Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature: _____ _____

Qin Sun Date

Apocrine Carcinoma of the Breast: Treatment Patterns and Outcomes

By

Qin Sun

Master of Public Health

Biostatistics and Bioinformatics

_____

Yuan Liu

Thesis Advisor

_____

Cletus Arciero

Reader

Apocrine Carcinoma of the Breast: Outcomes and Treatment Approaches

By

Qin Sun

B.S.

Rutgers University – New Brunswick

2014

Thesis Committee Chair: Yuan Liu, Ph.D.

An abstract of

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Public Health

in Biostatistics and Bioinformatics

2016

Abstract

Apocrine Carcinoma of the Breast: Outcomes and Treatment Approaches

By Qin Sun

Introduction: Invasive breast cancer is the most common malignancy affecting women worldwide and also an important cause of cancer death. Apocrine carcinoma is a rare subtype of invasive ductal carcinoma of the breast, accounting for 0.3–1% of all breast cancers. This type of breast cancer is commonly ER (estrogen receptor), PR (progesterone receptor) and HER2 (receptor type 2) negative (triple negative), while expressing AR (androgen receptor). Our study sought to explore the characteristics (e.g. receptor status, treatment patterns, demographics and clinical characteristics) of this rare subset of breast cancer and compare outcomes with that of the more common triple negative invasive ductal carcinoma (TNBC).

Methods: The National Cancer Database (NCDB) Participant User File (PUF) for breast cancer from 2004- 2013 was queried, with two cohorts - apocrine carcinoma versus triple negative invasive ductal carcinoma identified. Descriptive and univariate statistics were calculated for all variables. A multivariable Cox proportional hazard model was used to compare overall survival (OS) between the two cohorts. Univariate and multivariate stratified analyses were also performed to investigate possible interactions between covariates and cohorts. Propensity score matching was performed to further reduce selection bias.

Results: There were 2,807,541 patients in the NCDB Breast PUF 2012 database, 38,514 patients (2537 (6.6%) apocrine carcinoma and 35,977 (93.4%) TNBC) met study criteria. 72.1% of the total study population was at least 50 years old. In terms of race, 73.4% were white and 21.9% were black. In multivariable analysis, patients with TNBC were at a higher risk of death compared to apocrine carcinoma (HR=1.40, $p < 0.001$) and a HR=1.34, ($p < 0.001$) based on propensity score matching approach. Stratified analyses showed a better outcome in apocrine carcinoma cohort for younger group (<50), the receipt of chemotherapy, LN positive, LVI present and HER2 when compared with the TNBC cohort.

Conclusion: This retrospective analysis of a national cancer database suggests that apocrine carcinoma of the breast has a better outcome, in terms of survival, than TNBC.

Apocrine Carcinoma of the Breast: Outcomes and Treatment Approaches

By

Qin Sun

B.S.

Rutgers University – New Brunswick

2014

Thesis Committee Chair: Yuan Liu, Ph.D.

An abstract of

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Public Health

in Biostatistics and Bioinformatics

2016

Acknowledgements

I am extremely grateful for my amazing thesis advisor Yuan Liu, whose encouragement, guidance, and support made this thesis possible.

I am also thankful for Cletus Arciero taking time to read my thesis.

Finally, I would like to thank my parents, Yale Sun and Yun Qian, for their love and support throughout my education.

Contents

List of Tables

List of Figures

1. Introduction

Invasive breast cancer is the most common malignancy affecting women worldwide and also an important cause of cancer death[8]. Most tumors are derived from mammary ductal epithelium, and the most common epithelial type (up to 50-80%) is invasive ductal carcinoma. The second most common epithelial type (up to 5-15%) is invasive lobular carcinoma. Apocrine carcinoma is a rare subtype of invasive ductal carcinoma of the breast, accounting for as little as 0.3–1% of all breast cancers and showing distinct morphologic, immunohistochemical and molecular genetic features[6,11,16]. The demographics of apocrine carcinoma indicate a prevalence of the disease in older women, with the average age of 65[15]. Apocrine carcinoma is graded based on ICD-O-3 grading system, with the majority of apocrine carcinoma graded as grade II and III[7]. Mitotic activity in apocrine carcinoma is usually moderate to high, particular in triple negative[5].

In the recent literature, apocrine carcinoma has been immunohistochemically categorized as estrogen receptor (ER) negative, progesterone receptor (PR) negative and androgen receptor (AR) positive tumors, and can be recognized by the combination of immunohistochemical and morphologic characteristic[18]. Overexpression of human epidermal growth factor receptor type 2 (HER2) is also frequently seen in apocrine carcinoma, although studies on HER2 status of apocrine-eccrine carcinomas are few in number. [18]. Apocrine tumors have been shown to be HER2-positive about 50% of the time, while HER2-negative apocrine carcinoma can be phenotyped as triple-negative breast cancer (TNBC). Of note, according to the studies of U.S and British women, triple negative/basal-like tumors appear to less common among white women compared to black women[15].

There are two main therapeutic approaches for the treatment of breast cancer, including apocrine carcinoma. Local therapy, including surgery and often radiation, is aimed at preventing

the local recurrence of the cancer. Systemic therapy, which can include hormone therapy and chemotherapy, is aimed at the prevention and treatment of both locoregional and distant disease recurrence[17]. Often, therapy is guided based upon the features of tumor, such as tumor grade, hormone receptor status, and perceived aggressiveness.

Our aim was to explore the clinical characteristics, receptor status, treatment approaches, and long term survival for apocrine carcinoma patients in direct comparison to patients with triple negative invasive ductal carcinoma. Our analysis focused on outcomes as measured by overall survival (OS), which defined as time to death or last follow up from date of diagnosis based on NCDB data dictionary. Demographic information and clinical data was also examined by performing stratified analysis, to identify whether various covariates influence the OS of breast cancer patients. Specifically our analysis sought to identify how the receptor status (ER, PR and HER2) may play a role in OS in apocrine carcinoma of the breast compared to triple negative invasive ductal carcinoma.

2. Methodology

2.1 Study Design and Patients

This is an observational, cohort study based on the data from NCDB Breast PUF 2012, a nationwide oncology outcomes database that captures data on approximately 70% of all new invasive cancer diagnoses in the United States each year[2]. The database contains 2,807,541 patients treated for an invasive breast tumor from 1998 to 2012. Triple negative invasive ductal carcinoma group was selected based upon two criteria: the eligible histology (8500) and the combination of ER (-), PR (-) and HER2 (-) negative. The primary outcome, which was defined as overall survival (OS), is the time from date of diagnosis to death or the last follow up. The reason that we limited the year of diagnosis to 2003 and later is that an important confounder

variable - Charlson Comorbidity Score is only available starting from 2003 in the NCDB PUF. Patients were further excluded if their tumor had in-situ behavior, male patients, palliative care and missing value of outcome (OS), resulting in a dataset of 38,514 primary tumors. This group is divided into two study cohorts: 2,537 with apocrine carcinoma breast cancer and 35,977 with triple negative invasive ductal carcinoma (Figure 1).

2.2 Variables and Measurement

Study variables were defined by the PUF dictionary of the NCDB[13]. The following baseline data items of the NCDB included: (1) patient's socioeconomic characteristics (age at diagnose, race/ethnicity, facility type, facility location, primary payer, urban/rural, median income, educational status); (2) disease characteristics (Charlson-Deyo score/CDCC, tumor grade, tumor size, regional lymph nodes positive/examined, sites of distant metastasis, Lymph Vascular Invasion, ER, PR, HER2, triple negative); (3) treatment characteristics (surgery of primary site, chemotherapy, hormone, radiation).

For patients' socioeconomic characteristics: Age at diagnosis was arbitrarily divided into two groups (<50 vs. >=50); race was divided into White and Black for comparison; facility type was divided into three groups (community cancer program/other vs. comprehensive community cancer program vs. academic/research program); facility location was divided into four parts (Northeast vs. South vs. Midwest vs. West); primary payer was divided into four groups (private vs. Medicaid vs. Medicare/other government vs. not insured/unknown); urban/rural status was divided into three groups (metro vs. urban vs. rural); median income quartiles in 2000 which means median household income for each patient's area of residence, were categorized as quartiles based on equally proportioned income ranges among all US zip codes (less than $30,000, between $30,000 to $35,999, between $36,000 to $45,999, more than $46,000);

educational status which recorded as percent of no high school degree quartiles in 2000 are less than 14%, between 14% to 19.9%, between 20% to 28.9%, more than 29%.

For patients' diseases characteristics: Charlson-Deyo score or comorbid conditions was analyzed as a binary variable (0 vs. 1+); tumor grade was assessed according to ICD-O-3 grading system and divided into four groups (well differentiated vs. moderately differentiated vs. poorly differentiated/undifferentiated vs. cell type not determined); tumor size quartile were divided into 0-2 cm, >2-5 cm, >5 cm; expression of ER, PR and HER2 were demonstrated by site specific factors from the Collaborative Stage Data Collection System using SSF1 for ER, SSF2 for PR and SSF15 for HER2 and divided into two groups (negative vs. positive); sites of distant metastasis focused on bone, brain, liver and lung involvement; lymph vascular invasion (LVI) was coded as present vs. not present; lymph node metastases (LN) was divided into two groups (negative vs. positive).

For patients' treatment characteristics: surgery of primary site, chemotherapy, hormone therapy and radiation therapy were all divided into two groups (Yes vs. No) for comparison.

2.3 Statistical Analyses

Descriptive statistics for each variable were reported. The univariate association between each covariate and study cohorts (apocrine carcinoma and triple negative invasive ductal carcinoma (TNBC)) were assessed using the chi-square test for categorical covariates, and ANOVA for continuous covariates. Univariate associations between each covariate including study cohorts and study outcome were then calculated for all socio-demographic, disease and treatment variables of interest, using Cox proportional hazards model and long-rank tests. Kaplan-Meier (KM) plots were produced to directly compare the survival curves by subgroups along with log-rank p-value and 5-year survival rate. The endpoint of interest was overall survival, which is the

time in months until death or last follow up from the date of diagnosis. The variable Histology represents the patient's risk category (apocrine carcinoma vs. TNBC) and the variable "60 Mo Survival" represents the disease-free survival time. The variable Censored is the censoring indicator: a status of 1 indicates alive, and a status of 0 indicates dead. The variable Event is OS.

A multivariable Cox proportional hazard model was fitted by a backward variable selection method applying an alpha =0.20 removal criteria, two variables, which were ER and PR, were removed from the model. Stratified Cox proportional hazards models were then conducted to generated stratum-specific treatment hazard ratios for age distribution, ethnicity, receptor status (ER, PR, and HER2) and receipt of chemo, hormone and radiation therapy. Long-rank tests were used to determine whether there were significant differences in the association between apocrine carcinoma and overall survival among each level of a categorical variable.

Since the data in this study are observational, there may be inherent differences in the apocrine carcinoma and TNBC groups. Therefore, propensity score matching was performed to create a sample of patients with balanced confounding variables. Propensity score technique, introduced by Rosenbaum and Rubin[14], is a popular method that compliments multivariable modeling as it can visually create homogenous study cohorts in terms of baseline characteristics and allow a straightforward estimation of treatment versus control effects that reflects adjustment for differences in all observed background characteristics, and hence reduce the estimation and treatment selection bias. A propensity score is defined as the probability of treatment assignment conditional on observed baseline characteristics: $e_i = \Pr(Z_i = 1 \mid X_i)$[1]. Propensity score method (PS) was conducted and the same variables in multivariable analysis were used in PS calculation. Four basic steps are involved as below: 1) estimate propensity score, 2) match or weight patients by their propensity scores, 3) check balance of baseline characteristics after PS technique and

treat <0.1 as negligible imbalance, 4) estimate the treatment effect in the matched or weighted sample by conducting Cox model with a robust variance estimator for OS.

Statistical analysis was conducted using SAS Version 9.4, and SAS macros or software developed at the Biostatistics and Bioinformatics at Winship Cancer Institute. Hazard ratios (HR) are presented with 95% confidence intervals. Hypothesis tests were two-tailed and the significant level was set at 0.05.

3. Results

3.1. Descriptive Statistics

The descriptive analysis for all variables of interest including the receptor status (ER, PR and HER2) and treatment approaches (chemo, hormone and radiation therapy) is shown in Table 1. The final analytic cohort consisted of 38,514 patients after applying the inclusion and exclusion criteria. Women with rare apocrine carcinoma breast cancer comprised only 6.6% of the total study population from 2003 to 2012. The average age at diagnose was 58.3 and 72.1% of the total study population was at least 50-year-old. 73.4% were white and 21.9% were black. Most patients lived in a metropolitan region (85%). Almost half or 41.3% of study population had more than $46,000 median income per year. For the disease characteristics, most patients (76.2%) had poorly differentiated/undifferentiated tumor grade. 84% of patients had 0 score of Charlson-Deyo Score, which means most patients indicated "no comorbid conditions recorded". 59.4% of patients were LN negative compared to 30.5% LN positive. Almost all or 97.5% of patients demonstrated no bone/brain/liver/lung metastases. Most patients were ER- (96.2%), PR- (96.7%) and HER2- (89.4%), based largely on the inclusion criteria. For the treatment characteristics, most patients received primary site surgery (77.3%), chemotherapy (74.5%) and radiation

therapy (55%), but only a few of them had hormone therapy (4.5%) due to the lack of ER or PR positivity.

3.2. Univariate Association with Study Cohorts

The univariate association between each covariate and study cohorts (apocrine carcinoma vs. TNBC) is shown in Table 2. Apocrine carcinoma occurred more often in older patients (>=50 years; p <0.01), which accounted for 84.98% vs. 71.2% in TNBC patients. 80.8% of all apocrine carcinoma cases were white women and 72.91% of all TNBC were whited women (p <0.001). For the disease characteristics, 85.9% of apocrine carcinoma patients and 83.9% TNBC patients were Charlson-Deyo Score or CDCC = 0 (p=0.007). Among the three categories of breast cancer staging with 0-2cm (stage I), >2-5cm (stage II) and >5cm (stage III), 33.94% apocrine carcinoma patients in stage I, 43.16% in stage II and 22.90% in stage III but there was no distinct different among TNBC patients based on their tumor sizes. For treatment receipt, patients received primary site surgery (80.88% in apocrine carcinoma vs. 77% in TNBC; p < 0.001), chemotherapy (52.58% in apocrine carcinoma vs. 76% in TNBC, p<0.001) and radiation therapy (51.28% in apocrine carcinoma vs. 55.3% in TNBC, p < 0.001).

3.3. Univariate/Multivariate Association with Overall Survival

The univariate (UVA) and multivariable survival analysis (MVA) with OS for all variables is shown in Table 3. In UVA, apocrine carcinoma of the breast (N=2537) had a better outcome than TNBC (N=35977) (HR_{TNBC}=1.46, p < 0.001), which fulfilled our expectation above. A better outcome (OS) is associated with younger group (HR=0.77 [0.72-0.81], p < 0.001), white ethnicity (HR_{white}=1.00 vs. HR_{black}=1.35, p < 0.001), private insurance (HR=0.52, p-value < 0.001), median income more than $46,000 ($HR_{\$46,000+}$=1 vs. $HR_{\$36,000-\$45,999}$=1.31 vs. $HR_{30,000-\$35,999}$=1.36 vs. HR<$30,000=1.61, p < 0.001), 0 score of Charlson-Deyo Score (HR=0.57, p <

0.001), lymph node negative disease (HR=0.24, p < 0.001), lack of metastasis (bone/brain/liver/lung) (HR=0.34, p < 0.001) and a lack of lymph vascular invasion (HR=0.67, p < 0.001). There was a linear relationship between OS and covariates (tumor grade, tumor size and LN examined). As each level of these covariates increased, the OS in breast cancer patients with apocrine carcinoma decreased (p <0.001).

In order to determine whether age distribution, ethnicity, LN, LVI, receptor status and/or treatment approaches influences OS, a multivariate Cox proportional hazards model was performed in MVA by controlling all possible confounding effects. Two variables were removed during the backwards elimination process: ER and PR. After controlling covariates including socio-demographic, disease characteristics and treatment approaches, patients with TNBC had a higher overall mortality rate compared to apocrine carcinoma patients (HR = 1.40 [1.22-1.60], p<.001).

3.4. Univariate Stratified Analyses (Kaplan Meier Curve)

In order to identify OS in breast cancer patients with apocrine carcinoma stratified by clinical characteristics, receptor status and treatment approach in a large population, survival analysis or Kaplan Meier analysis and long-rank test were performed to compare survival curves between apocrine carcinoma breast cancer group and TNBC group. The endpoint of interest was the 5-year survival time, which was the time in months until death or last follow up from the date of diagnosis. The 5-year observed survival rate for the entire group was 78.2% with apocrine carcinoma and 69.5% with TNBC (Figure 2). Patients with apocrine carcinoma had a much higher 5-year observed survival rate than patients with TNBC in LN positive patients (73.3% vs. 57.1%) and younger age group at diagnose (88.4% vs. 72.7%) (Figure 3). Patients with apocrine carcinoma had a slightly higher 5-year observed survival rate for both white and black women

(78.1% vs. 70.4% of White; 72.3% vs. 64.9% of Black; Log-rank p <0.0001), for elder age groups at diagnose (76.3% vs.68.1% of >=50-year-old; Log-rank p <0.0001), patient who received radiation (85.1% vs. 74.3%) and for patients who received chemotherapy (82.7% vs. 71.8%) and not received chemotherapy (71.7% vs.61.4%) (Figure 3). There was no distinct difference in the 5-year survival rates between two study cohorts stratified by ER (+/-), PR (+/-) and LN negative (Figure 3). The 5-year observed survival time of TNBC patients was longer than apocrine carcinoma stratified by HER2 (-) (69.3% vs.60.1%), but much shorter than the group of apocrine carcinoma stratified by HER2+ (88.9% vs. 64.5%) although this might result from small group sizes (Figure 3). 3-year observed survival rate was used to compare LVI stratum between study cohorts because sample size of apocrine carcinoma is too small, and indicated a greater rate of survival with apocrine carcinoma patients stratified by LVI present (92.6% vs. 78.1%; Log-rank p <0.0047).

3.5. Multivariate Stratified Analyses

Results from stratified Cox proportional hazards model are presented in Table 4. After making interaction with clinical characteristics, receptor statues and treatment approaches, variables that indicated significant differences in association apocrine carcinoma and OS within their categories were age distribution (p =0.018) and receipt of chemotherapy (p =0.037). Patients with apocrine carcinoma had lower odds of death than TNBC in both age groups, especially in younger group which categorized as under 50 years old ($HR_{<50}$=0.50 [0.35, 0.71] vs. $HR_{>=50}$=0.77 [0.67, 0.88] , p <.001). Similarly, Patients within apocrine carcinoma cohort had lower odds of death than TNBC in both chemotherapy groups, especially lower in patients who accepted chemotherapy in the past (HR_{yes}=0.66 [0.55, 0.79], p-value<.001 vs. HR_{No}=0.81 [0.69, 0.95], p =0.010). Although there was no statistical significance between levels of other stratified

covariates, patients with apocrine carcinoma had significantly lower hazard in ER positive group ($HR_{positive}$=0.73 [0.51, 1.05], p-value=0.086 vs. $HR_{negative}$=0.76 [0.65, 0.88], p <.001), PR positive group ($HR_{positive}$=0.68 [0.46, 0.98], p =0.041 vs. $HR_{negative}$=0.76 [0.66, 0.88], p <.001], Black patients (HR_{black}=0.64 [0.49, 0.83], p <0.001 vs. HR_{white}=0.75 [0.64, 0.87], p <.001).

3.6. Propensity Score Matching

The distribution of the logit of the propensity scores calculated in the apocrine carcinoma and TNBC cohorts are shown in Figure 4. The two curves were not overlapping well indicating baseline heterogeneity among covariates. Matching yielded two cohorts (apocrine carcinoma and TNBC) of 1192 patients (Table 5). Compared with Table 2, patients in the matched cohorts had similar distributions of socio-demographic and disease characteristics, receptor status and treatment approaches.

3.7. Overall Survival in Matched Sample

Based on the K-M survival plot of matched sample, the 5-year-survival rate for the apocrine carcinoma patients was higher (76.3% (CI 73.2%, 79.1%)) then TNBC patients (69.9% (CI 65.5%, 73.8%)). These differences in OS were statistical significant, which indicating patients in the TNBC cohort were associated with a significantly shorter survival rate (HR=1.34 [1.12-1.61]; p-value=001) than the apocrine carcinoma cohort (Fig5).

4. Discussion

Our analysis sought to compare the outcomes (OS) between breast cancer patients with apocrine carcinoma and those with triple negative invasive ductal carcinoma, as well as to identify OS in apocrine carcinoma cohort stratified by socio- demographic characteristics, receptor status and treatment approaches in a large population based observational analysis using NCDB database.

Some studies indicated a higher rate of apocrine carcinoma (and AR expression) in the elderly[18] and the statistical prevalence of apocrine carcinoma is higher in older women with an average of patient age of 65. Our descriptive analysis showed that mean age at diagnose was 58.3 for all patients. Our univariate association analysis indicated apocrine carcinoma occurs more often in older patients (>=50 years; p-value <0.01) which accounted for 84.98%. In addition, studies showed that triple negative tumors were more likely to occur among black women, but analysis indicated white women accounted for a higher rate of both apocrine carcinoma and TNBC, maybe due to a large proportion of white patients in total cases (73.4%).

Based on the univariate (UVA) and multivariable survival analysis (MVA) with OS for all variables, apocrine carcinoma of the breast was associated with significantly better survival compared to TNBC after controlling for socio-demographic, receptor status and treatment approaches, which fulfilled our study objective and expectation. Moreover, patients with apocrine carcinoma of the breast have a higher survival rate than TNBC in the younger group, white ethnicity lymph node negative and no present of lymph vascular invasion. OS was also lower in apocrine carcinoma patients with receptor status of ER+, PR+ and HER2+ and receipt of chemotherapy, hormone and radiation therapy.

The stratified analyses performed in our study suggested that apocrine carcinoma within certain groups might have better OS. Patients with apocrine carcinoma had a higher 5-year observed survival rate than TNBC based on the results of Kaplan Meier analysis and long-rank test. According to multivariate stratified analysis, although both age distributions showed significantly better OS with apocrine carcinoma compared to TNBC, group under 50-year-old showed the lowest hazard ratio. Similarly, although some studies indicated a poor response to chemotherapy in apocrine carcinoma cohort[18], apocrine carcinoma patients who did receive

chemotherapy showed a lower hazard ratio compared to those who did not receive chemotherapy. These results suggest that apocrine carcinoma patients with ER+, PR+, HER2+ and black race had lower OS.

Propensity score matching resulted in two cohorts of 1114 patients which showed overlapping curves. The results of KM curve and Cox proportional hazards modeling of matched sample were the same as that of the multivariate Cox proportional hazards model controlling for baseline covariates in that TNBC had a higher death rate than apocrine carcinoma.

Our study has several strengths. To our knowledge, this is the largest retrospective cohort study used to access the comparison of overall survival between patients with apocrine carcinoma and TNBC. The use of the NCDB provided us with larger sample sizes than would be possible within clinical trial data or single institutional databases alone. But our study had limitations, due to the use of NCDB database and its methodology. Limitations for retrospective study include selection bias and unobserved confounding effect; and no disease specific survival data is available because of dataset limitations. The event will not occur in all individuals at the endpoint of study, and for these drop off patients, their full survival times are unknown. Propensity score analyses maybe useful to balance observed baseline covariate between study cohorts, but they do nothing to balance unmeasured characteristics and confounders[19]. It is possible that our results are confounded by the data we do not have, such as the larger missing data of HER2.

Reference

1. Austin, P. C. (2011a). An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behavioral Research.* 46: 399-424. doi:10.1080/00273171.2011.568786

2. Bilimoria, K. Y., Stewart, A. K., Winchester, D. P., & Ko, C. Y. (2008). The National Cancer Data Base: A Powerful Initiative to Improve Cancer Care in the United States. *Annals of Surgical Oncology*, *15*(3), 683–690. doi:10.1245/s10434-007-9747-3

3. Carey LA, Cheang MCU, Perou CM. Chapter 29: Genomics, Prognosis, and Therapeutic Interventions, in Harris JR, Lippman ME, Morrow M, Osborne CK. Diseases of the Breast, 5th edition, Lippincott Williams & Wilkins, 2014. 8

4. Clark TG, Bradburn MJ, Love SB, Altman DG. Survival analysis. Part I: basic concepts and first analyses.Br J Cancer. 2003;89:232–238.

5. Choi J., Jung W.H. and Koo J. (2012) Clinicopathologic features of molecular subtypes of triple negative breast cancer based on immunocytochemical markers. Histol.Histopathol. 27,1481-1493.

6. Dreyer G, Vandorpe T, Smeets A, Forceville K, Brouwers B, Neven P, Janssens H, Deraedt K, Moerman P, Van Calster B, Christiaens MR, Paridaens R, Wildiers H. Triple negative breast cancer: clinical characteristics in the different histological subtypes. Breast. 2013 Oct;22(5):761-6.

7. Eudebi V., Millis R., Cattani M., Bussoati G. and Azzopari J. (1986). Apocrine carcinoma of the breast. A morphologic and immunocytochemical study. Am. J. Pathol. 123,532-541

8. Gerhard R, Costa JL, Schmitt F. Benign and malignant apocrine lesions of the breast. Expert Rev Anticancer Ther. 2012;12:215–21. [PubMed]

9. Global cancer fact & figures 2007 – American Cancer Society

10. Kohler BA, Sherman RL, Howlader N, et al. Annual report to the nation on the status of cancer, 1975-2011, featuring incidence of breast cancer subtypes by race/ethnicity, poverty, and state. J Natl Cancer Inst. 107(6), 2015.

11. Li CI, Uribe DJ, Daling JR. Clinical Characteristics of different histology type of breast cancer. Br J Cancer 2005; 93(9): 1046-52

12. Lund MJ, Butler EN, Hair BY, et al. Age/race differences in HER2 testing and in incidence rates for breast cancer triple subtypes: a population-based study and first report. Cancer. 116(11):2549-59, 2010.

13. National Cancer Data Base. http://www.facs.org/cancer/ncdb/index.html. Retrieved 13 January 2016.

14. Rosenbaum P, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrka. 1983;70:41-55

15. Steven Halls (2016). Moose &Doc Breast Cancer. Accessed at http://breast-cancer.ca/apocr-canc/

16. Tavassoli FA, Devilee P. World health organization classification of tumors, tumors of the breast and femalegenital organs. 2nd ed. Lyon France: IARC press; 2003.

17. Tsutsumi Y. Apocrine carcinoma as triple-negative breast cancer: novel definition of apocrine-type carcinoma as estrogen/progesterone receptor-negative and androgen receptor-positive invasive ductal carcinoma. Jpn J Clin Oncol. 2012 May;42(5):375-86.

18. Vranic S, Schmitt F, Sapino A, Costa JL, Reddy S, Castro M, Gatalica Z. Apocrine carcinoma of the breast: a comprehensive review. Histol Histopathol. 2013 Nov;28(11):1393-409. E

19. Winkelmayer, W., & Kurth, T. (2004). Propensity Score: help or hype? Nephrol Dial Transplant, 19 (7), 1671-1673.

20. Yamazaki M, Nagata Y, Monji S, Shigematsu Y, Baba T, Shimokawa H, Uramoto H, Yamada S, Hanagiri T, Tanaka F. Apocrine carcinoma of the breast. J UOEH. 2011 Dec 1;33(4):293-301.

Appendix

A. Tables

Table 1. Descriptive Statistics for All Variables

Variable	Level	N (%) = 38514
Histology	Triple Negative Invasive Ductal Carcinoma	35977 (93.4)
	Apocrine Carcinoma	2537 (6.6)
Age Group	<50	10743 (27.9)
	>=50	27771 (72.1)
Facility Type	Community Cancer Program/Other	4491 (11.7)
	Comprehensive Community Cancer Program	21398 (55.6)
	Academic/Research Program	12625 (32.8)
Facility Location	Northeast	7747 (20.1)
	South	15597 (40.5)
	Midwest	9316 (24.2)
	West	5854 (15.2)
Race	White	28281 (73.4)
	Black	8448 (21.9)
	Others/Unknown	1785 (4.6)
Primary Payer	Not Insured/Unknown	1798 (4.7)
	Private	20684 (53.7)
	Medicaid	3642 (9.5)
	Medicare/Other Government	12390 (32.2)
Year of Diagnosis	2003	255 (0.7)
	2004	381 (1.0)
	2005	451 (1.2)
	2006	551 (1.4)
	2007	622 (1.6)
	2008	802 (2.1)
	2009	2395 (6.2)
	2010	15911 (41.3)
	2011	17146 (44.5)

Variable	Level	N (%) = 38514
Tumor Grade	Well Differentiated	778 (2.0)
	Moderately Differentiated	6845 (17.8)
	Poorly Differentiated/Undifferentiated	29331 (76.2)
	Cell Type Not Determined	1560 (4.1)
Urban/Rural 2003	Metro	31731 (85.0)
	Urban	4959 (13.3)
	Rural	623 (1.7)
	Missing	1201
Median Income Quartiles 2000	Not Available	1467
	< $30,000	5407 (14.6)
	$30,000 - $35,999	6323 (17.1)
	$36,000 - $45,999	10014 (27.0)
	$46,000 +	15303 (41.3)
Percent No High School Degree Quartiles 2000	Not Available	1470
	>=29%	6888 (18.6)
	20-28.9%	8529 (23.0)
	14-19.9%	8267 (22.3)
	< 14%	13360 (36.1)
Great Circle Distance (quartile)	>=0, <=5	9522 (24.7)
	>5, <=9	9522 (24.7)
	>9, <=19	9518 (24.7)
	>19, <=3868	9463 (24.6)
	Unknown	489 (1.3)
Charlson-Deyo Score	0	32348 (84.0)
	1+	6166 (16.0)
Tumor Size (quartile)	>=0, <=2	9974 (25.9)
	>2, <=3	9191 (23.9)
	>3, <=4	9087 (23.6)
	>4, <=99	9315 (24.2)
	Unknown	947 (2.5)
LN Examined (quartile)	>=1, <=2	10877 (28.2)
	>2, <=4	6862 (17.8)
	>4, <=11	7788 (20.2)
	>11, <=76	8146 (21.2)
	Unknown	4841 (12.6)

Variable	Level	N (%) = 38514
Lymph Node Postive	Negative	22875 (59.4)
	Positive	11759 (30.5)
	Unknown	3880 (10.1)
Metastatic Bone Involvement	None	32099 (97.1)
	Yes	651 (2.0)
	Unknown	307 (0.9)
	Missing	5457
Metatstatic Brain Involvement	None	32582 (98.6)
	Yes	166 (0.5)
	Unknown	309 (0.9)
	Missing	5457
Metastatic Liver Involvement	None	32284 (97.7)
	Yes	473 (1.4)
	Unknown	300 (0.9)
	Missing	5457
Metastatic Lung Involvement	None	32094 (97.1)
	Yes	657 (2.0)
	Unknown	306 (0.9)
	Missing	5457
Lymph Vascular Invasion	Not present	19672 (59.5)
	Prsent	6573 (19.9)
	Unknown	6812 (20.6)
	Missing	5457
Estrogen Receptor (ER) Assay	Negative	37069 (96.2)
	Positive	1020 (2.6)
	Unknown	425 (1.1)
Progesterone Receptor (PR) Assay	Negative	37236 (96.7)
	Positive	843 (2.2)
	Unknown	435 (1.1)
HER2: Summary Result of Testing	Negative	34446 (89.4)
	Positive	344 (0.9)
	Unknown	3724 (9.7)
Triple Negative	No	523 (1.4)
	Yes	36276 (94.2)
	Unknown	1715 (4.5)

Variable	Level	N (%) = 38514
Surgery of Primary Site	No	8724 (22.7)
	Yes	29764 (77.3)
	Unknown	26 (0.1)
Chemotherapy	No	9053 (23.5)
	Yes	28694 (74.5)
	Unknown	767 (2.0)
Hormone Therapy	No	36026 (93.5)
	Yes	1727 (4.5)
	Unknown	761 (2.0)
Radiation Therapy	No	16845 (43.7)
	Yes	21188 (55.0)
	Unknown	481 (1.2)
Age at Diagnosis	Mean	58.30
	Median	58.00
	Minimum	20.00
	Maximum	90.00
	Std Dev	13.74
	Missing	0.00
Great Circle Distance	Mean	21.49
	Median	8.60
	Minimum	0.00
	Maximum	3867.80
	Std Dev	82.85
	Missing	489.00
Tumor Size (cm)	Mean	2.70
	Median	2.10
	Minimum	0.00
	Maximum	98.80
	Std Dev	2.73
	Missing	947.00
Regional Lymph Nodes Examined	Mean	7.39
	Median	4.00
	Minimum	1.00
	Maximum	76.00
	Std Dev	7.38
	Missing	4841.00

Variable	Level	N (%) = 38514
Regional Lymph Nodes Positive	Mean	1.31
	Median	0.00
	Minimum	0.00
	Maximum	90.00
	Std Dev	3.48
	Missing	4551.00

Table 2. Univariate Association with Histology

			Histology		
Covariate	**Statistics**	**Level**	**Triple Negative Invasive Ductal Carcinoma N=35977**	**Apocrine Carcinoma N=2537**	**Parametric P-value***
Age Group	N (Col %)	<50	10362 (28.8)	381 (15.02)	**<.001**
	N (Col %)	>=50	25615 (71.2)	2156 (84.98)	
Facility Type	N (Col %)	Community Cancer Program/Other	4245 (11.8)	246 (9.7)	**<.001**
	N (Col %)	Comprehensive Community Cancer Program	19905 (55.33)	1493 (58.85)	
	N (Col %)	Academic/Research Program	11827 (32.87)	798 (31.45)	
Facility Location	N (Col %)	Northeast	6921 (19.24)	826 (32.56)	**<.001**
	N (Col %)	South	14832 (41.23)	765 (30.15)	
	N (Col %)	Midwest	8861 (24.63)	455 (17.93)	
	N (Col %)	West	5363 (14.91)	491 (19.35)	
Race	N (Col %)	White	26231 (72.91)	2050 (80.8)	**<.001**
	N (Col %)	Black	8132 (22.6)	316 (12.46)	
	N (Col %)	Others/Unknown	1614 (4.49)	171 (6.74)	
Primary Payer	N (Col %)	Not Insured/Unknown	1724 (4.79)	74 (2.92)	**<.001**
	N (Col %)	Private	19542 (54.32)	1142 (45.01)	
	N (Col %)	Medicaid	3512 (9.76)	130 (5.12)	
	N (Col %)	Medicare/Other Government	11199 (31.13)	1191 (46.95)	

Covariate	Statistics	Level	Histology: Triple Negative Invasive Ductal Carcinoma N=35977	Histology: Apocrine Carcinoma N=2537	Parametric P-value*
Year of Diagnosis	N (Col %)	2003	23 (0.06)	232 (9.14)	**<.001**
	N (Col %)	2004	114 (0.32)	267 (10.52)	
	N (Col %)	2005	132 (0.37)	319 (12.57)	
	N (Col %)	2006	210 (0.58)	341 (13.44)	
	N (Col %)	2007	349 (0.97)	273 (10.76)	
	N (Col %)	2008	555 (1.54)	247 (9.74)	
	N (Col %)	2009	2087 (5.8)	308 (12.14)	
	N (Col %)	2010	15647 (43.49)	264 (10.41)	
	N (Col %)	2011	16860 (46.86)	286 (11.27)	
Tumor Grade	N (Col %)	Well Differentiated	598 (1.66)	180 (7.09)	**<.001**
	N (Col %)	Moderately Differentiated	5690 (15.82)	1155 (45.53)	
	N (Col %)	Poorly Differentiated/Undifferentiated	28313 (78.7)	1018 (40.13)	
	N (Col %)	Cell Type Not Determined	1376 (3.82)	184 (7.25)	
Urban/Rural 2003	N (Col %)	Metro	29581 (84.81)	2150 (88.4)	**<.001**
	N (Col %)	Urban	4702 (13.48)	257 (10.57)	
	N (Col %)	Rural	598 (1.71)	25 (1.03)	
Median Income Quartiles 2000	N (Col %)	< $30,000	5098 (14.73)	309 (12.66)	**<.001**
	N (Col %)	$30,000 - $35,999	5958 (17.22)	365 (14.96)	
	N (Col %)	$36,000 - $45,999	9395 (27.15)	619 (25.37)	
	N (Col %)	$46,000 +	14156 (40.91)	1147 (47.01)	
Percent No High School Degree Quartiles 2000	N (Col %)	>=29%	6514 (18.82)	374 (15.33)	**<.001**
	N (Col %)	20-28.9%	8045 (23.25)	484 (19.84)	
	N (Col %)	14-19.9%	7705 (22.27)	562 (23.03)	
	N (Col %)	< 14%	12340 (35.66)	1020 (41.8)	
Great Circle Distance (quartile)	N (Col %)	>=0, <=5	8776 (24.39)	746 (29.4)	**<.001**
	N (Col %)	>5, <=9	8859 (24.62)	663 (26.13)	
	N (Col %)	>9, <=19	8947 (24.87)	571 (22.51)	
	N (Col %)	>19, <=3868	8955 (24.89)	508 (20.02)	
	N (Col %)	Unknown	440 (1.22)	49 (1.93)	
Charlson-Deyo Score	N (Col %)	0	30169 (83.86)	2179 (85.89)	**0.007**
	N (Col %)	1+	5808 (16.14)	358 (14.11)	

Covariate	Statistics	Level	Histology: Triple Negative Invasive Ductal Carcinoma N=35977	Histology: Apocrine Carcinoma N=2537	Parametric P-value*
Tumor Size (quartile)	N (Col %)	>=0, <=2	9113 (25.33)	861 (33.94)	**<.001**
	N (Col %)	>2, <=3	8568 (23.82)	623 (24.56)	
	N (Col %)	>3, <=4	8615 (23.95)	472 (18.6)	
	N (Col %)	>4, <=99	8843 (24.58)	472 (18.6)	
	N (Col %)	Unknown	838 (2.33)	109 (4.3)	
LN Examined (quartile)	N (Col %)	>=1, <=2	10208 (28.37)	669 (26.37)	**0.009**
	N (Col %)	>2, <=4	6443 (17.91)	419 (16.52)	
	N (Col %)	>4, <=11	7272 (20.21)	516 (20.34)	
	N (Col %)	>11, <=76	7566 (21.03)	580 (22.86)	
	N (Col %)	Unknown	4488 (12.47)	353 (13.91)	
Lymph Node Positive	N (Col %)	Negative	21493 (59.74)	1382 (54.47)	**<.001**
	N (Col %)	Positive	10895 (30.28)	864 (34.06)	
	N (Col %)	Unknown	3589 (9.98)	291 (11.47)	
Metastatic Bone Involvement	N (Col %)	None	31569 (97.11)	530 (96.36)	0.218
	N (Col %)	Yes	640 (1.97)	11 (2)	
	N (Col %)	Unknown	298 (0.92)	9 (1.64)	
Metastatic Brain Involvement	N (Col %)	None	32041 (98.57)	541 (98.36)	0.056
	N (Col %)	Yes	166 (0.51)	0 (0)	
	N (Col %)	Unknown	300 (0.92)	9 (1.64)	
Metastatic Liver Involvement	N (Col %)	None	31747 (97.66)	537 (97.64)	**0.017**
	N (Col %)	Yes	470 (1.45)	3 (0.55)	
	N (Col %)	Unknown	290 (0.89)	10 (1.82)	
Metastatic Lung Involvement	N (Col %)	None	31560 (97.09)	534 (97.09)	0.106
	N (Col %)	Yes	650 (2)	7 (1.27)	
	N (Col %)	Unknown	297 (0.91)	9 (1.64)	
Lymph Vascular Invasion	N (Col %)	Not present	19357 (59.55)	315 (57.27)	0.547
	N (Col %)	Prsent	6456 (19.86)	117 (21.27)	
	N (Col %)	Unknown	6694 (20.59)	118 (21.45)	
Estrogen Receptor (ER) Assay	N (Row %)	Negative	35533 (95.86)	1536 (4.14)	**<.001**
	N (Row %)	Positive	340 (33.33)	680 (66.67)	
	N (Row %)	Unknown	104 (24.47)	321 (75.53)	

Covariate	Statistics	Level	Histology: Triple Negative Invasive Ductal Carcinoma N=35977	Histology: Apocrine Carcinoma N=2537	Parametric P-value*
Progesterone Receptor (PR) Assay	N (Row %)	Negative	35564 (95.51)	1672 (4.49)	**<.001**
	N (Row %)	Positive	311 (36.89)	532 (63.11)	
	N (Row %)	Unknown	102 (23.45)	333 (76.55)	
HER2: Summary Result of Testing	N (Row %)	Negative	34045 (98.84)	401 (1.16)	**<.001**
	N (Row %)	Positive	213 (61.92)	131 (38.08)	
	N (Row %)	Unknown	1719 (46.16)	2005 (53.84)	
Triple Negative	N (Col %)	No	0 (0)	523 (20.61)	**<.001**
	N (Col %)	Yes	35977 (100)	299 (11.79)	
	N (Col %)	Unknown	0 (0)	1715 (67.6)	
Surgery of Primary Site	N (Col %)	No	8242 (22.91)	482 (19)	**<.001**
	N (Col %)	Yes	27712 (77.03)	2052 (80.88)	
	N (Col %)	Unknown	23 (0.06)	3 (0.12)	
Chemotherapy	N (Col %)	No	7945 (22.08)	1108 (43.67)	**<.001**
	N (Col %)	Yes	27360 (76.05)	1334 (52.58)	
	N (Col %)	Unknown	672 (1.87)	95 (3.74)	
Hormone Therapy	N (Col %)	No	34146 (94.91)	1880 (74.1)	**<.001**
	N (Col %)	Yes	1200 (3.34)	527 (20.77)	
	N (Col %)	Unknown	631 (1.75)	130 (5.12)	
Radiation Therapy	N (Col %)	No	15669 (43.55)	1176 (46.35)	**<.001**
	N (Col %)	Yes	19887 (55.28)	1301 (51.28)	
	N (Col %)	Unknown	421 (1.17)	60 (2.36)	
Age at Diagnosis	N		35977	2537	**<.001**
	Mean		57.84	64.79	
	Median		57	65	
	Min		20	21	
	Max		90	90	
	Std Dev		13.62	13.78	
Great Circle Distance	N		35537	2488	0.765
	Mean		21.52	21.01	
	Median		8.7	7.3	
	Min		0	0	
	Max		3867.8	2591.6	
	Std Dev		80.72	108.89	

Covariate	Statistics	Level	Histology: Triple Negative Invasive Ductal Carcinoma N=35977	Histology: Apocrine Carcinoma N=2537	Parametric P-value*
Tumor Size (cm)	N		35139	2428	**<.001**
	Mean		2.73	2.38	
	Median		2.1	1.7	
	Min		0	0	
	Max		98.8	30	
	Std Dev		2.76	2.26	
Regional Lymph Nodes Examined	N		31489	2184	**0.004**
	Mean		7.36	7.82	
	Median		4	5	
	Min		1	1	
	Max		76	53	
	Std Dev		7.37	7.6	
Regional Lymph Nodes Positive	N		31761	2202	**<.001**
	Mean		1.28	1.68	
	Median		0	0	
	Min		0	0	
	Max		52	90	
	Std Dev		3.41	4.35	

* The parametric p-value is calculated by ANOVA for numerical covariates and chi-square test for categorical covariates.

Table 3. Univariate (UVA) and Multivariable (MVA) Survival Analysis for the Main Effect of Apocrine Carcinoma versus Triple Negative Invasive Ductal Carcinoma

			Overall Survival (Months)			
			UVA		MVA	
Covariate	Level	N	Hazard Ratio (95% CI)	HR P-value	Hazard Ratio (95% CI)	HR P-value
Histology	Triple Negative Invasive Ductal Carcinoma	35977	1.46 (1.33-1.61)	**<.001**	1.40 (1.22-1.60)	<.001
	Apocrine Carcinoma	2537	-	-	-	-

			Overall Survival (Months)			
			UVA		MVA	
Covariate	**Level**	**N**	**Hazard Ratio (95% CI)**	**HR P-value**	**Hazard Ratio (95% CI)**	**HR P-value**
Age Group	<50	10743	0.77 (0.72-0.81)	**<.001**		
	>=50	27771	-	-		
Facility Type	Academic/Research Program	12625	0.83 (0.76-0.90)	**<.001**	0.91 (0.83-0.99)	0.038
	Comprehensive Community Cancer Program	21398	0.84 (0.78-0.91)	**<.001**	0.99 (0.91-1.08)	0.761
	Community Cancer Program/Other	4491	-	-	-	-
Facility Location	West	5854	0.96 (0.87-1.05)	0.322	0.92 (0.84-1.02)	0.113
	Midwest	9316	1.18 (1.09-1.27)	**<.001**	1.09 (1.01-1.19)	0.035
	South	15597	1.16 (1.08-1.24)	**<.001**	0.98 (0.91-1.06)	0.675
	Northeast	7747	-	-	-	-
Race	Others/Unknown	1785	0.70 (0.60-0.80)	**<.001**	0.77 (0.66-0.91)	0.002
	Black	8448	1.35 (1.28-1.43)	**<.001**	1.22 (1.14-1.31)	<.001
	White	28281	-	-	-	-
Primary Payer	Medicare/Other Government	12390	1.11 (0.99-1.24)	0.067	0.91 (0.80-1.03)	0.146
	Medicaid	3642	1.13 (0.99-1.28)	0.066	1.09 (0.95-1.25)	0.209
	Private	20684	0.52 (0.46-0.58)	**<.001**	0.62 (0.55-0.71)	<.001
	Not Insured/Unknown	1798	-	-	-	-
Year of Diagnosis	2003	255	0.54 (0.42-0.69)	**<.001**		
	2004	381	0.69 (0.57-0.84)	**<.001**		
	2005	451	0.64 (0.53-0.77)	**<.001**		
	2006	551	0.59 (0.49-0.71)	**<.001**		
	2007	622	0.61 (0.51-0.73)	**<.001**		
	2008	802	0.85 (0.73-0.99)	**0.032**		
	2009	2395	0.85 (0.77-0.94)	**0.001**		
	2010	15911	0.94 (0.88-1.00)	**0.036**		
	2011	17146	-	-		

			Overall Survival (Months)			
			UVA		MVA	
Covariate	**Level**	**N**	**Hazard Ratio (95% CI)**	**HR P-value**	**Hazard Ratio (95% CI)**	**HR P-value**
Tumor Grade	Cell Type Not Determined	1560	2.25 (1.79-2.84)	**<.001**	1.90 (1.48-2.44)	<.001
	Poorly Differentiated/Undifferentiated	29331	1.72 (1.39-2.12)	**<.001**	1.78 (1.43-2.22)	<.001
	Moderately Differentiated	6845	1.39 (1.12-1.73)	**0.003**	1.31 (1.05-1.65)	0.019
	Well Differentiated	778	-	-	-	-
Urban/Rural 2003	Metro	31731	0.91 (0.75-1.11)	0.363	1.14 (0.93-1.41)	0.196
	Urban	4959	0.95 (0.77-1.16)	0.592	1.04 (0.84-1.29)	0.715
	Rural	623	-	-	-	-
Median Income Quartiles 2000	< $30,000	5407	1.61 (1.50-1.73)	**<.001**	1.37 (1.22-1.53)	<.001
	$30,000 - $35,999	6323	1.38 (1.28-1.48)	**<.001**	1.23 (1.11-1.35)	<.001
	$36,000 - $45,999	10014	1.31 (1.22-1.39)	**<.001**	1.19 (1.10-1.29)	<.001
	$46,000 +	15303	-	-	-	-
Percent No High School Degree Quartiles 2000	>=29%	6888	1.43 (1.33-1.54)	**<.001**	0.87 (0.78-0.97)	0.014
	20-28.9%	8529	1.33 (1.24-1.43)	**<.001**	1.01 (0.92-1.10)	0.818
	14-19.9%	8267	1.20 (1.12-1.29)	**<.001**	1.04 (0.96-1.13)	0.305
	< 14%	13360	-	-	-	-
Great Circle Distance (quartile)	>=0, <=5	9522	0.45 (0.38-0.54)	**<.001**		
	>5, <=9	9522	0.41 (0.34-0.49)	**<.001**		
	>9, <=19	9518	0.37 (0.30-0.44)	**<.001**		
	>19, <=3868	9463	0.37 (0.31-0.44)	**<.001**		
	Unknown	489	-	-		
Charlson-Deyo Score	1+	6166	1.75 (1.65-1.86)	**<.001**	1.55 (1.46-1.66)	<.001
	0	32348	-	-	-	-

			Overall Survival (Months)			
			UVA		MVA	
Covariate	**Level**	**N**	**Hazard Ratio (95% CI)**	**HR P-value**	**Hazard Ratio (95% CI)**	**HR P-value**
Tumor Size (quartile)	>=0, <=2	9974	0.12 (0.11-0.14)	**<.001**		
	>2, <=3	9191	0.18 (0.16-0.20)	**<.001**		
	>3, <=4	9087	0.27 (0.24-0.30)	**<.001**		
	>4, <=99	9315	0.56 (0.50-0.62)	**<.001**		
	Unknown	947	-	**-**		
LN Examined (quartile)	>=1, <=2	10877	0.17 (0.16-0.19)	**<.001**		
	>2, <=4	6862	0.16 (0.15-0.18)	**<.001**		
	>4, <=11	7788	0.29 (0.27-0.31)	**<.001**		
	>11, <=76	8146	0.41 (0.38-0.44)	**<.001**		
	Unknown	4841	-	**-**		
Lymph Node Postive	Negative	22875	0.13 (0.13-0.14)	**<.001**		
	Positive	11759	0.48 (0.45-0.51)	**<.001**		
	Unknown	3880	-	**-**		
Metastatic Bone Involvement	None	32099	0.32 (0.25-0.39)	**<.001**		
	Yes	651	3.85 (3.06-4.85)	**<.001**		
	Unknown	307	-	**-**		
Metatstatic Brain Involvement	None	32582	0.32 (0.26-0.40)	**<.001**		
	Yes	166	6.87 (5.27-8.96)	**<.001**		
	Unknown	309	-	**-**		
Metastatic Liver Involvement	None	32284	0.34 (0.27-0.42)	**<.001**		
	Yes	473	4.91 (3.85-6.26)	**<.001**		
	Unknown	300	-	**-**		
Metastatic Lung Involvement	None	32094	0.30 (0.24-0.37)	**<.001**		
	Yes	657	3.71 (2.96-4.64)	**<.001**		
	Unknown	306	-	**-**		

			Overall Survival (Months)			
			UVA		MVA	
Covariate	**Level**	**N**	**Hazard Ratio (95% CI)**	**HR P-value**	**Hazard Ratio (95% CI)**	**HR P-value**
Lymph Vascular Invasion	Not present	19672	0.43 (0.40-0.46)	**<.001**		
	Prsent	6573	1.17 (1.09-1.26)	**<.001**		
	Unknown	6812	-	-		
Estrogen Receptor (ER) Assay	Negative	37069	1.42 (1.17-1.73)	**<.001**		
	Positive	1020	0.96 (0.75-1.22)	0.727		
	Unknown	425	-	-		
Progesterone Receptor (PR) Assay	Negative	37236	1.45 (1.19-1.76)	**<.001**		
	Positive	843	1.02 (0.79-1.31)	0.883		
	Unknown	435	-	-		
HER2: Summary Result of Testing	Negative	34446	1.24 (1.15-1.34)	**<.001**	1.03 (0.93-1.15)	0.593
	Positive	344	0.88 (0.64-1.21)	0.445	0.72 (0.50-1.04)	0.081
	Unknown	3724	-	-	-	-
Triple Negative	Unknown	1715	0.95 (0.77-1.18)	0.659		
	Yes	36276	1.44 (1.18-1.75)	**<.001**		
	No	523	-	-		
Surgery of Primary Site	Unknown	26	1.29 (0.61-2.70)	0.506	1.83 (0.75-4.50)	0.185
	Yes	29764	0.52 (0.49-0.55)	**<.001**	0.50 (0.47-0.53)	<.001
	No	8724	-	-		
Chemotherapy	Unknown	0.84 (0.71-1.01)	0.057	0.84 (0.71-1.01)	0.93 (0.76-1.14)	0.513
	Yes	0.60 (0.57-0.64)	**<.001**	0.60 (0.57-0.64)	0.76 (0.71-0.81) -	<.001
	No	-	-	-		-
Hormone Therapy	Unknown	761	0.73 (0.61-0.89)	**0.002**	0.84 (0.68-1.04)	0.106
	Yes	1727	0.68 (0.59-0.77)	**<.001**	0.77 (0.67-0.88)	<.001
	No	36026	-	-	-	-

			Overall Survival (Months)			
			UVA		MVA	
Covariate	**Level**	**N**	**Hazard Ratio (95% CI)**	**HR P-value**	**Hazard Ratio (95% CI)**	**HR P-value**
Radiation Therapy	Unknown	481	0.58 (0.46-0.74)	**<.001**	0.63 (0.48-0.82)	<.001
	Yes	21188	0.50 (0.48-0.53)	**<.001**	0.52 (0.49-0.55)	<.001
	No	16845	-	-	-	-
Age at Diagnosis		38514	1.02 (1.02-1.03)	**<.001**	1.01 (1.01-1.02)	<.001
Great Circle Distance		38025	1.00 (1.00-1.00)	**0.010**	-	
Tumor Size (cm)		37567	1.05 (1.05-1.05)	**<.001**	1.05 (1.05-1.05)	<.001
Regional Lymph Nodes Examined		33673	1.04 (1.04-1.04)	**<.001**		
Regional Lymph Nodes Positive		33963	1.06 (1.06-1.06)	**<.001**		

* Number of observations in the original data set = 38514. Number of observations used = 33550.

** Backward selection with an alpha level of removal of .20 was used. The following variables were removed from the model: Estrogen Receptor (ER) Assay, Progesterone Receptor (PR) Assay, Surgery of Primary Site, and Urban/Rural 2003.

Table 4. Multivariable Survival Analysis of Overall Survival – Interaction with Clinical Characteristics, Receptor Statues and Treatment Approaches

		Overall Survival (Months)		
Covariate	**Level**	**Hazard Ratio**	**HR P-value**	**Type3 P-value**
Stratified Comparisons by Age Group (INTERACTION):		-	-	**0.018**
<50	Apocrine Carcinoma vs. Triple Negative Invasive Ductal Carcinoma	0.50 (0.35-0.71)	**<.001**	-

		Overall Survival (Months)		
Covariate	**Level**	**Hazard Ratio**	**HR P-value**	**Type3 P-value**
>=50	Apocrine Carcinoma vs. Triple Negative Invasive Ductal Carcinoma	0.77 (0.67-0.88)	**<.001**	-
Stratified Comparisons by Race (INTERACTION):		-	-	0.250
White	Apocrine Carcinoma vs. Triple Negative Invasive Ductal Carcinoma	0.75 (0.64-0.87)	**<.001**	-
Black	Apocrine Carcinoma vs. Triple Negative Invasive Ductal Carcinoma	0.64 (0.49-0.83)	**<.001**	-
Stratified Comparisons by Estrogen Receptor (ER) Assay (INTERACTION):		-	-	0.849
Negative	Apocrine Carcinoma vs. Triple Negative Invasive Ductal Carcinoma	0.76 (0.65-0.88)	**<.001**	-
Positive	Apocrine Carcinoma vs. Triple Negative Invasive Ductal Carcinoma	0.73 (0.51-1.05)	0.086	-
Stratified Comparisons by Progesterone Receptor (PR) Assay (INTERACTION):		-	-	0.528
Negative	Apocrine Carcinoma vs. Triple Negative Invasive Ductal Carcinoma	0.76 (0.66-0.88)	**<.001**	-
Positive	Apocrine Carcinoma vs. Triple Negative Invasive Ductal Carcinoma	0.68 (0.46-0.98)	**0.041**	-
Stratified Comparisons by HER2: Summary Result of Testing (INTERACTION):		-	-	0.203
Negative	Apocrine Carcinoma vs. Triple Negative Invasive Ductal Carcinoma	0.90 (0.68-1.19)	0.446	-
Positive	Apocrine Carcinoma vs. Triple Negative Invasive Ductal Carcinoma	0.52 (0.23-1.15)	0.107	-
Stratified Comparisons by Chemotherapy (INTERACTION):		-	-	**0.037**

		Overall Survival (Months)		
Covariate	**Level**	**Hazard Ratio**	**HR P-value**	**Type3 P-value**
No	Apocrine Carcinoma vs. Triple Negative Invasive Ductal Carcinoma	0.81 (0.69-0.95)	**0.010**	-
Yes	Apocrine Carcinoma vs. Triple Negative Invasive Ductal Carcinoma	0.66 (0.55-0.79)	**<.001**	-
Stratified Comparisons by Hormone Therapy (INTERACTION):		-	-	0.564
No	Apocrine Carcinoma vs. Triple Negative Invasive Ductal Carcinoma	0.74 (0.64-0.86)	**<.001**	-
Yes	Apocrine Carcinoma vs. Triple Negative Invasive Ductal Carcinoma	0.81 (0.60-1.10)	0.171	-
Stratified Comparisons by Radiation Therapy (INTERACTION):		-	-	0.921
No	Apocrine Carcinoma vs. Triple Negative Invasive Ductal Carcinoma	0.72 (0.62-0.84)	**<.001**	-
Yes	Apocrine Carcinoma vs. Triple Negative Invasive Ductal Carcinoma	0.72 (0.60-0.86)	**<.001**	-

*Backward selection with an alpha level of removal of .2 was used. The following variables were removed from the model: Estrogen Receptor (ER) Assay, Progesterone Receptor (PR) Assay.

Table 5. Characteristics of Patients after Matching Apocrine Carcinoma vs. TNBC

		Histology			
Covariate	**Level**	**Triple Negative Invasive Ductal Carcinoma N=1192**	**Apocrine Carcinoma N=1192**	**Parametric P-value***	**Standardized Difference**
Facility Type	Community Cancer Program/Other	117 (9.82)	113 (9.48)	0.880	0.011
	Comprehensive Community Cancer Program	635 (53.27)	647 (54.28)		0.020
	Academic/Research Program	440 (36.91)	432 (36.24)		0.014

Facility Location	Northeast	350 (29.36)	340 (28.52)	0.933	0.018
	South	409 (34.31)	405 (33.98)		0.007
	Midwest	255 (21.39)	260 (21.81)		0.010
	West	178 (14.93)	187 (15.69)		0.021
Race	White	940 (78.86)	924 (77.52)	0.679	0.033
	Black	188 (15.77)	196 (16.44)		0.018
	Others/Unknown	64 (5.37)	72 (6.04)		0.029
Primary Payer	Not Insured/Unknown	33 (2.77)	38 (3.19)	0.706	0.025
	Private	572 (47.99)	549 (46.06)		0.039
	Medicaid	73 (6.12)	69 (5.79)		0.014
	Medicare/Other Government	514 (43.12)	536 (44.97)		0.037
Tumor Grade	Well Differentiated	59 (4.95)	71 (5.96)	0.683	0.044
	Moderately Differentiated	454 (38.09)	437 (36.66)		0.029
	Poorly Differentiated/Undifferentiated	602 (50.5)	609 (51.09)		0.012
	Cell Type Not Determined	77 (6.46)	75 (6.29)		0.007
Urban/Rural 2003	Metro	1031 (86.49)	1034 (86.74)	0.476	0.007
	Urban	149 (12.5)	140 (11.74)		0.023
	Rural	12 (1.01)	18 (1.51)		0.045
Median Income Quartiles 2000	< $30,000	168 (14.09)	176 (14.77)	0.479	0.019
	$30,000 - $35,999	197 (16.53)	192 (16.11)		0.011
	$36,000 - $45,999	283 (23.74)	311 (26.09)		0.054
	$46,000 +	544 (45.64)	513 (43.04)		0.052
Percent No High School Degree Quartiles 2000	>=29%	202 (16.95)	206 (17.28)	0.971	0.009
	20-28.9%	256 (21.48)	255 (21.39)		0.002
	14-19.9%	268 (22.48)	275 (23.07)		0.014
	< 14%	466 (39.09)	456 (38.26)		0.017
Charlson-Deyo Score	0	1002 (84.06)	997 (83.64)	0.781	0.011
	1+	190 (15.94)	195 (16.36)		0.011
Estrogen Receptor (ER) Assay	Negative	1053 (88.34)	1039 (87.16)	0.671	0.036
	Positive	112 (9.4)	122 (10.23)		0.028
	Unknown	27 (2.27)	31 (2.6)		0.022
Progesterone Receptor (PR) Assay	Negative	1062 (89.09)	1046 (87.75)	0.588	0.042
	Positive	104 (8.72)	116 (9.73)		0.035
	Unknown	26 (2.18)	30 (2.52)		0.022
HER2: Summary Result of Testing	Negative	362 (30.37)	369 (30.96)	0.708	0.013
	Positive	86 (7.21)	76 (6.38)		0.033
	Unknown	744 (62.42)	747 (62.67)		0.005
Surgery of Primary Site	No	217 (18.2)	217 (18.2)	0.606	0.000
	Yes	975 (81.8)	974 (81.71)		0.002
Chemotherapy	No	437 (36.66)	442 (37.08)	0.863	0.009
	Yes	723 (60.65)	722 (60.57)		0.002
	Unknown	32 (2.68)	28 (2.35)		0.021

Hormone Therapy	No	1032 (86.58)	1040 (87.25)	0.780	0.020
	Yes	128 (10.74)	118 (9.9)		0.028
	Unknown	32 (2.68)	34 (2.85)		0.010
Radiation Therapy	No	539 (45.22)	551 (46.22)	0.882	0.020
	Yes	630 (52.85)	619 (51.93)		0.018
	Unknown	23 (1.93)	22 (1.85)		0.006
Age at Diagnosis		63.68 (13.71)	63.66 (13.72)	0.969	0.002
Tumor Size (cm)		Mean (Std)	2.56 (2.49)	2.54 (2.43)	0.803

Table 6. Association with Survival for Apocrine Carcinoma versus TNBC in Matched Sample

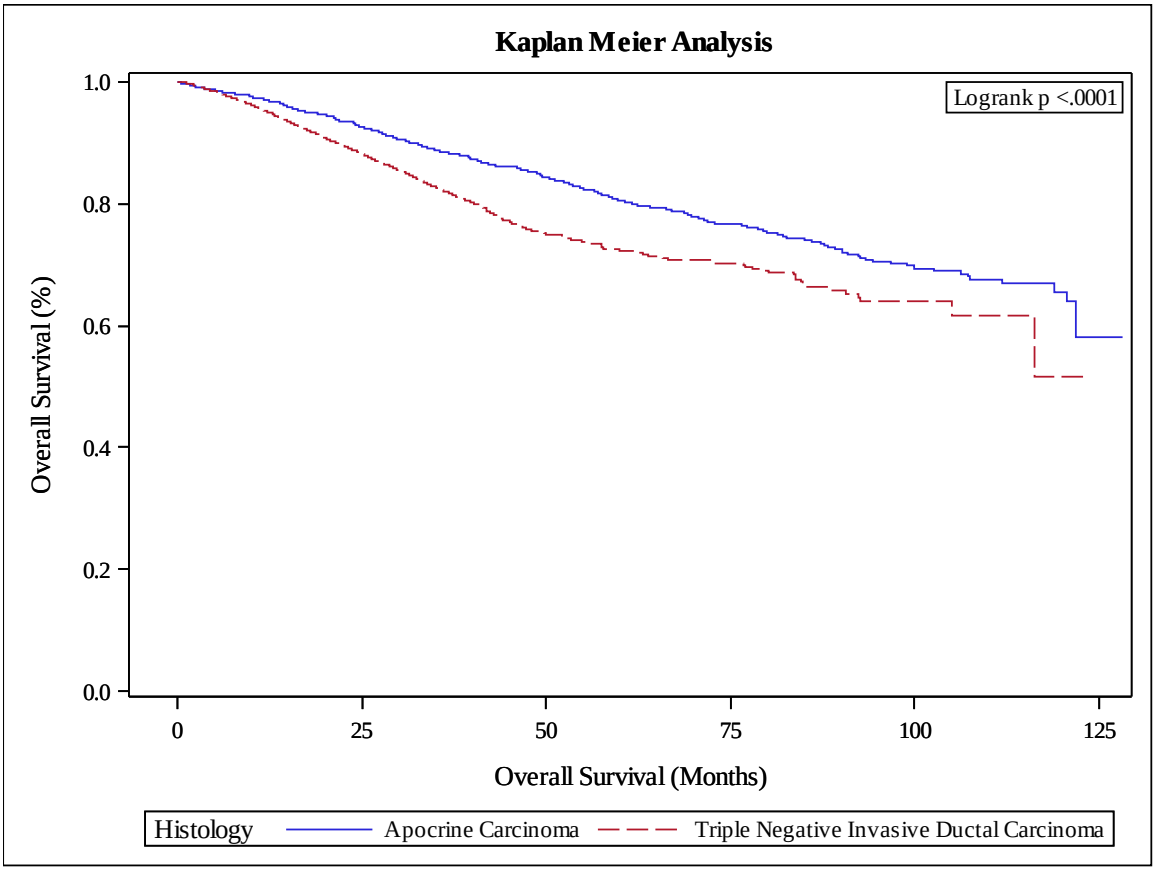
			Overall Survival (Months)		
Covariate	**Level**	**N**	**Hazard Ratio (95% CI)**	**HR P-value**	**Log-rank P-value**
Histology	Triple Negative Invasive Ductal Carcinoma	1192	1.34 (1.12-1.61)	**0.001**	**0.002**
	Apocrine Carcinoma	1192	-	-	

B. Figures

Figure 1. Diagram of Study Population Selection and Exclusion

Selection and Exclusion Criteria	Sample Size	Excluded
NCDB Breast PUF Cancer Cases	2807541	-
Year of diagnosis 2003 ~ 2011	1743422	1064119
Exclude Behavior in situ	1389281	354141
Exclude Male Patients	1375462	13819
Exclude Palliative Care	1369095	6367
Included Eligible Histology	38518	1330577
Exclude Missing Outcome	38514	4

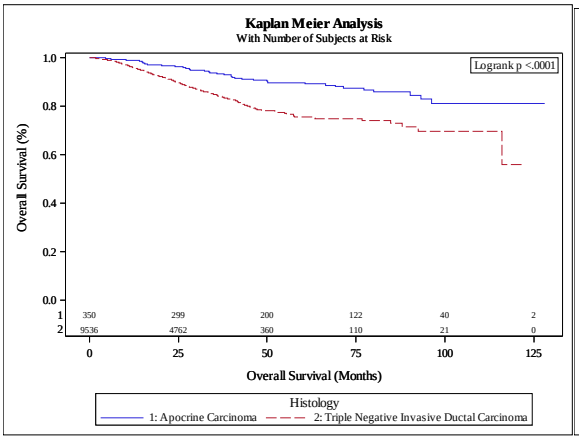
Figure 2. KM Plot for All Patients



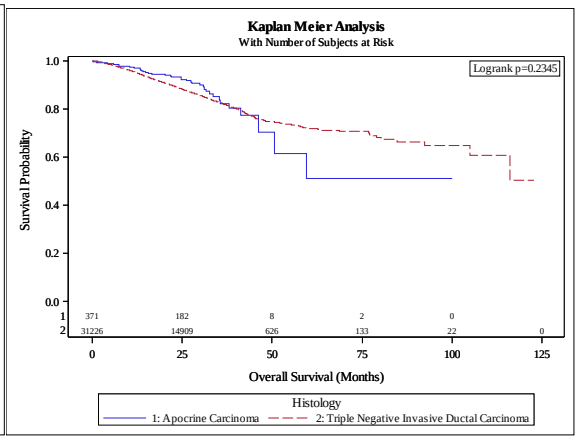
Histology	No. of Subject	Event	Censored	Median Survival (95% CI)	12 Mo Survival	60 Mo Survival
Apocrine Carcinoma	2317	448 (19%)	1869 (81%)	NA (121.9, NA)	97.0% (96.2%, 97.6%)	80.6% (78.7%, 82.4%)
Triple Negative Invasive Ductal Carcinoma	32982	4307 (13%)	28675 (87%)	NA (116.1, NA)	95.2% (95.0%, 95.5%)	72.4% (70.9%, 73.8%)

Figure 3. KM Plot between Apocrine Carcinoma and TNBC Stratified by Clinical Characteristics, Receptor Status and Treatment Approaches

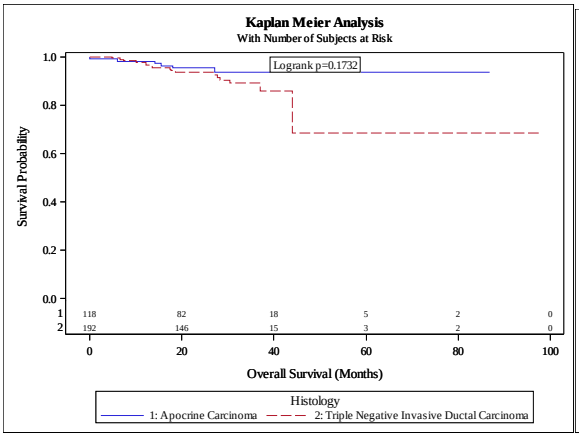
<50



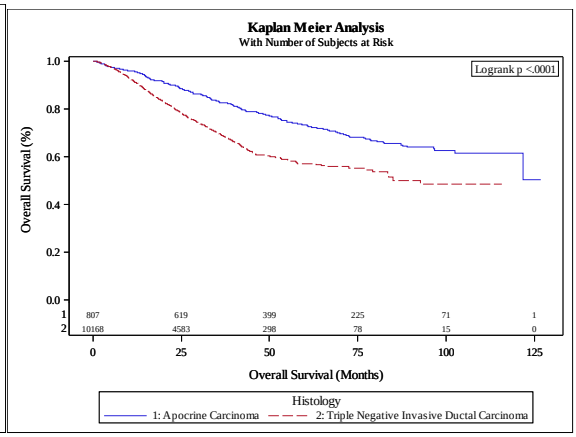
HER2 (-)



HER2(+)



LN(+)



Histology	No. of Subject	Event	Censored	Median Survival (95% CI)	24 Mo Survival	60 Mo Survival
Age (<50) Apocrine Carcinoma	350	39 (11%)	311 (89%)	NA (NA, NA)	96.7% (94.2%, 98.2%)	89.7% (85.6%, 92.8%)
Triple Negative Invasive Ductal Carcinoma	9536	1091 (11%)	8445 (89%)	NA (116.1, NA)	90.2% (89.6%, 90.9%)	75.7% (73.2%, 78.0%)
Age (>=50) Apocrine Carcinoma	1967	409 (21%)	1558 (79%)	NA (121.9, NA)	92.7% (91.4%, 93.8%)	78.9% (76.7%, 80.9%)
Triple Negative Invasive Ductal Carcinoma	23446	3216 (14%)	20230 (86%)	NA (104.9, NA)	88.2% (87.7%, 88.6%)	70.9% (69.0%, 72.7%)
Ethnicity (White) Apocrine Carcinoma	1895	383 (20%)	1512 (80%)	NA (121.9, NA)	93.0% (91.7%, 94.1%)	80.1% (77.9%, 82.1%)

Triple Negative Invasive Ductal Carcinoma	24325	3028 (12%)	21297 (88%)	NA (NA, NA)	89.3% (88.9%, 89.7%)	73.4% (71.7%, 75.0%)
Ethnicity (Black) Apocrine Carcinoma	271	54 (20%)	217 (80%)	NA (NA, NA)	93.0% (89.1%, 95.5%)	77.3% (70.6%, 82.6%)
Triple Negative Invasive Ductal Carcinoma	7260	1163 (16%)	6097 (84%)	116.1 (NA, NA)	86.3% (85.4%, 87.1%)	67.5% (64.3%, 70.5%)
ER Negative Apocrine Carcinoma	1422	281 (20%)	1141 (80%)	NA (NA, NA)	92.2% (90.7%, 93.6%)	78.9% (76.2%, 81.3%)
Triple Negative Invasive Ductal Carcinoma	32589	4250 (13%)	28339 (87%)	NA (NA, NA)	88.8% (88.4%, 89.1%)	72.3% (70.8%, 73.7%)
ER Positive Apocrine Carcinoma	616	90 (15%)	526 (85%)	NA (NA, NA)	95.3% (93.2%, 96.7%)	83.9% (80.2%, 87.1%)
Triple Negative Invasive Ductal Carcinoma	305	39 (13%)	266 (87%)	NA (72.3, NA)	89.2% (84.7%, 92.5%)	75.9% (63.0%, 84.8%)
PR Negative Apocrine Carcinoma	1546	298 (19%)	1248 (81%)	NA (NA, NA)	92.3% (90.8%, 93.5%)	79.2% (76.6%, 81.5%)
Triple Negative Invasive Ductal Carcinoma	32620	4253 (13%)	28367 (87%)	NA (NA, NA)	88.8% (88.4%, 89.1%)	72.3% (70.8%, 73.8%)
RP Positive Apocrine Carcinoma	483	74 (15%)	409 (85%)	NA (NA, NA)	95.6% (93.3%, 97.2%)	83.8% (79.6%, 87.2%)
Triple Negative Invasive Ductal Carcinoma	275	36 (13%)	239 (87%)	72.3 (72.3, NA)	89.4% (84.6%, 92.7%)	76.1% (63.3%, 84.9%)
HER2 Negative Apocrine Carcinoma	371	40 (11%)	331 (89%)	NA (50.6, NA)	93.3% (90.0%, 95.6%)	51.2% (24.6%, 72.7%)
Triple Negative Invasive Ductal Carcinoma	31226	3969 (13%)	27257 (87%)	NA (116.1, NA)	88.8% (88.4%, 89.2%)	72.0% (70.1%, 73.8%)
HER2 Positive Apocrine Carcinoma	118	6 (5%)	112 (95%)	NA (NA, NA)	95.3% (89.1%, 98.0%)	93.6% (85.8%, 97.2%)
Triple Negative Invasive Ductal Carcinoma	192	17 (9%)	175 (91%)	NA (44.1, NA)	93.7% (88.9%, 96.5%)	68.6% (28.9%, 89.2%)
Chemo YES Apocrine Carcinoma	1245	183 (15%)	1062 (85%)	NA (NA, NA)	95.7% (94.4%, 96.8%)	85.0% (82.5%, 87.1%)
Triple Negative Invasive Ductal Carcinoma	25608	3030 (12%)	22578 (88%)	116.1 (116.1, NA)	89.9% (89.5%, 90.3%)	74.7% (73.0%, 76.2%)
Chemo NO Apocrine Carcinoma	989	255 (26%)	734 (74%)	121.9 (118.8, NA)	89.8% (87.7%, 91.5%)	74.3% (71.0%, 77.3%)
Triple Negative Invasive Ductal Carcinoma	6918	1205 (17%)	5713 (83%)	NA (92.3, NA)	84.7% (83.7%, 85.6%)	64.4% (60.9%, 67.7%)
Hormone YES Apocrine Carcinoma	500	80 (16%)	420 (84%)	NA (121.7, NA)	96.0% (93.8%, 97.4%)	85.3% (81.3%, 88.5%)
Triple Negative Invasive Ductal Carcinoma	1118	123 (11%)	995 (89%)	NA (92.3, NA)	91.2% (89.2%, 92.8%)	78.8% (72.8%, 83.6%)
Hormone NO Apocrine Carcinoma	1703	356 (21%)	1347 (79%)	NA (121.9, NA)	92.2% (90.8%, 93.4%)	78.4% (76.0%, 80.6%)

Triple Negative Invasive Ductal Carcinoma	31357	4122 (13%)	27235 (87%)	NA (116.1, NA)	88.7% (88.3%, 89.1%)	71.9% (70.3%, 73.4%)
Radiation YES	1266	199 (16%)	1067 (84%)	NA (NA, NA)	95.6% (94.3%, 96.6%)	85.4% (83.0%, 87.6%)
Apocrine Carcinoma						
Triple Negative Invasive Ductal Carcinoma	19281	2141 (11%)	17140 (89%)	116.1 (116.1, NA)	91.2% (90.8%, 91.7%)	75.5% (73.6%, 77.2%)
Radiation NO	1000	243 (24%)	757 (76%)	121.9 (118.8, NA)	90.0% (87.9%, 91.8%)	73.5% (70.2%, 76.6%)
Apocrine Carcinoma						
Triple Negative Invasive Ductal Carcinoma	13388	2125 (16%)	11263 (84%)	NA (NA, NA)	85.0% (84.3%, 85.7%)	67.6% (65.0%, 70.0%)
LN Negative	1358	173 (13%)	1185 (87%)	NA (NA, NA)	96.8% (95.7%, 97.7%)	87.4% (85.1%, 89.4%)
Apocrine Carcinoma						
Triple Negative Invasive Ductal Carcinoma	20999	1483 (7%)	19516 (93%)	NA (116.1, NA)	94.6% (94.2%, 94.9%)	81.3% (79.4%, 83.0%)
LN Positive	807	213 (26%)	594 (74%)	NA (121.7, NA)	89.5% (87.1%, 91.5%)	**73.3% (69.7%, 76.6%)**
Apocrine Carcinoma						
Triple Negative Invasive Ductal Carcinoma	10168	2391 (24%)	7777 (76%)	92.6 (79.2, NA)	79.3% (78.4%, 80.2%)	**57.1% (54.5%, 59.7%)**
LVI Not Present	302	21 (7%)	281 (93%)	46.3 (NA, NA)	94.8% (91.3%, 96.9%)	0.0% (NA, NA)
Apocrine Carcinoma						
Triple Negative Invasive Ductal Carcinoma	18462	1552 (8%)	16910 (92%)	NA (47.9, NA)	92.6% (92.1%, 93.0%)	NA (NA, NA)
LVI Present	112	13 (12%)	99 (88%)	NA (NA, NA)	92.6% (85.0%, 96.4%)	NA (NA, NA)
Apocrine Carcinoma						
Triple Negative Invasive Ductal Carcinoma	6141	1432 (23%)	4709 (77%)	NA (45, NA)	78.1% (76.9%, 79.2%)	NA (NA, NA)

Figure 4. Distribution of the logit of the Propensity Scores Calculated for the TNBC (top) and Apocrine Carcinoma (bottom) Cohorts

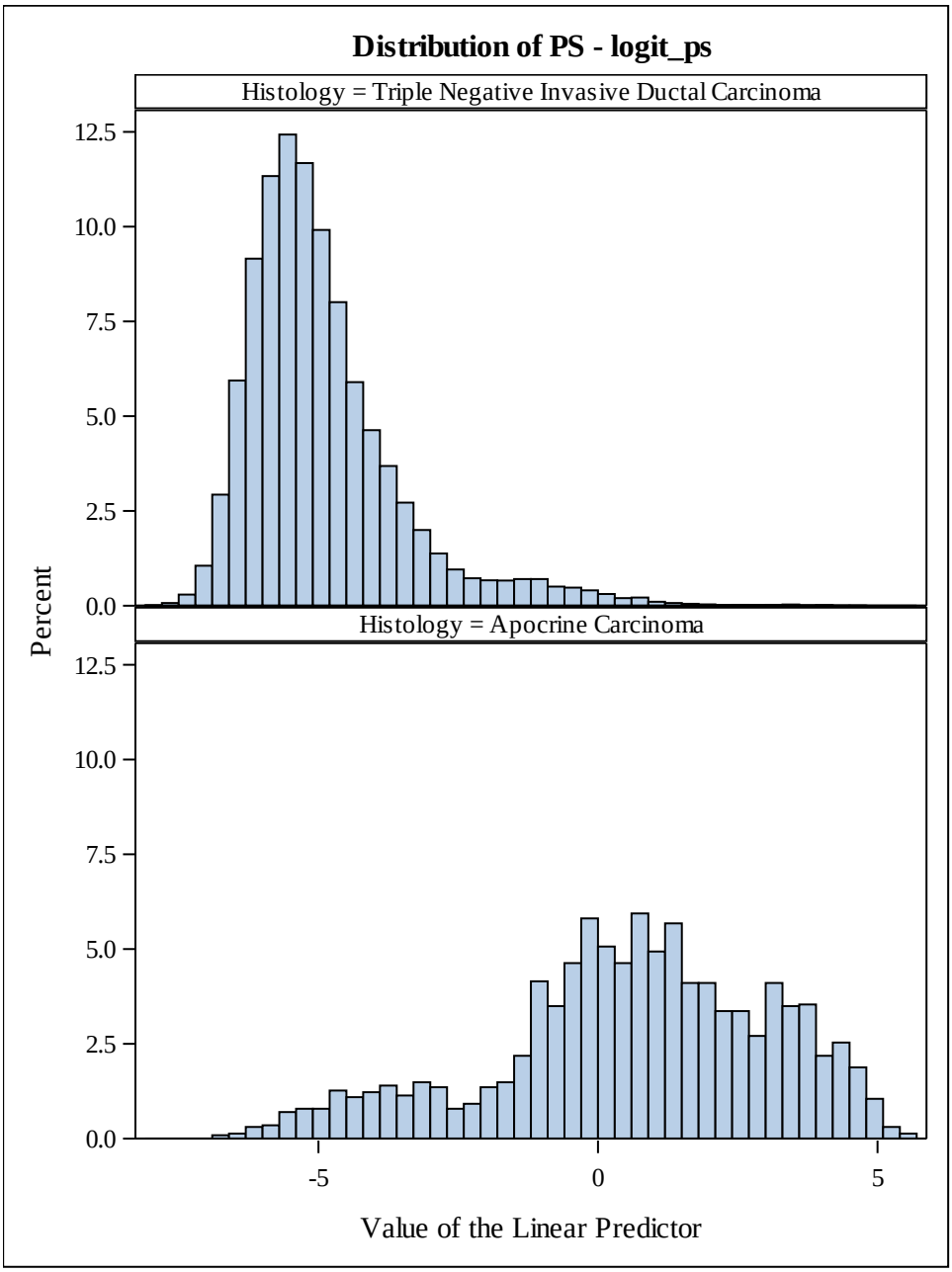
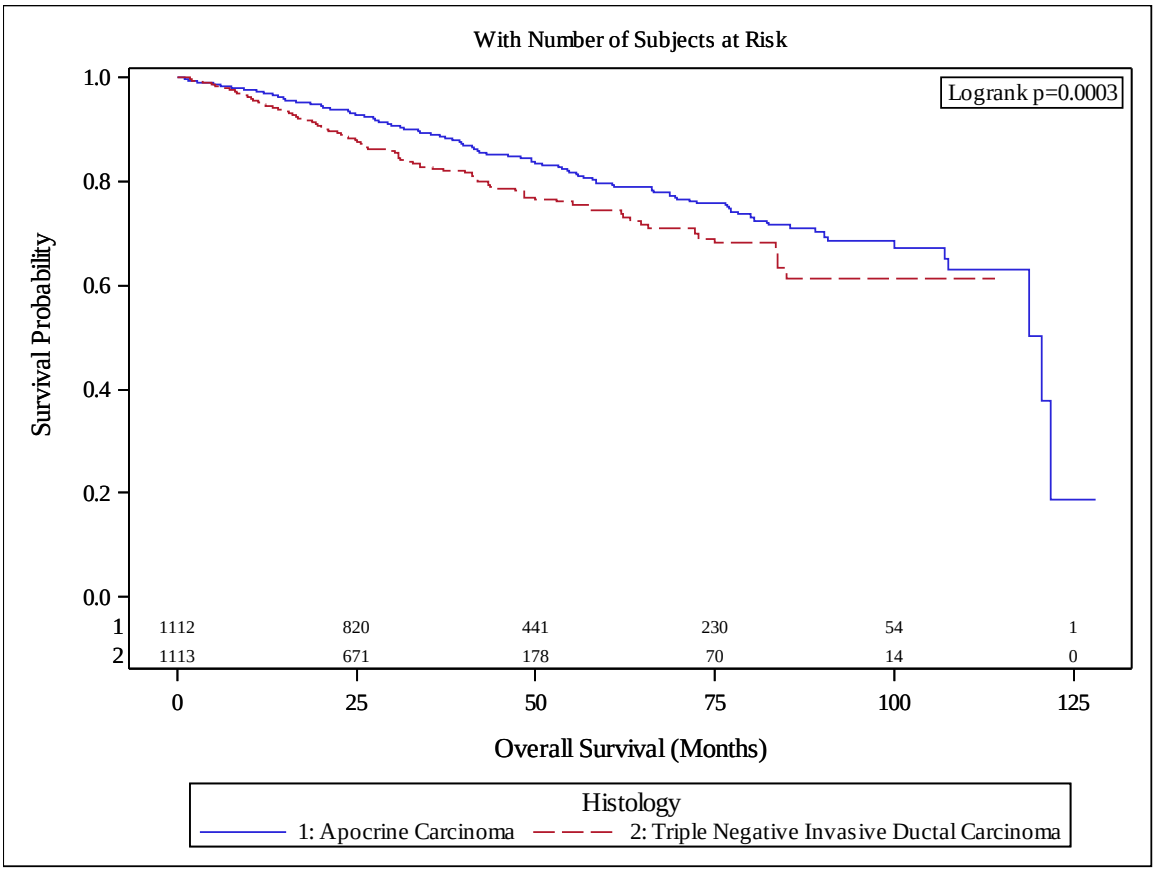


Figure 5. KM Plot for Apocrine Carcinoma versus Triple Negative Invasive Ductal Carcinoma in Matched Sample



Histology	No. of Subject	Event	Censored	Median Survival (95% CI)	12 Mo Survival	60 Mo Survival
Apocrine Carcinoma	1112	195 (18%)	917 (82%)	120.7 (118.8, NA)	97.1% (95.9%, 97.9%)	79.5% (76.2%, 82.3%)
Triple Negative Invasive Ductal Carcinoma	1113	187 (17%)	926 (83%)	NA (NA, NA)	94.9% (93.4%, 96.0%)	74.4% (70.0%, 78.2%)