The prognostic impact of age in different molecular subtypes of breast cancer: a population-based study

Dongjun Dai*1, Yiming Zhong*1, Zhuo Wang1, Neelum Aziz Yousaizai1, Hongchuan Jin2, Xian Wang* Corresponding Author: Xian Wang
Email address: wangx118@zju.edu.cn

1 Department of Medical Oncology, Sir Run Run Shaw Hospital, Medical School, Zhejiang University, Hangzhou, Zhejiang, China
2 Laboratory of Cancer Biology, Key Lab of Biotherapy, Sir Run Run Shaw Hospital, Medical School, Zhejiang University, Hangzhou, Zhejiang, China

**Background:** The aim of current study was to use competing risk model to calculate the potential differences that age played in different breast cancer subtypes.

**Methods:** The cohort was selected from Surveillance, Epidemiology, and End Results (SEER) program. The cumulative incidences of death (CID) were assessed for breast cancer caused deaths and other causes of mortality. Multivariate Cox proportional hazards regression model and multivariate subdistribution hazard (SH) model were used to evaluate the prognostic value of age in different breast cancer subtypes.

**Results:** We involved 33,968 breast cancer patients into our cohort. We found older patients had worse overall survival (OS) than young patients in hormone receptor positive and human epidermal growth factor receptor 2 positive (HR+/HER2+) (>= 40 vs. < 40, HR = 2.07, 95% CI = 1.28-3.35, p < 0.05). However, when we used competing risk model, we found young age was an independent risk factor only for triple negative breast cancer (TNBC) subtype (>= 40 vs. < 40, HR = 0.71, 95% CI = 0.56-0.89, p < 0.05). No association was found in other groups.

**Conclusion:** Our research was currently the largest sample size study and the first competing risk model-based study on the association between age and different breast cancer subtypes. We found < 40 years patients had worse breast cancer specific survival (BCSS) than older patients in TNBC subtype.
The prognostic impact of age in different molecular subtypes of breast cancer: a population-based study

Dongjun Dai¹¶, Yiming Zhong¹¶, Zhuo Wang¹, Neelum Aziz Yousafzai¹, Hongchuan Jin², Xian Wang¹,*

¹ Department of Medical Oncology, Sir Run Run Shaw Hospital, Medical School, Zhejiang University, Hangzhou, China;
² Laboratory of Cancer Biology, Key Lab of Biotherapy, Sir Run Run Shaw Hospital, Medical School, Zhejiang University, Hangzhou, China

¶ These authors contribute to this work equally.

Corresponding Author:
Xian Wang¹
Department of Medical Oncology, Sir Run Run Shaw Hospital, Medical School, Zhejiang University, 3 East Qingchun Road, Hangzhou, China
Email address: wangx118@zju.edu.cn

Abstract

Background: The aim of current study was to use competing risk model to calculate the potential differences that age played in different breast cancer subtypes.

Methods: The cohort was selected from Surveillance, Epidemiology, and End Results (SEER) program. The cumulative incidences of death (CID) were assessed for breast cancer caused deaths and other causes of mortality. Multivariate Cox proportional hazards regression model and multivariate subdistribution hazard (SH) model were used to evaluate the prognostic value of age in different breast cancer subtypes.

Results: We involved 33,968 breast cancer patients into our cohort. We found older patients had worse overall survival (OS) than young patients in hormone receptor positive and human epidermal growth factor receptor 2 positive (HR+/HER2+) (≥ 40 vs. < 40, HR = 2.07, 95%CI = 1.28-3.35, p < 0.05). However, when we used competing risk model, we found young age was an independent risk factor only for triple negative breast cancer (TNBC) subtype (≥ 40 vs. < 40, HR = 0.71, 95%CI = 0.56-0.89, p < 0.05). No association was found in other groups.
Conclusion: Our research was currently the largest sample size study and the first competing risk model-based study on the association between age and different breast cancer subtypes. We found < 40 years patients had worse breast cancer specific survival (BCSS) than older patients in TNBC subtype.

Introduction

Breast cancer is a heterogeneous disease. Subtypes of breast cancer show diverse phenotype and have different responsiveness to treatments (Perou et al. 2000). There are four main molecular subtypes of breast cancer that comprise luminal A, luminal B breast cancer, basal-like and HER2-like breast cancer. All these types of breast cancer were determined by expression of ER (estrogen receptor), PR (progesterone receptor), HER2 (human epidermal growth factor receptor 2) and proliferative marker such as Ki67 and cytokeratin CK5/6 (Reis-Filho & Pusztai 2011). Triple-negative breast cancer (TNBC) is referred to breast cancer with negative expression of ER, PR and HER2. Most of TNBC are genetically defined basal subtype (Alluri & Newman 2014). The American Cancer Society showed that in different races, the incidence rates per 100,000 of luminal A, luminal B, HER2 and TNBC subtypes range from 53-82, 11-14, 4-7 and 8-24, respectively.

Different breast cancer subtypes have different prognostic and therapeutic implications (Haque et al. 2012; Hennigs et al. 2016; Sorlie et al. 2001). Luminal A patients are often with low-grade tumors, and good prognosis (Blows et al. 2010; Perou & Borresen-Dale 2011). Luminal B, HER2 and TNBC types are widely recognized to have poorer survival and with higher grade (Tran & Bedard 2011). Hormone therapy is used to treat hormone receptors positive (HR+) breast cancer (Luminal A and Luminal B) (Abdulkareem & Zurmi 2012). HER2 inhibitor, such as trastuzumab, is usually administrated to HER2 and Luminal B breast cancers (Kim et al. 2017). For TNBC patients, chemotherapy remains the mainstay of treatment (Lebert et al. 2018).

Age at diagnosis of the patient was found to be an important prognostic factor for breast cancer (Beadle et al. 2011). Young age at diagnosis was observed to be correlated with worse prognosis (Anders et al. 2008; Anderson et al. 2009). Breast cancers in younger women were found to have lower mRNA expression of ER and PR, and higher expression of HER2 and EGFR (Anders et al. 2008). Previous study showed breast cancer in young patients was more aggressive (ER-negative and basal-like subtype) and it was more indolent (ER-positive and non-basal-like subtype) in older women breast cancer patients (Anderson et al. 2014). Whether different age frequencies could alter the prognosis in different breast cancer subtypes was currently a subject of discussion or speculation. Previous study found that in hormone receptor (HR) positive patients (ER or PR positive), the younger ones (< 40 years) had a worse relapse-
free-survival (RFS) than older ones. While there was no association between age and RFS in HER2 or TNBC subtypes (Azim et al. 2012; Liu et al. 2014). On the contrary, other studies had found < 40 years patients had a worse overall survival (OS) than older ones in TNBC subtype but not the HR+ or HER2 subtypes (Liedtke et al. 2013; Liedtke et al. 2015).

To be noted, breast cancers have a relative long survival rate, which means the OS might be influenced by other causes of death. Competing risk model considers both the disease-specific death and other causes of death, which is widely used for prognosis analysis of long survival disease (Latouche et al. 2013; Satagopan et al. 2004). There was no study using the competing risk model to calculate the prognostic value of age in different breast cancer molecular subtypes. Besides, previous studies investigated with relative small-scale samples. Surveillance, Epidemiology, and End Results (SEER) database of the National Cancer Institute is a national collaboration program of Unite State, covering almost 26% American population’s cancer incidence and survival data. To our best knowledge, there was no SEER-based study on the association between age and the prognosis in breast cancer subtypes. Hence, the current study performed a competing risk model-based analysis of age in the prognosis of breast cancer subtypes using SEER-based population.

**Material and methods**

**Patient screening**

We selected our cohort from SEER by using SEER*Stat 8.3.5 software. Since the five-year relative survival rates of breast cancer was over 90%, long-term follow-up of breast cancer patients were collected. Besides, as HER2 status was registered since 2010, we involved it with the patients diagnosed in 2010. We used the follow inclusion criteria to screen the patients: (1) it should be female primary breast cancer patients diagnosed between age 20 to 79 who had done surgery; (2) it should have records of ER, PR and HER2 status; (3) it should be unilateral invasive ductal carcinoma with specific location; (4) it should also include clinicopathological information of race, laterality, tumor location, grade, tumor size, 7th AJCC tumor stage, number of positive regional nodes, marital status, and records if the patients had radiotherapy or chemotherapy; (5) The survival time of the patient should be over 3 months, and the survival status should be recoded for survival analyses. Patients who did not meet these criteria would be excluded. To be noted, the SEER database defined breast cancer subtypes by immunohistochemistry HR and HER2 status. Hence, the molecular subtypes of breast cancer in current study was roughly defined as the follows: HR+/HER2-, HR+/HER2+, HER2 and TNBC.
**Study variables and endpoints**

We extracted the following variables from the selected cohorts that included age at diagnosis (20-39, 40-49, 50-79), race (Caucasian, African American, American Indian/Alaska Native, Asian or Pacific Islander), laterality (right or left side), tumor location, grade (well-differentiated, moderately differentiated, poorly differentiated, undifferentiated or anaplastic), tumor size, 7th AJCC tumor stage, number of positive regional nodes, marital status, and radiotherapy or chemotherapy experience. The tumor location was defined by SEER Site Specific Coding Modules ([https://seer.cancer.gov/manuals/2016/appendixc.html](https://seer.cancer.gov/manuals/2016/appendixc.html)), which comprised nipple, central portion of breast, upper-inner quadrant of breast, lower-inner quadrant of breast, upper-outer quadrant of breast, lower-outer quadrant of breast, axillary tail of breast, overlapping lesion of breast and breast that is not otherwise specified. The 7th AJCC tumor stage was roughly considered as I, II, III and IV. The widowed or single (never married or having a domestic partner) or divorced or separated patients was classified as unmarried. The value of tumor size and number of positive regional nodes were transformed into small categorical variables to fit the linear assumption. The median follow-up was estimated as the median observed survival time.

**Statistical analyses**

The cumulative incidences of death (CID) were assessed for breast cancer caused deaths and deaths from other causes. Multivariate SH model and multivariate Cox proportional hazards regression model were used to assess the breast cancer specific survival (BCSS) and OS, respectively. Hazard ratio (HR) and 95% confidence index (95%CI) were calculated. All the statistical analyses were performed by using R version 3.4.4. A p-value less than 0.05 was considered statistically significant.

**Results**

**Cohort selection**

After selection, we involved 33,968 breast cancer patients into our cohort. The distribution of breast cancer molecular subtypes and age is listed in Table 1. The age < 40 years presented with lower HR+ (72.67% vs. 82.28% and 84.61%, p < 0.001), more TNBC (20.14% vs. 13.32% and 11.28%, p < 0.001) and HER2 subgroup (7.19% vs. 4.39% and 4.12%, p < 0.001). Besides, it was worth to notice that the age < 40 years presented with more African Americans (15.13% vs. 11.86% and 9.44%, p < 0.001), higher grade (II-IV, 91.85% vs. 81.77% and 74.65%, p <
larger tumor size (> 2cm, 55.78% vs. 43.58% and 33.93%, p < 0.001), higher proportion of 7th AJCC tumor stage (II-IV, 69.53% vs. 55.48% and 43.84%, p < 0.001), more positive regional nodes (node > 0, 46.56% vs. 37.87% and 29.33%, p < 0.001), less experience of radiotherapy (50.03% vs. 54.53% and 58.62%, p < 0.001) and more treatment of chemotherapy (80.50% vs. 63.16% and 39.22%, p < 0.001). The median follow-up time was 64 months. The other causes of mortality increased with the age of patients. The death rate caused by breast cancer and other cause were 11.03% and 1.60 % for the age < 40 years, 5.74% and 1.88% for the age between 40 and 49 years, and 5.82% and 5.65% for the age over 50 years.

The CIF analysis of age in breast cancer subtype prognosis

As shown in Figure 1, competing risk model was used to evaluate the CID induced by age in different breast cancer subtypes. We found significant association between age and breast cancer subtypes in either the breast cancer caused deaths or other causes of death in the overall population and HR+ group (p < 0.001). The age < 40 years group showed a worse outcome than older age groups. The CID caused by cancer was lower in HR+ group, which might be a result that HR+ subtype had a better prognosis than HER2 or TNBC subtypes. Further subgroup analysis found age < 40 years could significantly increase the CID of HR+/HER2- (p < 0.001) and TNBC (p < 0.001). However, there was no association between age and the CID of HR+/HER2+ and HER2 group in our cohort (p > 0.05). As expected, the increased age could significantly elevate the CID of other causes of death in all groups (p < 0.001).

Multivariate analysis of age and the prognosis of breast cancer specific survival (BCSS) and OS in different breast cancer subtypes

Next, we evaluated the independent prognostic value of age in breast cancer subtypes by performing multivariate Cox proportional hazards regression model and multivariate SH model, which included all variables we had obtained. As shown in Table 2, the multivariate Cox model showed that the older patients had worse OS than young patients in HR+/HER2+ subtype (>= 40 vs. < 40, HR = 2.07, 95%CI = 1.28-3.35, p < 0.05), and there was no association between age and OS in other subtypes (p > 0.05). However, the SH model found the older patients did not show a worse prognosis in HR+/HER2+ subtype (p > 0.05), while the young patients were found to have a worse BCSS in TNBC (>= 40 vs. < 40, HR = 0.71, 95%CI = 0.56-0.89, p < 0.05), which was conflict to the multivariate Cox regression analysis. Judged from CID plot (Figure 1) of HR+/HER2+ subtype, the CID of BCSM of < 40 years group had a similar tendency with CID of other causes of deaths of > 50 years group, indicating that the worse OS of HR+/HER2+ in
older patients might be a result of other causes of deaths. While as shown in CID plot of TNBC, the CID of BCSM of < 40 years group had a much higher CID than other causes of deaths of > 50 years group, therefore, no association of age and TNBC OS might be caused by other reasons of death.

Consistent with multivariate Cox model, our SH model found no association between age and HER2 subtype of breast cancer (p > 0.05).

The association between treatment and prognosis of different breast cancer subtypes for <40 years patients.

We further evaluated the association between treatment and prognosis of different breast cancer subtypes for <40 years patients. As shown in Table 3, the radiotherapy could improve the OS and CSS in HR+/HER2- group according to Cox and SH models, respectively. Chemotherapy could also increase the OS of HR+/HER2- breast cancer patients. The radiotherapy was also found to improve the OS in HER2 group. However, this needed to be further validated because of the limited sample size of HER2 group (radiotherapy vs. no radiotherapy, HR = 0.02, 95%CI = 0.005-0.63, p < 0.05). To be noted, there was no association between the radiotherapy and TNBC prognosis, and unexpectedly, the chemotherapy would worse the CSS of TNBC patients in SH model (chemotherapy vs. no chemotherapy, HR = 6.90, 95%CI = 2.07-22.96, p < 0.05).

Discussion

Our study was currently the largest competing risk model-based study on the association between age and the prognosis in different breast cancer subtype. We observed < 40 years patients had a significant worse CID than older patients in HR+/HER2- and TNBC subtypes but not HR+/HER2+ and HER2 subtypes in breast cancer. Further Cox multivariate analyses found > 50 years patients had a worse OS than < 40 years patients in HR+/HER2+ subtypes. However, this was eliminated when we applied SH model, indicating that the worse OS in older patients was from other causes of death. Indeed, we observed the deaths rates caused by other reasons in > 50 years patients was 3 times more than the < 50 years patients. In addition, our SH model found < 40 years patients had a worse BCSS than older patients in TNBC subtypes. Moreover, we found chemotherapy would worse the CSS of TNBC patients under 40 years in SH model.

The previous study found < 40 years patients presented with more African Americans, higher grade, higher tumor stage, more positive lymph nodes than older ones(Anders et al. 2008;
Anders et al. 2009; Bharat et al. 2009), and young age patients presented with more aggressive subtypes such as HER2 and TNBC (Anders et al. 2011; Anders et al. 2008; Anderson et al. 2014; Azim et al. 2012). Our study confirmed these clinicopathologic results. The menopause transition was found to influence the ER positive rate (Tarone & Chu 2002; Yasui & Potter 1999). Age-specific rates of ER- breast cancer cease to increase after 50 years of age, while the age-specific rates of ER+ breast cancer continue to increase after 50 years of age (Tarone & Chu 2002), which might be explained by the finding that proliferation rate of ER+ cells increased with age (Shoker et al. 1999).

It is widely recognized that young breast cancer patients had worse outcomes than older ones (Anders et al. 2009; Nixon et al. 1994). The young people might not pay enough attention for breast cancer, which might lead to delay in diagnosis until worse stage has come. However, multivariate analysis found age was an independent factor associated with breast cancer (Nixon et al. 1994). Besides, younger women diagnosed with early-stage breast cancer were actually suggested to be more likely to die than older early-stage breast cancer patients (Gnerlich et al. 2009). The biological differences may distinct between the breast tumor of young and older patients. It was suggested that the incidence of breast cancer could be increased shortly after the first pregnancy (Albrektsen et al. 2005). The pregnancy could increase the incidence of aggressive ER/PR (-) breast cancer while it decreased the incidence of ER/PR (+) tumors (Britt et al. 2007; Hildreth et al. 1983; Ruder et al. 1989). Moreover, it was demonstrated that the gene expression pattern in breast cancers detected following a pregnancy was mainly attributable to TNBC, which was more prevalent in pregnancy-associated breast cancers than nulliparous group (Asztalos et al. 2015). We speculate that the gene pattern in TNBC was a key factor that related to its young age prevalence, such as BRAC1/2 mutation, which was frequently occurred in both the young patients and TNBC patients (Peshkin et al. 2010; Rosenberg et al. 2016). However, more gene patter studies are required to get new evidences. In our study, competing risk model found that the age was an independent risk factor for TNBC but not the other molecular subtypes of breast cancer, which was consistent with previous studies (Liedtke et al. 2013; Liedtke et al. 2015).

Furthermore, in patients under 40 years, we found chemotherapy would worse the CSS of TNBC patients and have no association with the OS of TNBC patients. In contrast, chemotherapy would improve the OS of HR+/HER2- patients, and radiotherapy could improve the OS and CSS of HR+/HER2- patients. The different response in the treatment of different breast cancer subtypes might partially explain the worse prognosis of <40 years TNBC patients and prove new hints for clinicians. However, the sample size of the < 40 years patients in current study was relatively small. Further larger scale study is needed for a more reliable result.
There are some limitations in our study. First, this is a retrospective study, which presented relative low level of clinical evidence. Further randomized controlled trials are required. Second, selection bias might exist as we only included patients with complete information of involved variables. Third, breast cancer subtypes are only defined by ER, PR and HER2 status in SEER database, the detailed molecular information was missing such as Ki-67 and other proliferating factors (without which the luminal A and luminal B subtypes could not be distinguished). Further detailed studies with more specific molecular classification are required.

Conclusions

In summary, the current study was the first competing risk model-based study with the largest sample size on the association between age and breast cancer subtypes. We found age < 40 years was an independent risk factor for TNBC but not the other subtypes of breast cancers.

References

Abdulkareem IH, and Zurmi IB. 2012. Review of hormonal treatment of breast cancer. Niger J Clin Pract 15:9-14. 10.4103/1119-3077.94088

Albrektsen G, Heuch I, Hansen S, and Kvale G. 2005. Breast cancer risk by age at birth, time since birth and time intervals between births: exploring interaction effects. Br J Cancer 92:167-175. 10.1038/sj.bjc.6602302

Alluri P, and Newman LA. 2014. Basal-like and triple-negative breast cancers: searching for positives among many negatives. Surg Oncol Clin N Am 23:567-577. 10.1016/j.soc.2014.03.003

Anders CK, Fan C, Parker JS, Carey LA, Blackwell KL, Klauber-DeMore N, and Perou CM. 2011. Breast carcinomas arising at a young age: unique biology or a surrogate for aggressive intrinsic subtypes? J Clin Oncol 29:e18-20. 10.1200/JCO.2010.28.9199

Anders CK, Hsu DS, Broadwater G, Acharya CR, Foekens JA, Zhang Y, Wang Y, Marcom PK, Marks JR, Febbo PG, Nevins JR, Potti A, and Blackwell KL. 2008. Young age at diagnosis correlates with worse prognosis and defines a subset of breast cancers with shared patterns of gene expression. J Clin Oncol 26:3324-3330. 10.1200/JCO.2007.14.2471

Anders CK, Johnson R, Litton J, Phillips M, and Bleyer A. 2009. Breast cancer before age 40 years. Semin Oncol 36:237-249. 10.1053/j.seminoncol.2009.03.001

Anderson SJ, Wapnr I, Dignam JJ, Fisher B, Mamounas EP, Jeong JH, Geyer CE, Jr., Wickerham DL, Costantino JP, and Wolmark N. 2009. Prognosis after ipsilateral breast tumor recurrence and locoregional recurrences in patients treated by breast-conserving
therapy in five National Surgical Adjuvant Breast and Bowel Project protocols of node-negative breast cancer. J Clin Oncol 27:2466-2473. 10.1200/JCO.2008.19.8424

Anderson WF, Rosenberg PS, Prat A, Perou CM, and Sherman ME. 2014. How many etiological subtypes of breast cancer: two, three, four, or more? J Natl Cancer Inst 106. 10.1093/jnci/dju165

Asztalos S, Pham TN, Gann PH, Hayes MK, Deaton R, Wiley EL, Emmadi R, Kajdacsy-Balla A, Banerji N, McDonald W, Khan SA, and Tonetti DA. 2015. High incidence of triple negative breast cancers following pregnancy and an associated gene expression signature. Springerplus 4:710. 10.1186/s40064-015-1512-7

Azim HA, Jr., Michiels S, Bedard PL, Singhal SK, Criscitiello C, Ignatiadis M, Haibe-Kains B, Piccart MJ, Sotiriou C, and Loi S. 2012. Elucidating prognosis and biology of breast cancer arising in young women using gene expression profiling. Clin Cancer Res 18:1341-1351. 10.1158/1078-0432.CCR-11-2599

Beadle BM, Woodward WA, and Buchholz TA. 2011. The impact of age on outcome in early-stage breast cancer. Semin Radiat Oncol 21:26-34. 10.1016/j.semradonc.2010.09.001

Bharat A, Aft RL, Gao F, and Margenthaler JA. 2009. Patient and tumor characteristics associated with increased mortality in young women (< or =40 years) with breast cancer. J Surg Oncol 100:248-251. 10.1002/jso.21268

Blows FM, Driver KE, Schmidt MK, Broeks A, van Leeuwen FE, Wesseling J, Cheang MC, Gelmon K, Nielsen TO, Blomqvist C, Heikkila P, Heikkinen T, Nevanlinna H, Akslen LA, Begin LR, Foulkes WD, Couch FJ, Wang X, Cafourek V, Olson JE, Baglietto L, Giles GG, Severi G, McLean CA, Southey MC, Rakha E, Green AR, Ellis IO, Sherman ME, Lissowska J, Anderson WF, Cox A, Cross SS, Reed MW, Provenzano E, Dawson SJ, Dunning AM, Humphreys M, Easton DF, Garcia-Closas M, Caldas C, Pharoah PD, and Huntsman D. 2010. Subtyping of breast cancer by immunohistochemistry to investigate a relationship between subtype and short and long term survival: a collaborative analysis of data for 10,159 cases from 12 studies. PLoS Med 7:e1000279. 10.1371/journal.pmed.1000279

Britt K, Ashworth A, and Smallley M. 2007. Pregnancy and the risk of breast cancer. Endocr Relat Cancer 14:907-933. 10.1677/ERC-07-0137

Gnerlich JL, Deshpande AD, Jeffe DB, Sweet A, White N, and Margenthaler JA. 2009. Elevated breast cancer mortality in women younger than age 40 years compared with older women is attributed to poorer survival in early-stage disease. J Am Coll Surg 208:341-347. 10.1016/j.jamcollsurg.2008.12.001

Haque R, Ahmed SA, Inzhakova G, Shi J, Avila C, Polikoff J, Bernstein L, Enger SM, and Press MF. 2012. Impact of breast cancer subtypes and treatment on survival: an analysis spanning two decades. Cancer Epidemiol Biomarkers Prev 21:1848-1855. 10.1158/1055-9965.EPI-12-0474

Hennigs A, Riedel F, Gondos A, Sinn P, Schirmacher P, Marne F, Jager D, Kauzcor HU, Stieber A, Lindel K, Debus J, Golatta M, Schutz F, Sohn C, Heil J, and Schneeweiss A. 2016.
314 Prognosis of breast cancer molecular subtypes in routine clinical care: A large
315 prospective cohort study. BMC Cancer 16:734. 10.1186/s12885-016-2766-3
316 Hildreth NG, Kelsey JL, Eisenfeld AJ, LiVolsi VA, Holford TR, and Fischer DB. 1983.
317 Differences in breast cancer risk factors according to the estrogen receptor level of the
318 tumor. J Natl Cancer Inst 70:1027-1031.
319 Kim J, Pareja F, Weigelt B, and Reis-Filho JS. 2017. Prediction of Trastuzumab Benefit in
320 HER2-Positive Breast Cancers: Is It in the Intrinsic Subtype? J Natl Cancer Inst 109.
321 10.1093/jnci/djw218
322 Latouche A, Allignol A, Beyersmann J, Labopin M, and Fine JP. 2013. A competing risks
323 analysis should report results on all cause-specific hazards and cumulative incidence
324 functions. J Clin Epidemiol 66:648-653. 10.1016/j.jclinepi.2012.09.017
325 Lebert JM, Lester R, Powell E, Seal M, and McCarthy J. 2018. Advances in the systemic
326 treatment of triple-negative breast cancer. Curr Oncol 25:S142-S150. 10.3747/co.25.3954
327 Liedtke C, Hess KR, Karn T, Rody A, Kiesel L, Hortobagyi GN, Pusztai L, and Gonzalez-
328 Angulo AM. 2013. The prognostic impact of age in patients with triple-negative breast
329 cancer. Breast Cancer Res Treat 138:591-599. 10.1007/s10549-013-2461-x
330 Liedtke C, Rody A, Gluz O, Baumann K, Beyer D, Kohls EB, Lausen K, Hanker L, Holtrich U,
331 Becker S, and Karn T. 2015. The prognostic impact of age in different molecular
332 subtypes of breast cancer. Breast Cancer Res Treat 152:667-673. 10.1007/s10549-015-
333 3491-3
334 Liu Y, Xin T, Huang DY, Shen WX, Li L, Lv YJ, Jin YH, Song XW, Teng C, and Jiang QY.
335 2014. Prognosis in very young women with triple-negative breast cancer: retrospective
336 study of 216 cases. Med Oncol 31:222. 10.1007/s12032-014-0222-2
337 Nixon AJ, Neuberg D, Hayes DF, Gelman R, Connolly JL, Schnitt S, Abner A, Recht A, Vicini
338 F, and Harris JR. 1994. Relationship of patient age to pathologic features of the tumor
339 and prognosis for patients with stage I or II breast cancer. J Clin Oncol 12:888-894.
340 10.1200/JCO.1994.12.5.888
341 Perou CM, and Borresen-Dale AL. 2011. Systems biology and genomics of breast cancer. Cold
342 Spring Harb Perspect Biol 3. 10.1101/cshperspect.a003293
343 Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT,
344 Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lonning PE,
345 Borresen-Dale AL, Brown PO, and Botstein D. 2000. Molecular portraits of human
346 breast tumours. Nature 406:747-752. 10.1038/35021093
347 Peshkin BN, Alabek ML, and Isaacs C. 2010. BRCA1/2 mutations and triple negative breast
348 cancers. Breast Dis 32:25-33. 10.3233/BD-2010-0306
349 Reis-Filho JS, and Pusztai L. 2011. Gene expression profiling in breast cancer: classification,
350 prognostication, and prediction. Lancet 378:1812-1823. 10.1016/S0140-6736(11)61539-0
351 Rosenberg SM, Ruddy KJ, Tamimi RM, Gelber S, Schapira L, Come S, Borges VF, Larsen B,
352 Garber JE, and Partridge AH. 2016. BRCA1 and BRCA2 Mutation Testing in Young
353 Women With Breast Cancer. JAMA Oncol 2:730-736. 10.1001/jamaoncol.2015.5941
Ruder AM, Lubin F, Wax Y, Geier A, Alfundary E, and Chetrit A. 1989. Estrogen and progesterone receptors in breast cancer patients. Epidemiologic characteristics and survival differences. Cancer 64:196-202.

Satagopan JM, Ben-Porat L, Berwick M, Robson M, Kutler D, and Auerbach AD. 2004. A note on competing risks in survival data analysis. Br J Cancer 91:1229-1235. 10.1038/sj.bjc.6602102

Shoker BS, Jarvis C, Clarke RB, Anderson E, Hewlett J, Davies MP, Sibson DR, and Sloane JP. 1999. Estrogen receptor-positive proliferating cells in the normal and precancerous breast. Am J Pathol 155:1811-1815. 10.1016/S0002-9440(10)65498-3

Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Lonning PE, and Borresen-Dale AL. 2001. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A 98:10869-10874. 10.1073/pnas.191367098

Tarone RE, and Chu KC. 2002. The greater impact of menopause on ER- than ER+ breast cancer incidence: a possible explanation (United States). Cancer Causes Control 13:7-14.

Tran B, and Bedard PL. 2011. Luminal-B breast cancer and novel therapeutic targets. Breast Cancer Res 13:221. 10.1186/bcr2904

Yasui Y, and Potter JD. 1999. The shape of age-incidence curves of female breast cancer by hormone-receptor status. Cancer Causes Control 10:431-437.
Figure 1

Univariate analysis based on the competing risk regression model.

The association between age and breast cancer in all cohort (A), HR+ group (D), and molecular subtypes (C-F)
1 = BCSM
2 = Other causes of death

A. All cohort

- 20-39y 1
- 40-49y 1
- 50-79y 1
- 20-39y 2
- 40-49y 2
- 50-79y 2

Death rate of 1: 6.09%  
P1 < 0.001

Death rate of 2: 4.70%  
P2 < 0.001

B. HR+ only

- 20-39y 1
- 40-49y 1
- 50-79y 1
- 20-39y 2
- 40-49y 2
- 50-79y 2

Death rate of 1: 4.54%  
P1 < 0.001

Death rate of 2: 4.59%  
P2 < 0.001

C. HR+/HER2-

- 20-39y 1
- 40-49y 1
- 50-79y 1
- 20-39y 2
- 40-49y 2
- 50-79y 2

Death rate of 1: 4.41%  
P1 < 0.001

Death rate of 2: 4.67%  
P2 < 0.001

D. HR+/HER2+

- 20-39y 1
- 40-49y 1
- 50-79y 1
- 20-39y 2
- 40-49y 2
- 50-79y 2

Death rate of 1: 5.48%  
P1 > 0.05

Death rate of 2: 3.99%  
P2 < 0.001

E. HER2

- 20-39y 1
- 40-49y 1
- 50-79y 1
- 20-39y 2
- 40-49y 2
- 50-79y 2

Death rate of 1: 9.84%  
P1 > 0.05

Death rate of 2: 5.16%  
P2 < 0.001

F. TNBC

- 20-39y 1
- 40-49y 1
- 50-79y 1
- 20-39y 2
- 40-49y 2
- 50-79y 2

Death rate of 1: 15.42%  
P1 < 0.001

Death rate of 2: 5.35%  
P2 < 0.001
Table 1 (on next page)

The characteristic of each involved variables

**Abbreviations:** HR, hazard ratio; 95%CI, 95% confidence index; HR+, hormone receptor positive; HER2, human epidermal growth factor receptor 2; TNBC, triple negative breast cancer.
Table 1. The characteristic of each involved variables

| Characteristics                          | <40y          | 40y-49y       | >50y         | p value |
|-----------------------------------------|---------------|---------------|--------------|---------|
| Breast cancer subtype                   |               |               |              |         |
| HR+/HER2-                               | 1023 (54.50%) | 4506 (69.48%) | 19351 (75.57%) | < 0.001 |
| HR-/HER2+                               | 135 (7.19%)   | 285 (4.39%)   | 1054 (4.12%)  |         |
| HR+/HER2+                               | 341 (18.17%)  | 830 (12.80%)  | 2313 (9.03%)  |         |
| HR-/HER2-                               | 378 (20.14%)  | 864 (13.32%)  | 2888 (11.28%) |         |
| Race                                    |               |               |              | < 0.001 |
| White                                   | 1347 (71.76%) | 4990 (76.95%) | 21060 (82.25%) |         |
| Black                                   | 284 (15.13%)  | 769 (11.86%)  | 2416 (9.44%)  |         |
| American Indian/Alaska Native           | 13 (0.69%)    | 44 (0.68%)    | 113 (0.44%)   |         |
| Asian or Pacific Islander               | 233 (12.41%)  | 682 (10.52%)  | 2017 (7.88%)  |         |
| Laterality                              |               |               |              | 0.02    |
| Right - origin of primary               | 935 (49.81%)  | 3274 (50.49%) | 12465 (48.68%) |         |
| Left - origin of primary                | 942 (50.19%)  | 3211 (49.51%) | 13141 (51.32%)|         |
| Location                                |               |               |              | < 0.001 |
| Nipple                                  | 4 (0.21%)     | 11 (0.17%)    | 89 (0.35%)    |         |
| Central portion                         | 80 (4.26%)    | 276 (4.26%)   | 1322 (5.16%)  |         |
| Upper-inner quadrant                   | 211 (11.24%)  | 809 (12.47%)  | 3308 (12.92%) |         |
| Lower-inner quadrant                   | 103 (5.49%)   | 357 (5.51%)   | 1571 (6.14%)  |         |
| Upper-outer quadrant                    | 653 (34.79%)  | 2325 (35.85%) | 8972 (35.04%) |         |
| Lower-outer quadrant                    | 170 (9.06%)   | 488 (7.53%)   | 1938 (7.57%)  |         |
| Axillary tail                           | 12 (0.64%)    | 36 (0.56%)    | 118 (0.46%)   |         |
| Overlapping lesion                      | 413 (22.00%)  | 1390 (21.43%) | 5588 (21.82%) |         |
| Breast, NOS                             | 231 (12.31%)  | 793 (12.23%)  | 2700 (10.54%) |         |
| Grade                                   |               |               |              | < 0.001 |
| Well differentiated; Grade I            | 153 (8.15%)   | 1182 (18.23%) | 6492 (25.35%) |         |
| Moderately differentiated; Grade II     | 641 (34.15%)  | 2734 (42.16%) | 11320 (44.21%)|         |
| Poorly differentiated; Grade III        | 1074 (57.22%) | 2540 (39.17%) | 7701 (30.07%) |         |
| Undifferentiated; anaplastic; Grade IV  | 9 (0.48%)     | 29 (0.45%)    | 93 (0.36%)    |         |
| Tumor size | <=1cm | <=2cm | <=3cm | <=4cm | <=5cm | >5cm |
|------------|-------|-------|-------|-------|-------|------|
|            |       |       |       |       |       |      |
| < 0.001    | 234 (12.47%) | 1336 (20.60%) | 7199 (28.11%) |       |       |      |
| <=1cm      | 596 (31.75%) | 2323 (35.82%) | 9719 (37.96%) |       |       |      |
| <=2cm      | 469 (24.99%) | 1483 (22.87%) | 4760 (18.59%) |       |       |      |
| <=3cm      | 267 (14.22%) | 611 (9.42%) | 1794 (7.01%) |       |       |      |
| <=4cm      | 115 (6.13%) | 281 (4.33%) | 875 (3.42%) |       |       |      |
| >5cm       | 196 (10.44%) | 451 (6.95%) | 1259 (4.92%) |       |       |      |

| Tumor stage | I | II | III | IV |
|-------------|---|----|-----|----|
| < 0.001     |   |    |     |    |
| I           | 572 (30.47%) | 2887 (44.52%) | 14381 (56.16%) |       |
| II          | 897 (47.79%) | 2572 (39.66%) | 8132 (31.76%) |       |
| III         | 366 (19.50%) | 928 (14.31%) | 2730 (10.66%) |       |
| IV          | 42 (2.24%) | 98 (1.51%) | 363 (1.42%) |       |

| Regional nodes positive | >=10 | 0 | 1-3 | 4-9 |
|-------------------------|------|---|-----|-----|
| < 0.001                 |      |  |     |     |
| >=10                    | 87 (4.64%) | 211 (3.25%) | 808 (3.16%) |       |
| 0                       | 1003 (53.44%) | 4029 (62.13%) | 18097 (70.67%) |       |
| 1-3                     | 616 (32.82%) | 1735 (26.75%) | 5196 (20.29%) |       |
| 4-9                     | 171 (9.11%) | 510 (7.86%) | 1505 (5.88%) |       |

| Marital status | Married | Unmarried |
|----------------|---------|-----------|
|                |         |           |
| < 0.001        |         |           |
| Married        | 1145 (61.00%) | 4364 (67.29%) | 15458 (60.37%) |       |
| Unmarried      | 732 (39.00%) | 2121 (32.71%) | 10148 (39.63%) |       |

| Radiotherapy | No | Yes |
|--------------|----|-----|
| < 0.001      |    |     |
| No           | 938 (49.97%) | 2949 (45.47%) | 10596 (41.38%) |       |
| Yes          | 939 (50.03%) | 3536 (54.53%) | 15010 (58.62%) |       |

| Chemotherapy | No | Yes |
|--------------|----|-----|
| < 0.001      |    |     |
| No           | 366 (19.50%) | 2389 (36.84%) | 15564 (60.78%) |       |
| Yes          | 1511 (80.50%) | 4096 (63.16%) | 10042 (39.22%) |       |

**Abbreviations:** HR, hazard ratio; 95%CI, 95 confidence index; HR+, hormone receptor positive; HER2, human epidermal growth factor receptor 2; TNBC, triple negative breast cancer.
Table 2 (on next page)

Multivariate Cox and SH analyses breast cancer subtypes

**Abbreviations:** HR, hazard ratio; 95%CI, 95% confidence index; HR+, hormone receptor positive; HER2, human epidermal growth factor receptor 2; TNBC, triple negative breast cancer; Significant results with p <0.05 were bolded.
**Table 2.** Multivariate Cox and SH analyses breast cancer subtypes

|                      | OS   | SH   |
|----------------------|------|------|
|                      | Age  | HR (95%CI) | HR (95%CI) |
| All cohort          |      |            |            |
| <40y                 | References | References |
| >=40y               | 1.12 (0.98-1.28) | 0.87 (0.75-1.02) |
| HR+ only            |      |            |            |
| <40y                 | References | References |
| >=40y               | 1.27 (1.05-1.52) | 0.95 (0.77-1.17) |
| HR+/HER2-           |      |            |            |
| <40y                 | References | References |
| >=40y               | 1.10 (0.91-1.35) | 0.86 (0.68-1.07) |
| HR+/HER2+           |      |            |            |
| <40y                 | References | References |
| >=40y               | **2.07 (1.28-3.35)** | 1.31 (0.78-2.20) |
| HER2                |      |            |            |
| <40y                 | References | References |
| >=40y               | 1.17 (0.73-1.88) | 0.86 (0.51-1.46) |
| TNBC                |      |            |            |
| <40y                 | References | References |
| >=40y               | 0.87 (0.70-1.08) | **0.71 (0.56-0.89)** |

**Abbreviations:** HR, hazard ratio; 95%CI, 95% confidence index; HR+, hormone receptor positive; HER2, human epidermal growth factor receptor 2; TNBC, triple negative breast cancer; Significant results with p <0.05 were bolded.
Table 3 (on next page)

Multivariate Cox and SH analyses of breast cancer subtypes

HR, hazard ratio; 95%CI, 95% confidence index; HR+, hormone receptor positive; HER2, human epidermal growth factor receptor 2; TNBC, triple negative breast cancer; Significant results with p < 0.05 were bolded. To be noted, all the variables were involved in the multivariate Cox analyses. While in the Competing Risks Regression, in case the iterative algorithm was not converged, the race, laterality, tumor stage, marital status, positive regional nodes, radiotherapy and chemotherapy were involved in SH model for HR+/HER2+ and TNBC groups, the tumor stage, radiotherapy and chemotherapy were involved in SH model for HER2 group, and all variables were involved in SH model for HR+/HER2- group.
Table 3. Multivariate Cox and SH analyses of breast cancer subtypes

| Group (number) | Treatment  | <40y | <40y |
|---------------|-----------|-----|-----|
|               |           | OS  | SH  |
| HR+/HER2-     | Radiotherapy | 0.60 (0.41-0.88) | 0.59 (0.35-0.99) |
| (1023)        | Chemotherapy | References | References |
| HR+/HER2+     | Radiotherapy | 1.35 (0.53-3.45) | 0.93 (0.29-3.01) |
| (341)         | Chemotherapy | References | References |
| HER2          | Radiotherapy | 0.02 (0.005-0.63) | 0.53 (0.10-2.89) |
| (135)         | Chemotherapy | References | References |
| TNBC          | Radiotherapy | 0.98 (0.51-1.88) | 0.55 (0.23-1.35) |
| (378)         | Chemotherapy | References | References |

HR, hazard ratio; 95%CI, 95% confidence index; HR+, hormone receptor positive; HER2, human epidermal growth factor receptor 2; TNBC, triple negative breast cancer; Significant results with p < 0.05 were bolded. To be noted, all the variables were involved in the multivariate Cox analyses. While in the Competing Risks Regression, in case the iterative algorithm was not converged, the race, laterality, tumor stage, marital status, positive regional nodes, radiotherapy and chemotherapy were involved in SH model for HR+/HER2+ and TNBC groups, the tumor stage, radiotherapy and chemotherapy were involved in SH model for HER2 group, and all variables were involved in SH model for HR+/HER2- group.