medicine, Seoul, Korea.

⁹Department of Pathology, Dongguk University Ilsan Hospital, Dongguk University College of Medicine, Goyang, Korea.

¹⁰ Division of Breast and Endocrine Surgery, Department of Surgery, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea

*These authors contributed equally to the work.

§Corresponding author

Young Kee Shin, M.D., Ph.D.

Molecular Pathology Laboratory, Department of Pharmacy, Seoul National University

College of Pharmacy, San 56-1, Sillim-dong, Gwanak-gu, Seoul, 151-742, Korea.

Telephone: +82-02-880-9126

Fax: +82-02-872-1795

E-mail: ykeeshin@snu.ac.kr

Email addresses:

YLC: ylachoi@skku.edu

EO: eoh001@snu.ac.kr

SP: parksr@catholic.ac.kr

YK: yeonju@ncc.re.kr

YHP: yeonh.park@samsung.com

KS: songseoul@hotmail.com

EYC: eunyoon.cho@samsung.com

YCH: ychong1@snu.ac.kr

JSC: jscmd@naver.com

JEL: jeongeon.lee@samsung.com

JHK: jinnee.kim@samsung.com

SJN: seokjin.nam@samsung.com

YHL: yh00.im@samsung.com

JHY: drjh.yang@samsung.com

YKS: ykeeshin@snu.ac.kr

Abstract

Background: Triple-negative breast cancers (TNBCs) and basal-like breast cancers (BLBCs) are known as poor outcome subtypes with a lack of targeted therapy. Previous studies have shown conflicting results regarding the difference of prognostic significance between TNBCs and BLBCs. In this study, we aimed to characterize the prognostic features of TNBCs, in view of BLBCs and quintuple-negative breast cancers (QNBC/5NPs).

Methods: Using tissue microarray-based immunohistochemical analysis, we categorized 951 primary breast cancers into four or five subtypes according to the expression of ER, PR, HER2, and basal markers (CK5/6, EGFR).

Results: The results of this study showed that both TNBCs and BLBCs were associated with high histological and/or nuclear grades. When the TNBCs are divided into two subtypes by the presence of basal markers, the clinicopathologic characteristics of TNBCs were mainly maintained in the BLBCs. The 5-subgrouping was the better prediction model for both disease free and overall survival in breast cancers than the 4-subgrouping. After multivariate analysis of TNBCs, the BLBCs did not have a worse prognosis than the QNBC/5NPs. Interestingly, the patients with BLBCs showed significant adjuvant chemotherapy benefit. In addition, QNBC/5NPs comprised about 6~8% of breast cancers in publicly available breast cancer datasets

Conclusion: The QNBC/5NP subtype is a worse prognostic subgroup of TNBCs, especially in higher stage and this result may be related to adjuvant chemotherapy benefit of BLBCs, calling for caution in the identification of subgroups of patients for therapeutic classification.

Keywords: breast carcinoma; triple-negative; basal-like; prognosis

Background

Breast cancer is a heterogeneous disease, and basal-like breast cancers (BLBCs) have attracted particular attention because they have been shown to have shorter survival compared to the other subtypes [1, 2]. With the practical difficulties of gene-expression profiling as a routine diagnostic tool, immunohistochemical surrogate panels have been introduced to identify BLBCs that are ‘triple-negative breast cancers’ (TNBCs) (estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, HER2 not overexpressed). [2, 3]. The immunohistochemical panel that is considered to be the gold standard in the identification of BLBCs is that proposed by Nielsen et al., in which BLBCs are defined as breast cancers expressing neither ER/PR nor HER2 and expressing CK5/6 and/or EGFR [2]. Both TNBCs and BLBCs have attracted clinical interest in regard to their prognoses and associated therapeutic approaches [4]. Several clinical trials have been conducted on the routine diagnostic immunohistochemical assay of TNBCs [5]. However, it should be noted that identification of a subgroup of tumors based solely on the lack of expression of immunohistochemical markers affords the risk of incorrect assignment due to technical artifacts [6].

Recently, Kreike et al. examined the gene expression and pathological characteristics of 97 TNBCs and concluded that “BL tumors can be reliably defined by TN immunohistochemistry (IHC)” and that “TN tumors are synonymous with BL tumors” [7]. However, Rakha et al. mentioned that equating TNBC with BLBC is misleading, and insisted that while most TNBCs fall into the BL molecular subtype of breast cancer, the terms are not completely synonymous [8]. Many studies have shown that TNBCs have the worst prognosis without the option of chemotherapy. On the other hand, other studies have shown that the expression of “basal markers” (CK5/6, CK14, CK17, and/or EGFR) is associated with a poor prognosis [9-11]. Significantly, Rakha et al. and Cheang et al. reported that BLBCs defined by five biomarkers showed worse response to chemotherapy, and shorter survival [12, 13].

Therefore, the aim of our study was to compare prognosis between a model in which breast cancers were divided by four subtypes (including TNBC) and five subtypes (including BLBC), and to define the associated clinicopathological factors. To perform this study, we designed a longitudinal study, and used immunohistochemical markers to identify breast tumor intrinsic subtypes using formalin-fixed, paraffin-embedded (FFPE) tumor blocks, and

examined the clinicopathologic parameters of BLBCs and TNBCs. Also, we explored BLBC and TNBC in two large microarray datasets by comparing the expression levels of the five markers according to the subtypes.

Methods

Study population

We collected breast cancer cases from the Samsung Medical Center in Seoul, Korea. Inclusion criteria for this study were: 1) histologically verified incident breast cancer; 2) female; 3) between 20-80 years of age; 4) enrolled between January 1, 1995 and December 31, 2002; 5) breast tissue samples available for study. All of the subjects were diagnosed stage I to III primary breast cancer and underwent surgical treatment and adjuvant chemotherapy according to standard treatment protocols. Radiation therapy was performed in the cases of breast conserving surgery or on patients with stage III breast cancer. All ER-positive patients underwent hormonal therapy with tamoxifen. No patients underwent anti-HER-2 therapy. We reviewed each patient's medical record for clinical information, including follow-up status and outcome information. Breast cancer stage was classified according to the American Joint Committee on Cancer (AJCC) TNM criteria (6th edition). In the case of live subjects, the last date of follow-up was June 30, 2007. The protocol for the study was approved by the institutional review board (IRB) of the Samsung Medical Center.

Immunohistochemistry and fluorescent in situ hybridization

The original H&E-stained slides from the patients in the retrospective cohort were reviewed, and representative tumor regions without secondary change, such as hemorrhage, necrosis, and fibrosis, were marked by a pathologist (YL Choi). Corresponding 1290 FFPE breast cancer tissue blocks were obtained. Two 2mm cores from each case were obtained, and two sets of tissue microarray (TMA) paraffin blocks were made. The sections were deparaffinized with xylene, hydrated in serial dilutions of alcohol, and then immersed in 3% hydrogen

peroxide solution to neutralize endogenous peroxidase activity. Next, sections were microwaved in citrate buffer for antigen retrieval. Slides were incubated with monoclonal antibodies against CK5/6 (1:100, M7237, DAKO, Carpinteria, CA, USA), HER2 (1:250, A0485, DAKO), and EGFR (1:30, M7239, Novocastra) for 1 hour at room temperature. After washing, the tissue section was reacted with the biotinylated anti-mouse secondary antibody, followed by incubation with streptavidin-horseradish-peroxidase complex. Slides were washed, and the chromogen was developed for 5 minutes with liquid 3,3'-diaminbenzidine (DAKO). HER2 fluorescent in situ hybridization assay was performed with the PathVysion HER2 DNA Probe Kit (Abbott Molecular, Inc.) according to the manufacturer's instructions. The average copy number for each probe was determined and the amplification ratio was calculated as a ratio between the average copy per cell for HER2 and the average copy number for centromere 17.

Two pathologists (YL Choi, JS Choi) were blinded to the clinical outcomes of the patients, and independently scored the results of the staining. ER and PR stain data were acquired from the pathologic report. The staining studies were scored using the Allred score (AS), a method that semi-quantitates the proportion of positive cells (scored on a 0 to 5 scale) and staining intensity (scored on a 0 to 3 scale), with a maximum score of 8; an AS > 2 was considered positive [14]. The CK5/6, EGFR and HER2 immunohistochemical results were from TMA and considered positive with the following criteria in at least one core. CK5/6 stains were considered positive if any cytoplasmic and/or membranal staining was observed. Immunostaining for EGFR was interpreted as positive when at least 10% of the tumor cells showed moderate to strong membranal staining [15]. HER2 positivity was defined as an intensity of 3+ by IHC or as gene amplification ratio of ≥ 2.0 by FISH in the case of an intensity of 1+ or 2+ by IHC [16].

Definition of breast cancer subtypes by immunohistochemistry

The immunohistochemical surrogate panel (ER, PR, HER2, EGFR, and CK5/6) used to define the breast cancer subtypes has been previously published [2, 17]. In this study, we used two subtyping schemes. Each case was classified as one of five IHC-based subtypes: luminal A (ER+ and/or PR+, HER2-), luminal B (ER+ and/or PR+, HER2+), HER2 (ER-, PR-, and HER2+), and TNBC (ER-, PR-, and HER2-) [17]. TNBCs were further divided into BLBCs and QNBC/5NPs according to the basal-markers. TNBC expressing either EGFR or CK5/6 was defined as BLBC (ER-, PR-, HER2-, CK5/6+, and/or EGFR+). Breast tumors which were TN and expressed neither CK5/6 nor EGFR were defined as ‘quintuple-negative breast cancer’ (QNBC/5NP) (ER-, PR-, HER2-, CK5/6-, and EGFR-).

Statistical Analysis

Disease free survival (DFS) was defined as the time from the date of diagnosis to the date of the documentation of relapse, including locoregional recurrence and/or distant metastasis. Overall survival (OS) was expressed as the number of months from diagnosis to the date of death. Differences in the frequencies of basic characteristics, clinical parameters, and subtypes were statistically analyzed using the chi-square test, or Fisher’s exact test in the case of less than five expected cases. For multiple statistical comparisons, chi-square test was corrected by Bonferroni’s correction. Survival curves were constructed using the Kaplan-Meier method, and the log-rank test was used to compare mean survival rates across subtypes. For multivariate analysis, Cox regression models were built to estimate the adjusted hazard ratios (HRs) of breast cancer subtypes with tumor size, lymph node involvement and adjuvant chemotherapy. To test the statistical significance between model 1 (5-subgrouping) and model 2 (4-subgrouping), a likelihood ratio test of the differences was used. The null

hypothesis was that the model 2 did not describe survival differently than model 1. Statistical significance was defined as $P < 0.05$. All statistical analyses were performed using SPSS 15.0 and SAS 9.1 statistical software packages.

Microarray Analysis

We classified TNBC into BLBC and QNBC/5NP with two public independent gene expression datasets, Vijver et al.(316 samples) and Wang et al.(286 samples) [18-20] Vijver et al. was generated with 2-color oligo chips (Agilent, Hu25K) and Wang et al. was 1-color oligo chips (Affymetrix, U133X3P). Each dataset consists of a large number of random breast cancer patients. Vijver et al. is available at <http://www.rii.com/publications/2002/nejm.html>, and Wang et al. can be downloaded from the NCBI GEO data repository (GSE2034). For Wang et al, gene expression values were centered by subtracting the mean value of each probe set across the samples from each measured value. Both datasets included only ER IHC information, so the other four IHC results (PR, HER2, CK5/6 and EGFR) were dichotomized into 'positive (+)' and 'negative (-)' by the mRNA expression levels of the corresponding genes on the microarray chips. Under the assumption that Cheang et al.'s cohort with 4,046 breast samples was representative of a random sample of breast cancer population and the proportion of ER IHC result (ER+ : 70.5%) was similar to the ER IHC results (Vijver et al.: 76%, Wang et al.: 72%) in the selected microarray data sets, we used the proportion of the status of the IHC results of each marker in his dataset to determine the cut-off for the surrogate mRNA expression for the selected microarray datasets [13]. For instance, the cut-off for KRT5 which corresponds to CK5/6 was determined at the point where the proportion of '+' to '-' was the same as the proportion of 'CK5/6 +' to 'CK5/6-' in Cheang et al.'s IHC results (Additional File 1, Figure S1). The cut-offs for PGR, KRT5 and EGFR were

determined by synchronizing the proportion of their statuses in Cheang et al.'s IHC results and in each microarray dataset (Additional File 1, Figure S1). The cut-off for ERBB2 was determined from the clear bimodal distribution, by assigning '+' for right side and '-' for left side (Additional File 1, Figure S1). The only available IHC result, ER, was not replaced by the expression of ESR1. Each sample from the microarray datasets was assigned to one of the five subtypes according to the status of the five markers.

Results

Patients Characteristics

After excluding 339 (26.3% (339/1290)) cases due to failure of staining, 951 cases that had informative immunohistochemical results were included in the study. The median age in the study population at diagnosis was 47 years (range, 20–80 years). The clinicopathologic characteristics of patients are summarized in Table 1. Ductal histology was the most prevalent breast cancer type, and was present in 91.6% (872/951). Lobular histology was present in 2.7% (26/951). The remaining 5.7% (53/951) had cancers of other histological types, including mucinous, tubular, medullary, and metaplastic types. Mastectomy was performed in 62.3% (592/951), and 37.7% (359/951) underwent breast conserving surgery. Out of the 951 patients, 83.7% (796/951) received adjuvant chemotherapy; 491 were treated with nonanthracycline-based chemotherapy - CMF (cyclophosphamide, methotrexate, 5-fluorouracil) and 194 were treated with anthracycline-based chemotherapy - AC (doxorubicin, cyclophosphamide) and FAC (5-fluorouracil, doxorubicin, cyclophosphamide). Remaining 155 patients (16.3% (155/951)) did not receive any adjuvant systemic chemotherapy. The median follow-up time was 75.0 months (from 2.47 to 152.1 months).

Clinicopathologic characteristics of breast cancer subtypes

51.1% (486/951) were luminal A subtype, 12.9% (123/951) were luminal B subtype, 11.9% (113/951) were HER2 subtype, and 24.1% (229/951) were TN-subtype (Table 1). Among the TN-subtype, 139 were defined as BLBCs (14.6% (139/951)), whereas 90 cases were defined as QNBC/5NPs 90 (9.5% (90/951)). The clinicopathologic characteristics of each breast cancer subtype are shown in Table 1. Women with luminal A subtypes showed lower nuclear and histologic grade. Women with tumors of the TNBCs were younger, and

less often had lymph node involvement, but had higher nuclear and poorer histological grades. When the TNBCs were subdivided into BLBCs and QNBC/5NPs according to the basal-markers, women with BLBCs maintained the features of TNBCs, showing younger age, less often lymph node involvement, higher nuclear and poorer histological grades. 8 cases out of 10 medullary carcinomas were assigned to BLBCs and 5 out of 6 metaplastic carcinomas were to TNBCs (3 in BLBCs and 2 in QNBC/5NPs). Histologic features of representative case of QNBC/5NPs and BLBCs are described (Additional File 1, Figure S2).

Disease free survival and overall survival by breast cancer subtypes

DFS time ranged from 0 to 152.1 months with median of 67.5 months. During the study period, 256 women (26.9% (256/951)) had local recurrence and/or metastasis. HR and 95% confidence intervals (95% CI) for DFS according to breast cancer subtypes are shown in Table 2 and survival analyses were demonstrated in Figure 1A and 1C. Advanced stage did not significantly increase the recurrence risk of luminal B and HER2 subtype (Table 2, Additional File 1, Figure S3). Interestingly, BLBCs showed obvious DFS benefit with adjuvant chemotherapy (HR, 0.24; 95% CI, 0.10-0.60; P=0.002), whereas TNBCs did not show statistical benefit from chemotherapy.

OS time ranged from 2.47 to 152.1 months with median of 75.0 months. During the study period, 137 women (14.4% (137/951)) died and 814 were alive at the end of the study. HR and 95% CI for OS are shown in Table 3 and survival analyses were demonstrated in Figure 1B and 1D. Tumor size, lymph node involvement, and AJCC stage were significant prognostic factors in the analysis with all cases of breast cancers but the significance were maintained only in luminal A, TNBCs (including BLBCs and QNBC/5NPs) after subgroup analysis with molecular subtyping. Remarkably, OS in BLBCs and QNBC/5NPs, and TNBCs

showed a dramatic decrease in stage III (Table 3, Additional File 1, Figure S3). Adjuvant chemotherapy was associated with prolonged OS in Luminal B (HR, 0.23; 95% CI, 0.09-0.52; P=0.007), BLBCs (HR, 0.17; 95% CI, 0.07-0.43; P <0.001), and TNBCs (HR, 0.41; 95% CI, 0.18-0.93; P =0.06).

Correlation between adjusted variables and survival

We also performed multivariate analysis for DFS and OS including statistically significant variables (Table 4). Although the risk of recurrence was not significant in all subtypes, the HER2 and luminal B subtype was associated with recurrence. Compared to Luminal A, all other molecular subtypes were associated with a worse OS. From the survival analyses with DFS and OS, HER2 subtype was the worst prognostic subtype among four-or five subtypes (Table 4) (HR, 3.07, 95% CI, 1.86-4.87; P <0.001). The risk of death in women with luminal B subtypes was increased 1.86-fold (95% CI, 1.11-3.11; P=0.018), with the BLBCs the risk was increased 2.80-fold (95% CI, 1.68-4.68; P <0.001), with the QNBC/5NPs the risk was increased 3.04-fold (95% CI, 1.75-5.26; P <0.001), and with the TNBCs the risk was increased 2.91-fold (95% CI, 1.88-4.48; P <0.001) compared to those with luminal A breast cancer. The likelihood ratio test between two Cox models was significant for DFS (P=0.020) and OS (P=0.016), indicating that adding model 1 (5-subgrouping) significantly improved the goodness of fit of the model for survival analyses in breast cancer patients. We showed that BLBC was not associated with worse prognosis than QNBC/5NP from the multivariate analyses for DFS and OS of TNBC patients (Table 5).

Chemotherapy effects on the subtypes

When the patients were divided into two groups depending on the adjuvant chemotherapy, the survival analysis revealed that patients with luminal B and TNBCs had overall survival benefit from chemotherapy ($P < 0.001$ for luminal B, $P = 0.027$ for TNBCs) (Figure 2). From the five-subtype analysis, the BLBCs who did not receive adjuvant chemotherapy had a shortest DFS and OS showing dramatic survival gain after chemotherapy ($P = 0.001$ for DFS, $P < 0.001$ for OS) (Figure 3). On the other hand, QNBC/5NP did not show chemotherapy benefit in both DFS and OS. However, this was difficult to interpret due to small numbers etc (only 1 event in the QNBC/5NP group) (Figure 3). Furthermore, multivariate analysis of BLBC patients for DFS and OS including tumor size and lymph node involvement confirmed survival gain of adjuvant chemotherapy (Table 6). The survival gain with anthracycline-based and non-anthracycline-based adjuvant chemotherapy did not show statistical difference in BLBCs (data not shown).

Molecular subtyping in public microarray datasets

In order to support our data that TNBC is substantially different from BLBC, we explored the mRNA expression levels of the five genes (ESR1, PGR, ERBB2, KRT5 and EGFR) according to five subtypes in two large-sized microarray datasets. We classified each dataset into five subtypes from the definition of subtypes described in Methods. From the classification, the proportion of five subtypes in the two microarray datasets turned out to be very similar, and it was similar to the proportion of subtypes in Cheang et al.'s data set as well (Fig 4A). The proportion of QNBC/5NP within TNBC was about 36% and 50% in each microarray dataset, and our data set showed similar proportion which was about 40%. The expression patterns of each gene in five subtypes were consistent across the datasets displaying the clear characteristics of each subtype. Elevated expression of KRT5 and EGFR

was a distinct feature of BLBC discriminating it from the other subtypes, and their high expression in BLBC were independent from the expression of other three genes - ESR1, PGR and ERBB2 - which showed no difference between BLBC and QNBC/5NP (Figure 4B and C). KRT5 showed a wide range of expression level across samples, while the expression range of EGFR was relatively very narrow in both datasets, and most of the TNBC samples showing high expression of KRT5 accompanied high expression of EGFR (data not shown). From the overall expression patterns of five genes, KRT5 and EGFR seemed to be one of the major causes of the heterogeneity of TNBC, and a large portion of TNBC was non-BLBC subtypes supporting that TNBC should not be replaced interchangeably with BLBC.

Discussion

In summary, with 951 breast cancer cases, we observed that TNBCs showed a poor OS prognosis, showing higher nuclear and/or histological grade, next to the HER2 subtypes. The worse OS was observed in QNBC/5NPs among TNBCs. In previous studies, the frequency of TNBCs and BLBCs has ranged from 17.1% to 30.5% and 8.0% to 55.7% respectively, depending upon the definition or criteria used [21]. The difference of the frequency might be explained with the method and criteria of the IHC interpretation. In this study, the proportion of BLBCs was 14.6% (60.7% of TNBCs), which was similar with that (14.7%) in the previous Kim et al.'s study, which used four basal makers (CK5, CK14, EGFR, and c-kit) [15]. We used two basal markers (CK5/6, EGFR) in accordance with the Carolina Breast Cancer Study [17].

When the TNBCs tumors were divided into BLBCs and QNBC/5NPs based upon basal markers, QNBC/5NP had the poorer OS in our study. The Luminal A and BLBCs had the best DFS. Previous studies have shown that the expression of basal markers (basal cytokeratins and EGFR) in TNBCs also correlates with a worse prognosis and identifies a clinically distinct subgroup within the TNBCs [9, 10, 13]. In our study, CK5/6 was a poor prognostic marker but EGFR was not (data not shown). In Cheang et al.'s data from 3,744 cases, 17% were TNBCs and 9% were BLBCs, using the five-marker method [13]. They showed that the poor prognosis of the TNBCs is conferred almost entirely by those tumors positive for basal markers, that is, BLBCs. In our study, the clinicopathologic characteristics of TNBCs were maintained in BLBCs after dividing into two groups (BLBCs and QNBC/5NPs), but the poor prognosis of the TNBCs was primarily due to the QNBC/5NPs.

Without chemotherapy, BLBCs were worst prognostic subtype in both DFS and OS. Interestingly, BLBCs showed dramatic clinical benefit from both anthracycline-based and

CMF adjuvant chemotherapy. On the other hand, adjuvant chemotherapy showed little clinical benefit for QNBC/5NP in this study. However, this result simply reflects that some QNBC/5NP patients may not benefit from standard chemotherapy. Several studies of breast cancer patients treated with neoadjuvant anthracycline-based chemotherapy have reported higher response rates in TNBCs compared to luminal breast cancers [22, 23]. In an adjuvant setting, the HR for relapse or death among patients with BLBCs (HR, 0.54; 95% CI, 0.27–1.08) was reduced after treatment with anthracycline-based chemotherapy [24]. Along with those results, our study implies that the basal markers may be useful to identify TNBC patients who are most likely to benefit from anthracycline-containing adjuvant chemotherapy.

The finding that QNBC/5NPs do worse with regard to DFS and OS than BLBCs is different from the findings by Cheang et al. [13], Carey et al. [17] and Rakha et al. [8]. This difference might be due to the different proportion of breast cancer patients with no chemotherapy (Cheang et al: BLBCs treated/untreated: 103/179, here: 129/10). Furthermore, our results could be explained by the fact that BLBCs have been reported to benefit from chemotherapy [24]. Similarly the finding that BLBCs do just as well on CMF as on anthracycline-based chemotherapy, while QNBC/5NPs do worse on chemotherapy contradicts with the findings of Cheang et al. This discrepancy could be drawn from the different cohort, especially Asian population in this study, or small patient numbers used in this study.

Since 2001, breast cancer has been the most common cancer in women in Korea [25]. Recently, Rhee et al. reported that, in node-negative breast cancers, TNBC has a higher relapse rate and more aggressive clinicopathologic characteristics than non-TNBC [26]. However, the author did not divide the cases into five subtypes including BLBCs. Kim et al. reported the clinicopathologic significance of the BLBCs based on the expression of basal cytokeratins [15] indicating that BLBCs were associated with high histological and/or nuclear

grades, but no statistically significant survival differences were evident between BLBCs and those of other subtypes. BLBCs and/or TNBCs have been known to be associated with BRCA mutations, which have been investigated mostly in western countries. In the contrary, in Korea, the mutation rates of BRCA1 and 2 are extremely low, which implies that most BLBCs or TNBCs are not associated with BRCA mutations. In the population-based Carolina Breast Cancer Study, the prevalence of the BLBCs and luminal A subtypes was strongly influenced by race and menopausal status [17]. Some different tumorigenic mechanism may be involved in BLBCs or TNBCs breast cancers in Asia.

This study was limited in that the cohort is derived from a single institution in a large city (Seoul, Korea). Considering protein profiling by IHC is not homogenous in whole sections and the results of IHC from the TMAs usually give the lower frequencies than from the whole sections, BLBCs by TMA-based IHC can be contaminated with luminal A subtypes of relatively high proportion and their prognosis can be a little bit better than using whole sections. To rule out this possibility, ER and PR status were determined by assessment of whole sections of the tumor, not from cores as in TMA, because the patient was treated with anti-hormonal therapy based upon the result from whole sections. The remaining IHC results were from the TMA setting with two copies of 2-mm core to represent above at least 98% of the whole lesion (theoretical area is 8π), since usual one 1-mm core (theoretical area is π) represents 91% of the whole lesion and three cores (theoretical area is 3π) do 98% [27]. In this study, positive rate of ER and/or PR might be a little bit higher than the studies that are evaluated in TMA, implying that luminal A or B subtypes could be more prevalent in this study.

The QNBC/5NPs was observed in about 6~8% of total samples from publicly available breast cancer microarray datasets, and about 40% of TNBCs were non-BLBCs, which

supports that TNBCs is not synonymous with BLBCs. TNBCs have the poorer associated prognosis of breast cancer subtypes, a finding which is consistent with earlier studies. By adding EGFR and CK5/6 as basal markers, QNBC/5NPs were found to have the worse outcomes among TNBCs. Kreike et al. showed that, based on the gene-expression profiling, BL-subtype tumors (classified as TN tumors in their study) are heterogeneous and can be subdivided into at least five distinct subtypes [7]. Bertucci et al. called for caution in the interpretation of ongoing trials and the selection of patients for future trials based on their data, which showed that TNBCs represent a more heterogeneous group than BLBCs since TNBCs include basal and non-basal tumors, which are very different at the histological and molecular levels, most notably in regard to mRNA expression of molecules targeted by specific therapies that are under evaluation in clinical trials [28]. The risk of regarding the two tumor types as the same is that TNBCs included in clinical trials may not be identical to BLBCs, possibly leading to a falsely negative conclusion of unresponsiveness to drugs that may actually be capable of treating true BLBCs [28]. Bertucci et al. also revealed that BLBCs are a more homogenous group than TNBCs. The incomplete concordance between BLBCs and TNBCs has been reported by using various immunohistochemical definitions, including the ER-, HER2-, EGFR+, and/or CK5/6+ immunohistochemical profile, which is currently considered to be the most reliable definition [2]. In the majority of previous studies, the basal status of many TNBCs was thought to confer a poorer clinical outcome when compared to non-basal TNBCs; however, our results suggest that the QNBC/5NPs may be primarily responsible for conferring poor clinical outcomes.

Conclusions

In conclusion, TNBCs and BLBCs were associated with high histologic and/or nuclear grades. However, in contrast to previous data, QNBC/5NPs showed worse prognosis in this study, which might be from the poor response to the adjuvant chemotherapy. Based on these findings, the prognostication and the identification of subgroup of patients for therapeutic classification should be re-considered with more applicable markers. Furthermore, our study suggests that patients with QNBC/5NP might not benefit from anthracycline-based chemotherapy.

Competing interests

The authors declare that they have no competing interests

Authors' contributions

YLC and EO: designed the study, carried out the experiments, interpreted the data and wrote the manuscript; SP, YHP, EYC and YHL: critically revised the manuscript; JSC, YK and YCH: analyzed the immunohistochemical data and drafting of the manuscript; KS: carried out the experiments and data acquisition; JEL, JHK, SJN, JHY: participated in design and data collection of the data; YKS: participated in study design and coordination, data analysis, data interpretation and drafting of the manuscript. All authors read and approved the final manuscript.

Acknowledgments

We are very grateful to Jung Sun Lee for her technical assistance, which included slide cutting and immunohistochemical analysis and to Sook Young Woo (Samsung Biomedical Research Institute) for statistical support. This work was supported by the Korea Science and Engineering Foundation (KOSEF) grant funded by the Korea government (MOST) (M10729000029-08N2900-02910 and No.20100016847).

References

1. Sorlie, T., et al., *Repeated observation of breast tumor subtypes in independent gene expression data sets*. Proc Natl Acad Sci U S A, 2003. 100(14): p. 8418–23.
2. Nielsen, T.O., et al., *Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma*. Clin Cancer Res, 2004. 10(16): p. 5367–74.
3. Irvin, W.J., Jr. and L.A. Carey, *What is triple-negative breast cancer?* Eur J Cancer, 2008. 44(18): p. 2799–2805.
4. Nishimura, R. and N. Arima, *Is triple negative a prognostic factor in breast cancer?* Breast Cancer, 2008. 15(4): p. 303–8.
5. Silver, D.P., et al., *Efficacy of neoadjuvant Cisplatin in triple-negative breast cancer*. J Clin Oncol. 28(7): p. 1145–53.
6. Rakha, E.A., J.S. Reis-Filho, and I.O. Ellis, *Basal-like breast cancer: a critical review*. J Clin Oncol, 2008. 26(15): p. 2568–81.
7. Kreike, B., et al., *Gene expression profiling and histopathological characterization of triple-negative/basal-like breast carcinomas*. Breast Cancer Res, 2007. 9(5): p. R65.
8. Rakha, E., I. Ellis, and J. Reis-Filho, *Are triple-negative and basal-like breast cancer synonymous?* Clin Cancer Res, 2008. 14(2): p. 618; author reply 618–9.
9. Tischkowitz, M., et al., *Use of immunohistochemical markers can refine prognosis in triple negative breast cancer*. BMC Cancer, 2007. 7: p. 134.
10. Tan, D.S., et al., *Triple negative breast cancer: molecular profiling and prognostic impact in adjuvant anthracycline-treated patients*. Breast Cancer Res Treat, 2008. 111(1): p. 27–44.
11. Rakha, E.A., et al., *Prognostic markers in triple-negative breast cancer*. Cancer, 2007. 109(1): p. 25–32.
12. Rakha, E.A., et al., *Triple-negative breast cancer: distinguishing between basal and nonbasal subtypes*. Clin Cancer Res, 2009. 15(7): p. 2302–10.
13. Cheang, M.C., et al., *Basal-like breast cancer defined by five biomarkers has superior prognostic value than triple-negative phenotype*. Clin Cancer Res, 2008. 14(5): p. 1368–76.
14. Harvey, J.M., et al., *Estrogen receptor status by immunohistochemistry is superior to the ligand-binding assay for predicting response to adjuvant endocrine therapy in breast cancer*. J Clin Oncol, 1999. 17(5): p. 1474–81.
15. Kim, M.J., et al., *Clinicopathologic significance of the basal-like subtype of breast cancer: a comparison with hormone receptor and Her2/neu-overexpressing phenotypes*. Hum Pathol, 2006. 37(9): p. 1217–26.

16. Yaziji, H., et al., *HER-2 testing in breast cancer using parallel tissue-based methods*. *Jama*, 2004. 291(16): p. 1972-7.
17. Carey, L.A., et al., *Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study*. *JAMA*, 2006. 295(21): p. 2492-502.
18. van 't Veer, L.J., et al., *Gene expression profiling predicts clinical outcome of breast cancer*. *Nature*, 2002. 415(6871): p. 530-6.
19. van de Vijver, M.J., et al., *A gene-expression signature as a predictor of survival in breast cancer*. *N Engl J Med*, 2002. 347(25): p. 1999-2009.
20. Wang, Y., et al., *Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer*. *Lancet*, 2005. 365(9460): p. 671-9.
21. Fadare, O. and F.A. Tavassoli, *The phenotypic spectrum of basal-like breast cancers: a critical appraisal*. *Adv Anat Pathol*, 2007. 14(5): p. 358-73.
22. Carey, L.A., et al., *The triple negative paradox: primary tumor chemosensitivity of breast cancer subtypes*. *Clin Cancer Res*, 2007. 13(8): p. 2329-34.
23. Rouzier, R., et al., *Breast cancer molecular subtypes respond differently to preoperative chemotherapy*. *Clin Cancer Res*, 2005. 11(16): p. 5678-85.
24. Conforti, R., et al., *Breast cancer molecular subclassification and estrogen receptor expression to predict efficacy of adjuvant anthracyclines-based chemotherapy: a biomarker study from two randomized trials*. *Ann Oncol*, 2007. 18(9): p. 1477-83.
25. Ahn, S.H., *Clinical characteristics of breast cancer patients in Korea in 2000*. *Arch Surg*, 2004. 139(1): p. 27-30; discussion 31.
26. Rhee, J., et al., *The clinicopathologic characteristics and prognostic significance of triple-negativity in node-negative breast cancer*. *BMC Cancer*, 2008. 8: p. 307.
27. Rosen, D.G., et al., *Validation of tissue microarray technology in ovarian carcinoma*. *Mod Pathol*, 2004. 17(7): p. 790-7.
28. Bertucci, F., et al., *How basal are triple-negative breast cancers?* *Int J Cancer*, 2008. 123(1): p. 236-40.

Figure Legends

Figure 1. Kaplan-Meier curves of disease-free survival and overall survival

Disease-free survival according to (A) four and (C) five IHC-based subtypes of breast cancers in the study subjects. Overall survival according to (B) four and (D) five subtypes of breast cancers in the study subjects (Estimated mean survival with 95% CI).

Figure 2. Association between chemotherapy and four-subtypes of breast cancers.

Kaplan-Meier curves of disease-free survival (A and C) and overall survival (B and D). Cases without chemotherapy (A and B) and with chemotherapy (C and D) (Estimated mean survival with 95% CI). *P<0.001, **P=0.027

Figure 3. Association between chemotherapy and five-subtypes of breast cancers.

Kaplan-Meier curves of disease-free survival (A and C) and overall survival (B and D). Cases without chemotherapy (A and B) and with chemotherapy (C and D) (Estimated mean survival with 95% CI). *P =0.001, **P <0.001

Figure 4. Comparison of the expression levels of five genes according to five subtypes in the microarray data sets.

(A) Proportion of five breast cancer subtypes in two large-sized microarray data sets. (B) Expression levels of the five markers according to the subtypes in Vijver et al. (C) Expression levels of the five markers according to the subtypes in Wang et al. Distribution of expression levels of each marker was described with box plots. The bottom and top of the box are the 25th and 75th percentile, and the band near the middle of the box is the 50th percentile.

Table 1. Clinicopathologic characteristics of breast cancer subtypes

Variables	Molecular Subtyping													P value between four subtypes ^{***,†}					
	All												P value between five subtypes ^{*,†}						
	IHC-Luminal A			IHC-Luminal B			IHC-HER2		IHC-BLBC		IHC-QNBC/5NP				Four subtypes				
No.	(%)	N=486 51.1%	No.	(%)	N=123 12.9%	No.	(%)	N=113 11.9%	No.	(%)	N=139 14.6%	No.	(%)	N=90 9.5%	No.	(%)	No.	(%)	
Age Group (years)																			
<50	590	62.0	297	61.1	87	70.7	51	45.1	97	69.8	58	64.4	155	67.7	<0.001	<0.001			
≥50	361	38.0	189	38.9	36	29.3	62	54.9	42	30.2	32	35.6	74	32.3					
Family history of breast cancer																			
No	916	96.3	470	96.7	121	98.4	111	98.2	126	90.6	88	97.8	214	93.4					
Yes	35	3.7	16	3.3	2	1.6	2	1.8	13	9.4	2	2.2	15	6.6					
Tumor Size																			
≤ 2 cm	392	41.2	228	46.9	39	31.7	35	31.0	52	37.4	38	42.2	90	39.3					
2-5 cm	493	51.8	231	47.5	75	61.0	64	56.6	76	54.7	47	52.2	123	53.7					
> 5cm	66	6.9	27	5.6	9	7.3	14	12.4	11	7.9	5	5.6	16	7.0					
N Staging																			
N0	498	52.4	246	50.6	52	42.3	63	55.8	89	64.0	48	53.3	137	59.8					
N1	247	25.9	131	27.0	40	32.5	22	19.5	31	22.3	23	25.6	54	23.6					
N2	118	12.4	64	13.2	19	15.4	13	11.5	12	8.6	10	11.1	22	9.6					
N3	88	9.2	45	9.3	12	9.8	15	13.3	7	5.0	9	10.0	16	7.0					
AJCC stage																			
I	254	26.7	148	30.5	19	15.4	22	19.5	38	27.3	27	30.0	65	28.4					
II	467	49.1	221	45.5	69	56.1	57	50.4	78	56.1	42	46.7	120	52.4					
III	230	24.2	117	24.1	35	28.5	34	30.1	23	16.5	21	23.3	44	19.2					
LN involvement																			
Negative	498	52.4	246	50.6	52	42.3	63	55.8	89	64.0	48	53.3	137	59.8					
Positive	453	47.6	240	49.4	71	57.7	50	44.2	50	36.0	42	46.7	92	40.2					
Nuclear Grade																			
Low	92	9.7	67	13.8	5	4.1	3	2.7	9	6.5	8	8.8	17	7.4					
Intermediate	486	51.1	308	63.4	65	52.8	40	35.4	26	18.7	47	52.2	73	31.9					
High	373	39.2	111	22.8	53	43.1	70	61.9	104	74.8	35	38.9	139	60.7					
Histological Grade																			
Well	100	10.5	63	13.0	10	8.1	11	9.7	8	5.8	8	8.9	16	7.0					
Moderate	597	62.8	334	68.7	76	61.8	65	57.5	71	51.1	51	56.7	122	53.3					
Poor	254	26.7	89	18.3	37	30.1	37	32.7	60	43.2	31	34.4	91	39.7					
Estrogen Receptor (ER)																			
Negative	364	38.3	11	2.3	11	8.9	113	100.0	139	100.0	90	100.0	229	100.0					
Positive	587	61.7	475	97.7	112	91.1	0	0.0	0	0.0	0	0.0	0	0.0					

Table 2. Hazard ratios of breast cancer disease-free survival for several basic characteristics by breast cancer subtyping

Variables	Molecular subtyping												
	All											Four subtypes	
	IHC-Luminal A (N=486)			IHC-Luminal B (N=123)			IHC-HER2 (N=113)		IHC-BLBC (N=139)		IHC-QNBC/5NP (N=90)		IHC-TNBC (N=229)
HR(N of event) (95%CI)	p-value	p-value	HR(N of event) (95%CI)	p-value	p-value	HR(N of event) (95%CI)	p-value	HR(N of event) (95%CI)	p-value	HR(N of event) (95%CI)	p-value	HR(N of event) (95%CI)	
Age Group (yr)													
<50	1.00(174/590)	1.00(79/297)	1.00(32/87)	1.00(25/51)	1.00(22/97)	1.00(16/58)	1.00(6/38)	1.00(15/90)	1.00(38/155)				
≥50	0.73(82/361)	0.72(38/189)	0.88(12/36)	0.42(16/62)	0.70(7/42)	1.07(10/32)	1.04(14/76)	1.59(31/123)	0.89(17/74)				
	(0.55-0.96)	(0.49-1.07)	(0.45-1.71)	(0.21-0.79)	(0.29-1.60)	1.0	(0.87-4.27)	(0.50-1.58)	1.0				1.0
Tumor size (cm)													
≤ 2 cm	1.00(77/392)	1.00(38/228)	1.00(16/39)	1.00(8/35)	1.00(9/52)	1.00(6/38)	1.00(6/38)	1.00(15/90)	1.00(16/58)				
2-5 cm	1.62(146/493)	1.86(23/75)	0.73(23/75)	1.93(25/64)	1.04(14/76)	2.72(17/47)	1.04(14/76)	1.59(31/123)	1.07(10/32)				
	(1.23-2.14)	(1.25-2.77)	(0.38-1.38)	(0.87-4.27)	(0.45-2.39)	1.0	(0.87-4.27)	(0.86-2.94)	(0.48-2.37)				1.0
> 5cm	3.14(33/66)	3.32(12/27)	1.34(5/9)	2.65(7/14)	4.04(6/11)	4.67(3/5)	4.04(6/11)	4.30(8/16)	4.67(3/5)				1.0
	(2.08-4.72)	(1.73-6.35)	(0.49-3.67)	(0.96-7.30)	(1.43-11.22)	1.0	(0.96-7.30)	(1.90-9.82)	(1.16-18.22)				0.008
LN involvement													
Negative	1.00(85/503)	1.00(31/248)	1.00(15/53)	1.00(18/64)	1.00(12/89)	1.00(9/49)	1.00(12/89)	1.00(15/90)	1.00(16/58)				
Positive	2.54(171/448)	3.28(86/238)	1.60(29/70)	1.88(22/49)	2.75(17/50)	2.83(17/41)	2.75(17/50)	2.83(34/91)	2.83(17/41)				
	(1.96-3.30)	(2.17-4.95)	(0.86-2.98)	(1.01-3.51)	(1.32-5.77)	0.376	(1.01-3.51)	(1.64-4.89)	(1.26-6.36)				<0.001
AJCC stage													
I	1.00(36/254)	1.00(14/148)	1.00(8/19)	1.00(4/22)	1.00(5/38)	1.00(5/27)	1.00(5/38)	1.00(10/65)	1.00(5/27)				
II	1.62(106/467)	2.62(50/221)	0.55(18/69)	2.02(19/57)	1.24(13/78)	0.79(6/42)	1.24(13/78)	1.03(19/120)	0.79(6/42)				
	(1.02-2.14)	(1.45-4.74)	(0.24-1.26)	(0.68-5.94)	(0.44-3.48)	1.0	(0.68-5.94)	(0.48-2.20)	(0.24-3.48)				1.0
III	3.14(114/230)	5.86(53/117)	1.32(18/35)	3.59(17/34)	4.31(11/23)	6.30(15/21)	4.31(11/23)	5.14(26/44)	6.30(15/21)				
	(2.08-4.72)	(3.24-10.55)	(0.58-3.04)	(1.21-10.67)	(1.49-12.41)	0.176	(1.21-10.67)	(2.47-10.66)	(2.27-17.5)				<0.001
Adjuvant chemotherapy													
Not done	1.00(43/155)	1.00(17/95)	1.00(9/20)	1.00(10/22)	1.00(6/10)	1.00(1/8)	1.00(6/10)	1.00(7/18)	1.00(1/8)				
Done	0.98(213/796)	1.57(100/391)	0.63(35/103)	0.69(30/91)	0.24(23/129)	3.13(25/82)	0.24(23/129)	0.59(48/211)	3.13(25/82)				
	(0.71-1.36)	(0.94-2.64)	(0.31-1.32)	(0.34-1.42)	(0.10-0.60)	0.002	(0.34-1.42)	(0.26-1.30)	(0.43-23.2)				1.0

IHC: immunohistochemistry, LN: lymph node

Table 3. Hazard ratios of overall survival for several basic characteristics by breast cancer subtyping

Variables	Molecular subtyping											
	All				Five subtypes				Four subtypes			
	IHC-Luminal A (N=486)		IHC-Luminal B (N=123)		IHC-HER2 (N=113)		IHC-BLBC (N=139)		IHC-QNBC/5NP (N=90)		IHC-TNBC (N=229)	
HR(N of event) (95%CI)	P- value	HR(No. of event) (95%CI)	P- value	HR(No. of event) (95%CI)	P- value	HR(No. of event) (95%CI)	P- value	HR(No. of event) (95%CI)	P- value	HR(No. of event) (95%CI)	P- value	
Age Group (yr)												
<50	1.00(85/590)	1.00(22/297)	1.00(14/87)	1.00(21/51)	1.00(17/97)	1.00(11/58)	1.00(28/155)					
≥50	0.99(52/361) (0.70-1.39)	1.43(20/189) (0.78-2.61)	1.42(8/36) (0.59-3.37)	0.32(9/62) (0.15-0.71)	0.92(7/42) (0.38-2.20)	1.23(8/32) (0.49-3.04)	1.08(15/74) (0.58-2.03)	1.0	1.0	1.0	1.0	
Tumor size (cm)												
≤ 2 cm	1.00(28/392)	1.00(7/228)	1.00(5/39)	1.00(6/35)	1.00(6/52)	1.00(4/38)	1.00(10/90)					
2-5 cm	2.61(87/493) (1.70-4.00)	4.21(30/231) (1.94-10.1)	1.29(14/75) (0.46-3.61)	1.93(19/64) (0.77-4.83)	1.36(12/76) (0.51-3.64)	2.78(12/47) (0.89-8.63)	1.86(24/123) (0.88-3.88)	0.010	0.62	0.010	0.010	
> 5cm	5.34(22/66) (3.05-9.33)	6.55(5/27) (2.08-20.6)	2.17(3/9) (0.51-9.16)	2.46(5/14) (0.75-8.08)	6.65(6/11) (2.13-20.3)	7.76(3/5) (1.73-34.5)	7.54(9/16) (3.12-18.2)	0.046	0.008	0.046	<0.001	
LN involvement												
Negative	1.00(39/498)	1.00(7/246)	1.00(7/52)	1.00(12/63)	1.00(8/89)	1.00(5/48)	1.00(13/137)					
Positive	2.93(98/453) (2.03-4.24)	4.88(35/240) (2.25-10.45)	1.55(15/71) (0.63-3.81)	2.23(18/50) (1.07-4.62)	4.6(16/50) (1.73-9.49)	3.96(14/42) (1.43-11.0)	3.99(30/92) (2.09-7.35)	<0.001	0.004	<0.001	<0.001	
AJCC stage												
I	1.00(10/254)	1.00(0/148)	1.00(2/19)	1.00(3/22)	1.00(2/38)	1.00(3/27)	1.00(5/65)					
II	3.29(58/467) (1.68-6.43)	NA*(21/221)	1.09(10/69) (0.23-5.01)	1.70(13/57) (0.48-5.97)	2.74(11/78) (0.61-12.3)	0.72(3/42) (0.14-3.58)	1.58(14/120) (0.57-4.40)	1.0	1.0	1.0	1.0	
III	8.74(69/230) (4.50-16.9)	NA*(21/117) (0.57-11.9)	2.60(10/35) (0.57-11.9)	3.53(14/34) (1.01-12.34)	11.36(11/23) (2.57-51.3)	7.85(13/21) (2.23-27.5)	9.16(24/44) (3.49-24.0)	0.005	0.016	0.005	<0.001	
Adjuvant chemotherapy												
Not done	1.00(35/155)	1.00(12/95)	1.00(9/20)	1.00(7/22)	1.00(6/10)	1.00(1/8)	1.00(7/18)					
Done	0.59(102/796) (0.40-0.87)	0.74(30/391) (0.37-1.45)	0.23(13/103) (0.09-0.52)	0.75(23/91) (0.32-1.74)	0.17(18/129) (0.07-0.43)	2.23(18/82) (0.29-16.7)	0.41(36/211) (0.18-0.93)	0.06	<0.001	1.0	0.06	

* There is no event in AJCC stage I of luminal A subtype.

IHC: immunohistochemistry, LN: lymph node

Table 4. Adjusted Hazard ratio of breast cancer survival

Variables	Model 1 (5-subgrouping)				Model 2 (4-subgrouping)			
	Disease-free survival		Overall survival		Disease-free survival		Overall survival	
	HR (95%CI)	P-value	HR (95%CI)	P-value	HR (95%CI)	P-value	HR (95%CI)	P-value
Molecular subtyping								
IHC-Luminal A					1.00		1.00	
IHC-Luminal B					1.48 (1.04-2.09)	0.027	1.86 (1.11-3.11)	0.018
IHC-HER2					1.60 (1.11-2.31)	0.011	3.07 (1.86-4.87)	<0.001
IHC-TNBC					1.19 (0.89-1.68)	0.293	2.91 (1.88-4.48)	<0.001
Molecular subtyping								
IHC-Luminal A	1.00		1.00					
IHC-Luminal B	1.48 (1.04-2.09)	0.028	1.86 (1.11-3.11)	0.018				
IHC-HER2	1.60 (1.11-2.31)	0.012	3.07 (1.86-4.87)	<0.001				
IHC-BLBC	1.03 (0.68-1.56)	0.874	2.80 (1.68-4.68)	<0.001				
IHC-QNBC/5NP	1.43 (0.93-2.18)	0.102	3.04 (1.75-5.26)	<0.001				
Tumor size (cm)								
≤ 2 cm	1.00		1.00		1.00		1.00	
2-5 cm	1.39 (1.06-1.86)	0.02	2.26 (1.46-3.50)	<0.001	1.39 (1.05-1.85)	0.021	2.26 (1.46-3.39)	<0.001
> 5cm	2.19 (1.44-3.36)	<0.001	3.21 (2.13-4.67)	<0.001	2.16 (1.42-3.32)	<0.001	3.19 (1.78-5.73)	<0.001
LN involvement								
Negative	1.00		1.00		1.00		1.00	
Positive	2.40 (1.82-3.16)	<0.001	3.17 (2.14-4.61)	<0.001	2.41 (1.83-3.18)	<0.001	3.17 (2.14-4.61)	<0.001
Adjuvant Chemotherapy								
Not done	1.00		1.00		1.00		1.00	
Done	0.70 (0.50-0.99)	0.049	0.34(0.22-0.50)	<0.001	0.70 (0.50-0.99)	0.048	0.34 (0.22-0.50)	<0.001

IHC: immunohistochemistry

Table 5. Adjusted Hazard ratio of breast cancer survival in IHC-TNBC

	Disease-free survival		Overall survival	
	HR (95%CI)	P-value	HR (95%CI)	P-value
Molecular subtyping				
IHC-BLBC	1.00		1.00	
IHC-QNBC/5NP	1.43 (0.84-2.45)	0.188	1.20 (0.65-2.22)	0.558
Tumor size (cm)				
≤ 2 cm	1.00		1.00	
2-5 cm	1.44 (0.76-2.73)	0.258	1.61 (0.75-3.46)	0.219
> 5cm	3.06 (1.27-7.35)	0.012	4.07 (1.56-10.5)	0.004
LN involvement				
Negative	1.00		1.00	
Positive	2.54 (1.44-4.48)	0.001	3.83 (1.92-7.65)	<0.001
Adjuvant Chemotherapy				
Not done	1.00		1.00	
Done	0.54 (0.23-1.28)	0.161	0.31(0.13-0.76)	0.011

Table 6. Adjusted Hazard ratio of breast cancer survival in IHC-BLBC

	Disease-free survival		Overall survival	
	HR (95%CI)	P-value	HR (95%CI)	P-value
Tumor size (cm)				
≤ 2 cm	1.00		1.00	
2-5 cm	1.49 (0.51-2.61)	0.740	1.48 (0.57-3.87)	0.422
> 5cm	1.72 (0.55-5.32)	0.349	2.45 (0.73-8.24)	0.147
LN involvement				
Negative	1.00		1.00	
Positive	3.21 (1.49-6.89)	0.003	4.78 (1.95-11.73)	0.001
Adjuvant Chemotherapy				
Not done	1.00		1.00	
Done	0.34 (0.12-0.88)	0.028	0.22(0.08-0.59)	0.011

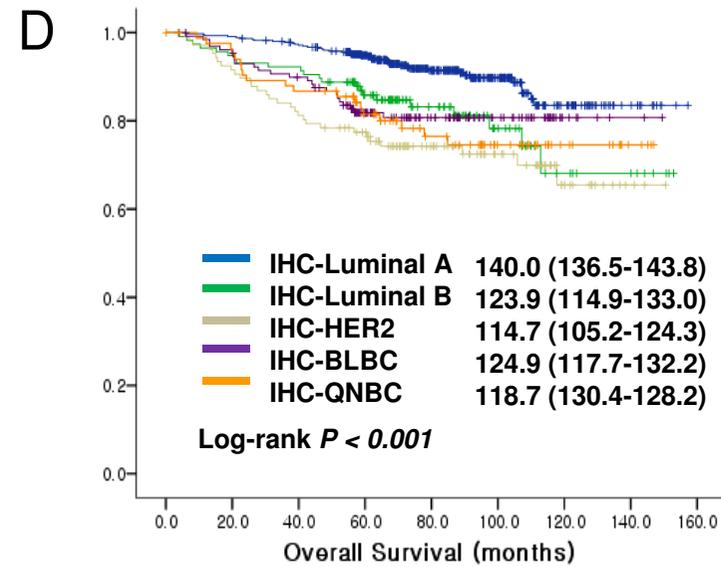
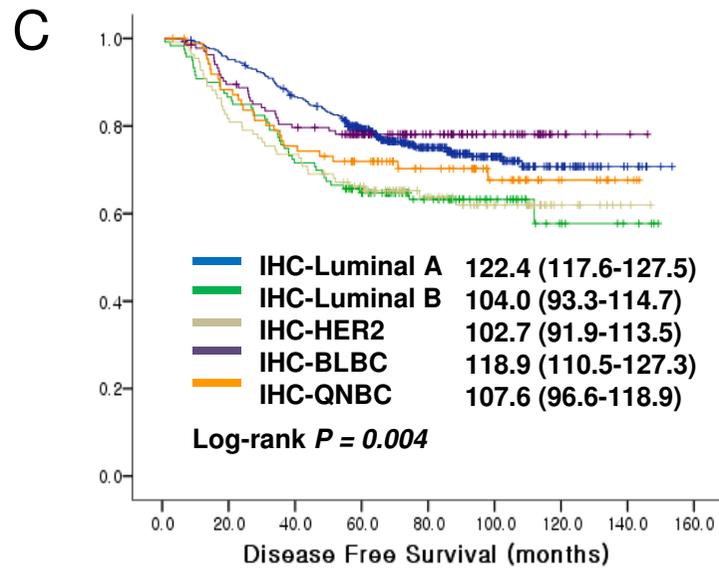
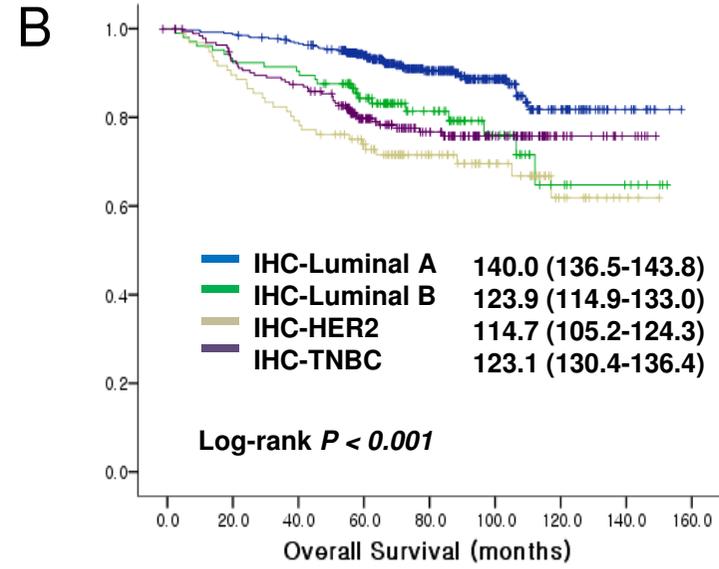
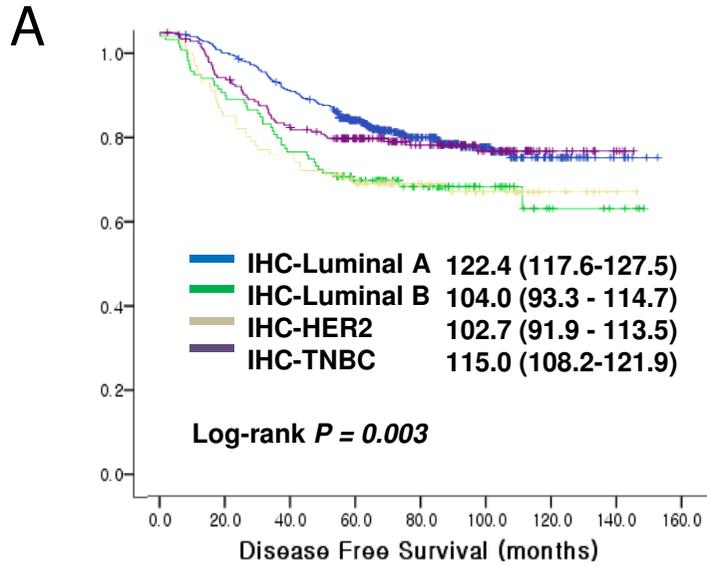
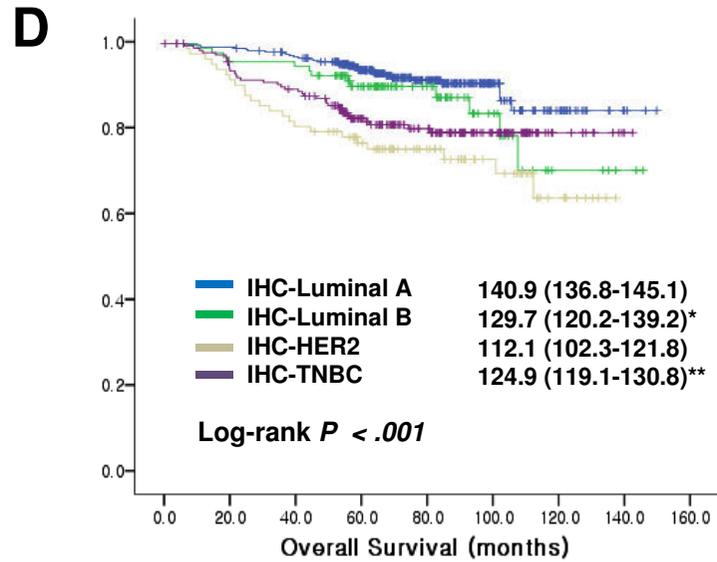
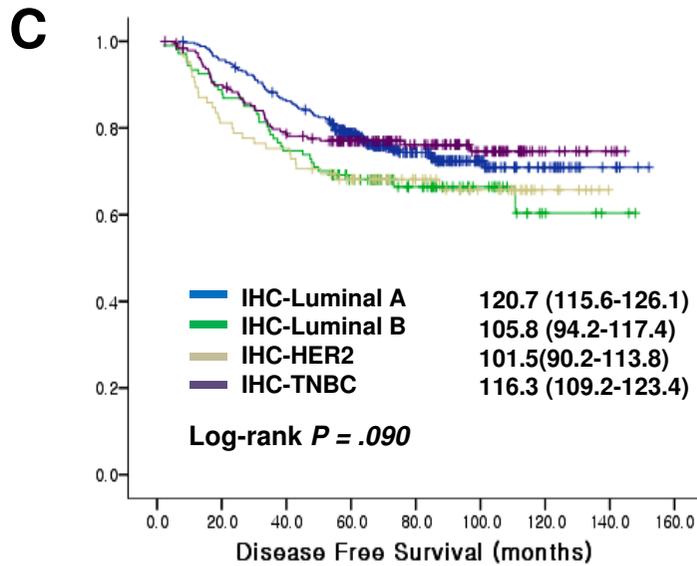
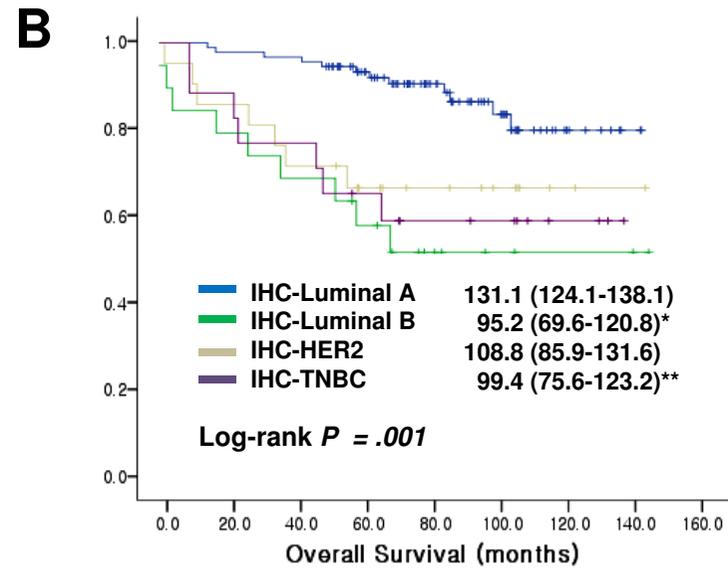
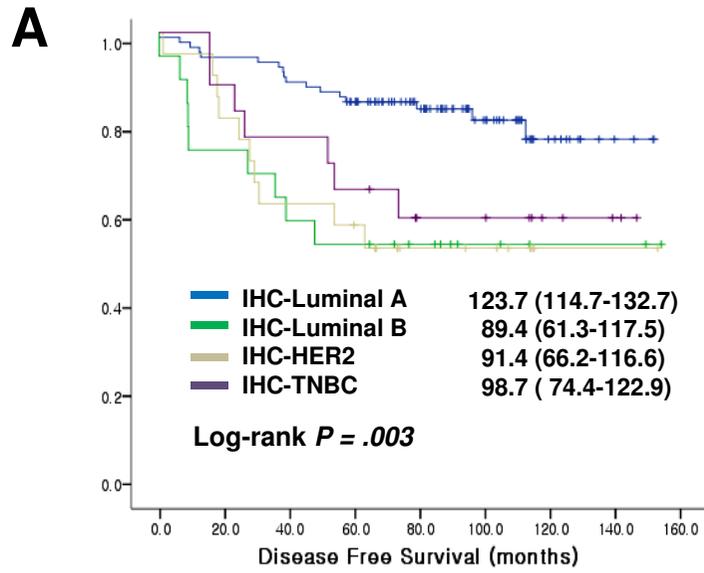


Figure 1



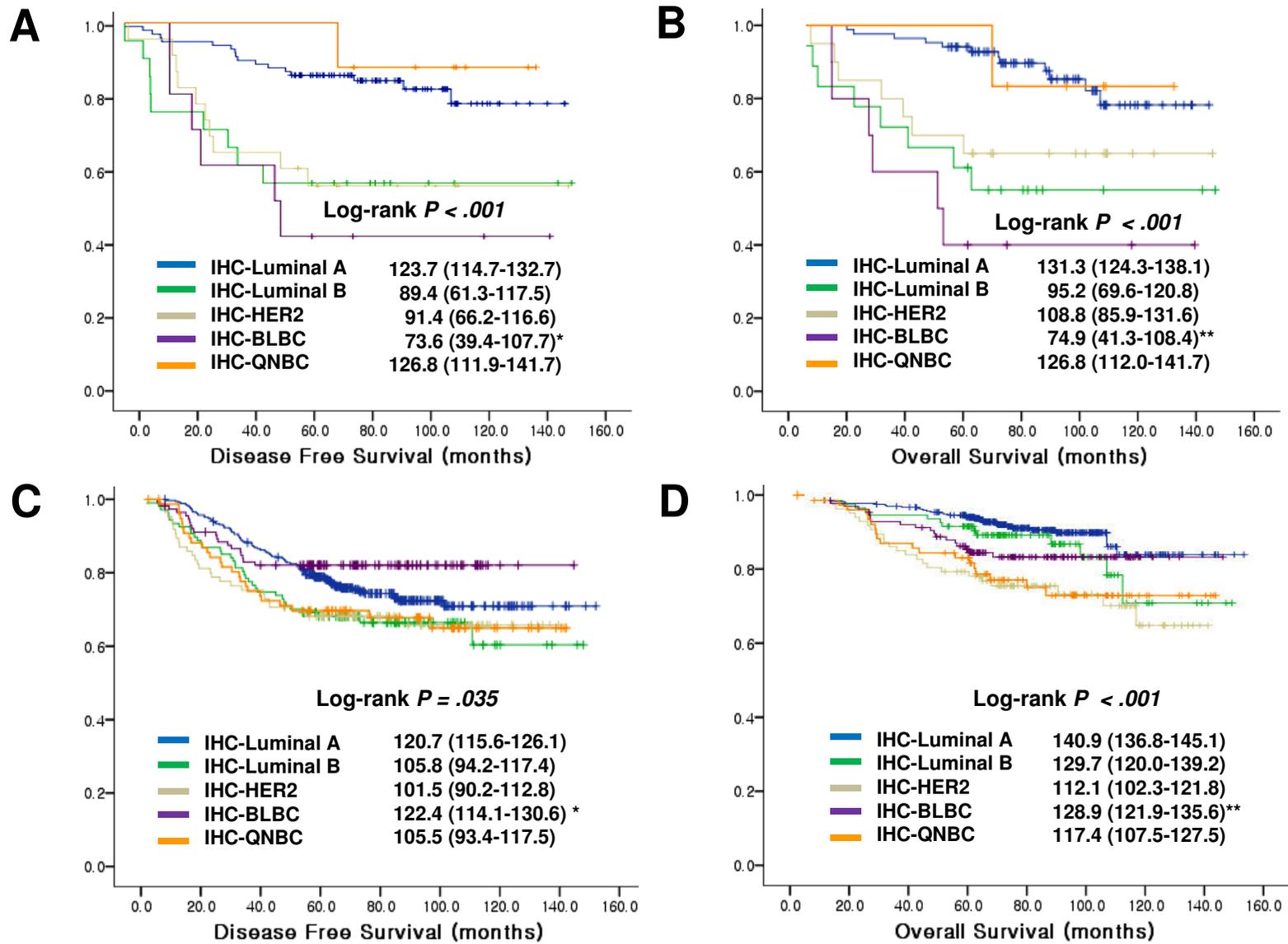
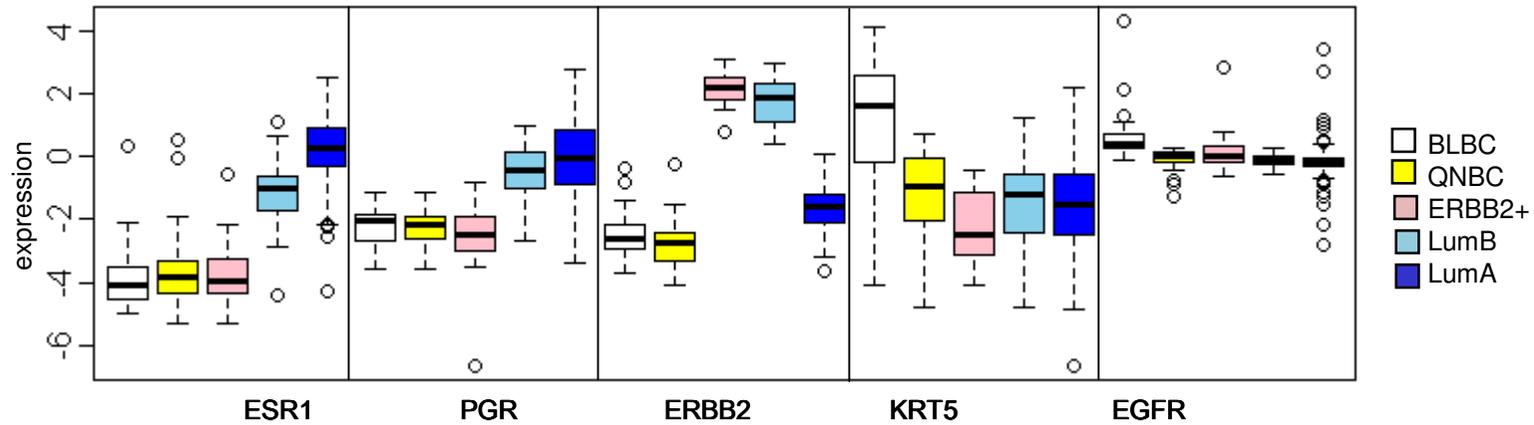


Figure 3

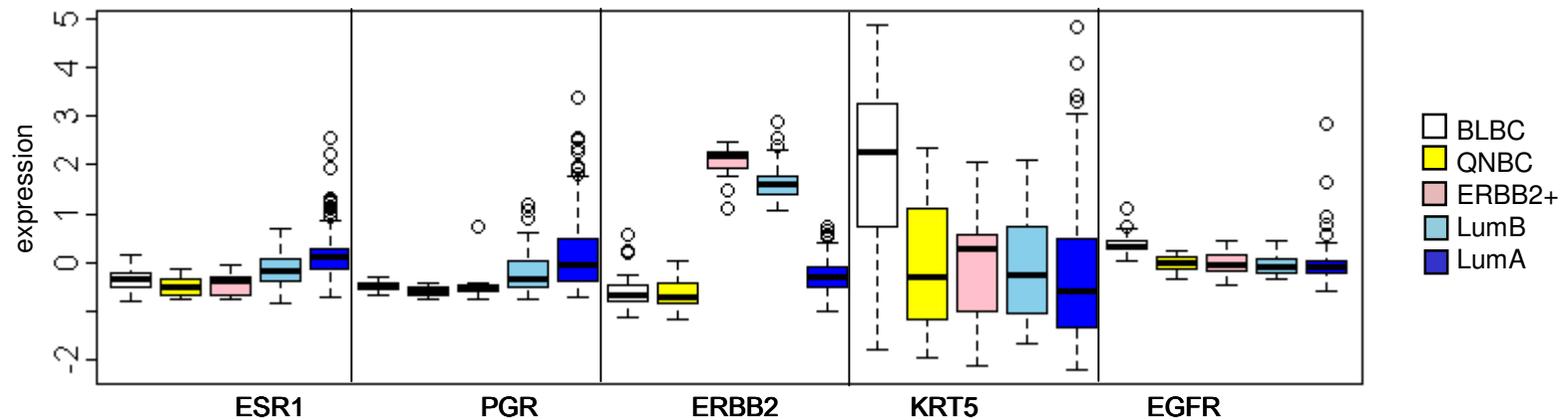
(A) Composition of subtypes in two microarray datasets

Data set	Chip	Total	Lum A	Lum B	ERBB2+	BLBC	QNBC
Vijver et al.	Agilent	316	207 (65.5%)	36 (11.4%)	19 (6.0%)	27 (8.5%)	27 (8.5%)
Wang et al.	Affymetrix	286	191 (66.8%)	29 (10.1%)	17 (5.9%)	31 (10.8%)	18 (6.3%)
Cheang et al.		3744	2625 (70.1%)	222 (5.9%)	258 (6.9%)	336 (9.0%)	303 (8.1%)

(B) Vijver et al.



(C) Wang et al.



Additional files provided with this submission:

Additional file 1: Supplemental Figures.pdf, 354K

<http://www.biomedcentral.com/imedia/9041122194545091/supp1.pdf>