Journal of Cancer and Tumor International

12(1): 16-28, 2022; Article no.JCTI.83754
ISSN: 2454-7360

A Comparison between Accelerated Failure Time Models in Analyzing the Survival of Breast Cancer Patients

Bosson-Amedenu Senyefia [a*], Acquah Joseph [a], Nyarko Christiana Cynthia [a] and Noureddine Ouerfelli [b]

[a] Department of Mathematical Sciences, University of Mines and Technology, P.O. Box 237, Tarkwa, Ghana.
[b] Institut Supérieur des Technologies Médicales de Tunis, LR13SE07, Laboratoire de Biophysique et Technologies Médicales, Université de Tunis El Manar, Tunis, Tunisia.

Authors' contributions

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/JCTI/2022/v12i130166

Open Peer Review History:
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://www.sdiarticle5.com/review-history/83754

Original Research Article

Received 15 November 2021
Accepted 19 January 2022
Published 20 January 2022

ABSTRACT

Background: Female breast cancer (BC) has surpassed lung cancer as the most prevalent reason for cancer-related diagnosis in the world. BC has geographical disparities in the intensity of effect of its associated risk factors on patients' survival. Several models can be employed to determine the effect of risk factors on patients' survival. The present study aims at evaluating these models.
Methods: The secondary data of 558 BC patients diagnosed at Korle Bu teaching hospital during 2010-2015 and followed-up (right censored) to the end of 2015 were analysed. The survival status, demographic and tumour characteristics of these patients were determined by event history analysis. To compare various models of survival, Akaike Information Criterion (AIC) , Bayesian Information Criteria (BIC) and Receiver Operation Characteristic (ROC) curve were used. R software was used for data analyses. The data consisted of BC patients in the age range of 13 to 97 years. The dataset was partitioned into training (holding 70%) and validation set (30%).
Results: Based on AIC, BIC and ROC curve values the Gompertz (AIC=2322, BIC=2391) was the best model fit for the survival data. Generalised Gamma (AIC=2378, BIC=2451) and Weibull (AIC=2382, BIC=2452) models were respectively the next alternatives among the nine (9)

**Corresponding author: Email: sbosson-amedenu@st.umat.edu.gh, senyefia@yahoo.com;*

accelerated failure time (AFT) models considered in our study. Results from the three best fitted AFT models showed that covariates such as Age at diagnosis, Progesterone receptor, Molecular Subtype, Grade, Stage, Metastasis, number of Lymph nodes involved and genetic status were the significant factors that have an effect on the survival time of BC patients in Ghana (P<0.05). The Area under the ROC curve (AUC=0.945) shows an outstanding performance of the Gompertz AFT model to discriminate the true disease status of patients.

Conclusion: Although the Cox proportional hazard model has seen wide usage and remains a robust approach in survival analyses for the past four decades; its proportional hazards assumption is most often violated by some covariates in medical research. Under such violations, AFT models are a strong alternative.

Keywords: Accelerated Failure-time models; akaike information criterion; bayesian information criteria; cox proportional hazard model; breast cancer; ROC curve.

1. INTRODUCTION

Breast cancer has been ranked the first cause of cancer death (14.3% of the total) in low-and-middle income countries; the second cause of death (15.4%) in developed countries; and the fifth cause of death worldwide [1]. Breast cancer in young women is relatively rare compared to breast cancer occurring in older women. Younger women diagnosed with breast cancer also tend to have a more aggressive biology and consequently a poorer prognosis than older women. There is a need for more research in the area to optimize clinical outcomes [2]. The mean age at diagnosis of breast cancer in Africa has been found to be at a relatively young age of 54 years and occurs a decade earlier than patients from high income countries. Breast cancer patients in underdeveloped countries comparatively present high staging (III and IV) of the disease at diagnosis [3]. GLOBOCAN report of 2016 shows that about 80% of Ghana's breast cancer cases are in the lower age-groups and often associated with lower survival, relatively poorer prognosis and higher mortalities although with relatively lower morbidities. The mean age of incidence of cancer in Ghana has been found to be 48 years. The peak age of incidence of BC among Africans occurs in the premenopausal while it occurs at the postmenopausal period among non-Africans [4,5]. About 82% of Ghanaian BC patients have been found to be diagnosed with triple negative molecular subtype versus 26% of African Americans and 16% of white Americans [6-8].

When the Cox Proportional Hazard (CPH) assumption is not tenable, accelerated failure-time (AFT) model is an alternative to the CPH model. However, AFT modelling relies heavily on fitting as many parametric models to find a more appropriate form of the parametric family that best models the covariates. This search for appropriate parametric models is limited to the availability of software for fitting the parametric models. Explained in another way, unlike in AFT models, it is a difficult task to make different distributional assumptions for PH model [9,10]. It has been established that the family of Gompertz distributions is not only a collection of Proportional hazard (PH) families but also a collection of AFT families [11,12]. [11,12] demonstrated that Gompertz distribution fits in very well into accelerated failure time modelling especially in modelling mortality [13] found that exponenential and Gompertz AFT models were the best model fit after comparing six AFT models on the basis of AIC and Cox-Snell Residuals. In their work, six AFT were compared, namely; exponential, loglogistic, lognormal, Gamma, Weibull and Gompertz. [14] demonstrated Gompertz regression parameterized as accelerated failure time model. They argued that Gompertz AFT models are appropriate for treatment effects modelling such as that of BC.

In this study we considered a novel comparison of nine [10] AFT models which has never been elucidated in our search in literature. The comparison of a larger number of AFT models makes it feasible in fitting as many parametric models to find the most robust AFT model that best fit the survival data.

2. MATERIALS AND METHODS

In this retrospective study, 558 patients with Breast cancer tumor characteristics and demographics have been studied: the data span the year 2010 through 2015 on diagnosed BC patients at Korle Bu Teaching Hospital of Ghana. The data was collected from patients' information in the archives of the hospital, which contained

their survival status and treatment information over a period of time. Only patients with complete information over the five year period of interest to this study were selected for the study. Patients who did not experience the event of interest (death to breast cancer) by the end of the study were right-censored. The effects of demographic variables such as Age at diagnosis , Recurrent status, HER2 status, ER status, PR status, Molecular Subtype, Grade, Disease Stage(I-II-III-IV), Distance metastases, Number of Lymph Nodes involved, Menopause status at diagnosis, Ethnicity, Hospitalization status, and Hereditary or genetic status (BRCA 1 or 2) were evaluated and compared among various AFT models. The data consisted of BC patients in the age range of 13 to 97 years. The dataset was partitioned into training (holding 70%) and validation set (30%).

To compare different survival models, Akaike Information Criterion (AIC), Bayesian Information Criteria (BIC) were applied. Q-Q plot was used to assess the appropriateness of AFT model (Shapiro-Wilk normality test: W = 0.99588, p-value = 0.1516). The Q-Q plot in approximates well to a straight line from the origin with most of the points lying on the line, the data is normally distributed (for two age groups of <=50years and >50years); an indication that the AFT model may provide an appropriate model. AIC, and BIC are used to measure the goodness of models' fitness. The smaller the AIC and BIC the better the model fit. AIC and BIC for the models used in this study has been calculated according to the following formula:

$$AIC = -2\log(\text{maximum likelihood}) + 2p$$

$$BIC = -2\log(\text{maximum likelihood}) + plogn$$

Where p is the number of model parameters example, p = 1 for the exponential model, p = 2 for the Weibull
model and p = 3 for the generalized gamma model.

TNM (7th edition) was employed to determine the stage of the disease (Edge et al., 2010).

R software was used for all analyses and the significance level was set at 5%.

BC tumors had been separated by expression of hormone receptor status which defines distinct biological phenotypes used to guide BC prognosis and treatment. Consider the classification in table 1 as explained by [15].

The Receiver Operation Characteristic (ROC) curve has emerged as the method of choice for assessing performance of prognostic and predictive models, notably for cancer subtype classification, disease diagnosis and prognosis, patient risk group stratification, and prediction of response to treatment [16].The receiver Operation characteristic (ROC) curve is presently a standard and popular goodness of fit test for comparing predictive models and assessing how accurate they predict clinical risk by stratifying patients into higher or lower risk categories of clinical importance [17]. The area under the ROC curve (AUC) provides a recommended index of accuracy of the measure of discrimination of these two groups which is used for the purpose of statistical analysis and for comparing at least two classification models [18]. The immense popularity of ROC curves in contemporary medical research stems partly from its usefulness as a means to evaluating predictive ability of covariates and ease of implementation and interpretation; as well as invariance in strictly increasing transformations of the feature and shifts in prevalence. ROC curve has been extensively utilised as a performance measure for binary classifiers, particularly in medical diagnostic situations [19]. When ROC curves are used in comparing regression models, the models are dichotomised into various thresholds to compute the resulting sensitivity and specificity.

Table 1. Molecular Classification of Breast Cancer

Receptor	Luminal A	Luminal B	Triple Negative	HER2 enriched(overexpression)
ER	Positive	Positive	Negative	Negative
PR	Positive	Positive	Negative	Negative
HER2	Negative	Positive or Negative	Negative	Positive

Table 2. Interpretation of ROC Curve

c-statistic	Interpretation
AUC=0.5	No discrimination
0.6≥AUC>0.5	Poor discrimination
0.7≥AUC>0.6	Acceptable discrimination
0.8≥AUC>0.7	Excellent discrimination
AUC>0.9	Outstanding discrimination

The area under the ROC curve (AUC) is a popular summary index of an ROC curve (sensitivity and specificity). To compare the area under the curve for any two different models,

$$Z = \frac{A_1 - A_2}{\sqrt{(se(A_1))^2 + (se(A_2)^2 - 2 \cdot r \cdot se(A_1) \cdot se(A_2)}}$$

where z is the standard normal variate and r represents the correlation induced between the two areas under the curve due to application of the two models on the same sample.

In general, the rule of thumb for interpreting AUC value is represented in Table 2 [20].

3. RESULTS

It has been established that BC affects mostly young pre-menopausal women in Ghana [21]. The average age of menopause among Ghanaian women have been found to be 48.05 $\pm$ 3.62 years [22]. This sets the basis for comparing survivorship or hazard among patients based on young premenopausal women (≤ 50 years) and postmenopausal women ages (> 50 years). BC patients in this study were mostly treated with a combination of surgery (mastectomy), radiation therapy, and chemotherapy. Although Hormonal therapies and HER2-targeted therapies have been widely efficient in helping slow down or even halt the growth of BC cells triple-negative BC does not respond to these hormonal therapy medicines and hence has the poorest prognosis [23]. Women in younger ages (≤ 50 years) were the most diagnosed (54.7%) of breast cancer (BC) at the start of the follow-up; compared to their counterparts in older ages (> 50 years, 45.3%). The mean age at diagnosis is 50 years with standard deviation 14.3. Concerning staging of tumor, about 38% and 11% of the women were respectively at degrees of III and IV. Another 18% and 28% of the women, at the day of diagnosis were at the stages of I and II respectively. Regarding Grading of the disease among the women diagnosed of BC, about 32% and 22% were well differentiated (Grade 1) and moderately differentiated (Grade 2) respectively. However, about 46% of were poorly differentiated (Grade 3). In the cause of follow-up, approximately 17% of the patients were hospitalized at some point and about 8% experiencing recurrence of the disease. Of the women diagnosed of BC, about 65% were postmenopausal. Among the molecular subtypes of BC, Triple negative (Basal type) had the most incidence (43%) followed by Luminal A (about 32%), Luminal B (22%) and HER2+ (about 3%). With regards to Metastases, about 281 (50.4%) of the tumours metastasized where 261(46.8%) did not metastasize and 16 (2.9%) could not be measured. About 15 % had no lymph node involvement, 38% cancer spread to 1 to 3 lymph nodes, 28% cancer spread to 4 to 9 lymph nodes and about 19% has spread to 10 or more lymph nodes. Inherited mutation in BRCA 1 and BRACA 2 accounted for 43% of the causes.

Young women (≤ 50 years) were the most diagnosed (54.7%); which is similar to the 54.8% found by [24]. The average age of women with breast cancer in this study was $50.0 \pm 14.3\ years$ which is fits the expected young profile of breast cancer patients in the region. This is comparable with prior studies in Cape Coast (49.9 years); (14), Kumasi (49.1 years); [25] and in Africa (50.2 years); [26]. Triple negative (Basal type) was the most (43%) prevalent molecular subtype while there less prevalent subtype was HER2+; this finding is supported by [24-29]. About 49% BC patients were found to be at late stage (III and IV), consistent with that of [30]. However, our study reported 46% of BC patients with Grade 3 tumors which is about 4% below the 50% reported by [30]. The 50.4% breast cancer metastasis found by our study fits well within the range (39.8% - 55.3%) of three commonest sites of distant metastases found by [31]. Our study found that 32% of the patients had Luminal A breast cancer which compares with the 32.8% found by [31]. The 85% BC lymph node metastasis found by our study is slightly higher than the 80% reported

by [32]. Although BRCA 1 or 2 accounts for 5% to 10% of BC disease in the general population, it has however been found to increase in the young age bracket of patients [33]. The 43% BRCA 1 or 2 prevalence found in our study data was the same as found in a similar study in Ghana [33]. The increased levels of Triple negative (or Basal type) molecular subtype of BC among sub-Saharan Africans have been associated with hereditary causes due to BRCA 1 or 2 [33]. We found that BRCA 1 or 2 was about 67% prevalent in the younger ages (≤ 50 years) of patients.

The proportional hazard (PH) assumption is violated by a covariate that shows a significant relationship between residuals and time by Goodness of Fit Test. It is therefore evident from Table 3 that Grade and Progesterone receptor (PR) are time-variant and so violate the PH assumption. Under such situations using the Cox PH model will produce erroneous estimates.

Table 3. Goodness of Fit Test for Proportional Hazard Assumption

Covariate	chisq	df	p-value
Age	0.00169	1	0.967
Recurrent	0.32053	1	0.571
HER2	0.00176	1	0.966
PR	5.16717	1	**0.023**
ER	3.77155	1	0.052
MSubtype	2.00089	1	0.157
Grade	5.63717	1	**0.018**
Stage	2.35698	1	0.125
Metastasis	0.28847	1	0.591
LymphNode	0.57144	1	0.45
Menopause	0.59849	1	0.439
Ethnicity	0.77946	1	0.377
Hospitalization	2.07502	1	0.15
Genetics	2.72288	1	0.099
GLOBAL	22.62211	14	0.067

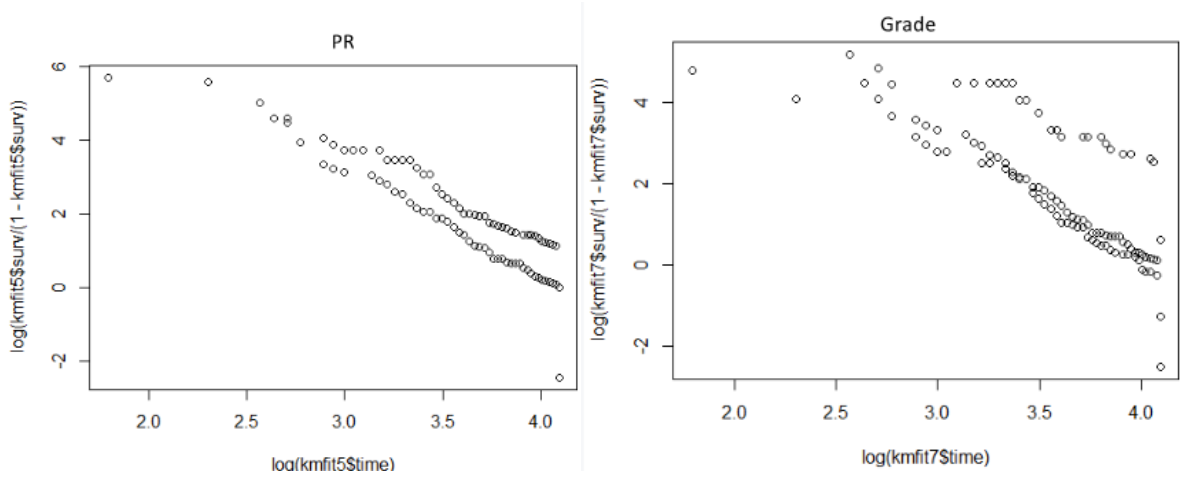

Fig. 1. Graphical Test for Proportional Hazard Assumption

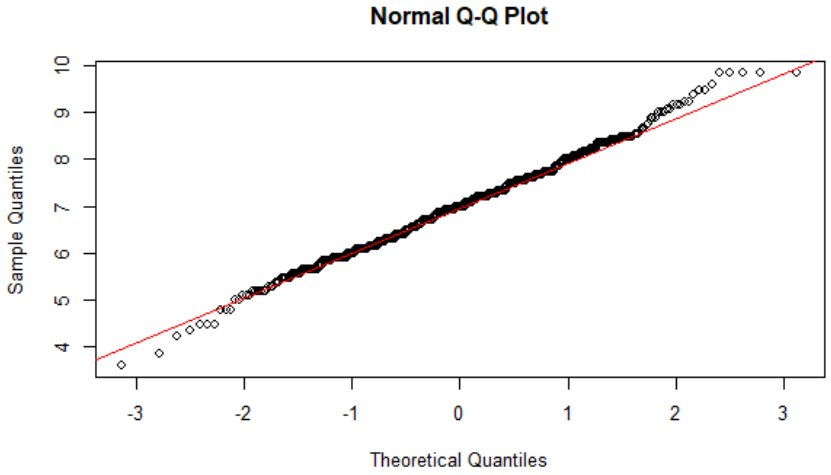

Fig. 2. Q-Q Plot Test for Appropriateness of AFT Model

Table 4. Goodness of Fit Test for AFT Candidate Models

AFT MODEL	AIC	BIC	Loglik
AFTloglogis	2436	2505	-1202
AFTExpo	2895	2959	-1432
AFTlognor	2455	2524	-1211
AFTGompertz	**2322**	**2391**	**-1145**
AFTgamma	2426	2495	-1172
AFTWeibull	2382	2452	-1175
AFTGengama	2378	2451	-1172
AFTLogGAuss	2455	2524	-1211
AFTreleigh	2567	2632	-1269

Table5. Gompertz AFT Model Output

Predictors	Estimates	CI	p
Age	1	0.99-1	**0.037**
Recurrent	0.95	0.87-1.03	0.233
HER2	0.93	0.74-1.17	0.546
ER	1.12	0.72-1.75	0.615
PR	1.55	1.16-2.06	**0.003**
MSubtype	0.5	0.27-0.9	**0.022**
Grade	3.59	1.18-10.87	**0.024**
Stage	0.91	0.88-0.95	**<0.001**
Metastasis	0.73	0.67-0.86	**<0.001**
LymphNode	0.94	0.91-0.98	**0.001**
Menopause	1.03	0.94-1.12	0.564
Ethnicity	0.99	0.96-1.02	0.602
Hospitalization	0.98	0.91-1.06	0.628
BRCA 1or2	4.16	1.26-13.7	**0.019**

Fig. 1 shows that the hazards cross for Progesterone receptor (PR) and Grade covariates which is consistent with outcome of the Goodness of Fit test; suggesting a violation to the PHA.

The Q-Q plot was used to check the Accelerated Failure Time (AFT) assumption. The Q-Q plot in Figure 2 approximates well to a straight line from the origin with most of the points lying on the line; the data is normally distributed (for two age groups of <=50years and >50years); an indication that the AFT model may provide an appropriate model.

To find the best fitting model the Akaike's Information Criterion (AIC) and Bayesian Information Criteria (BIC) were used. Out of the nine [10] AFT models compared, the Gompertz AFT model was found to be the best fitting model to the study data, with the least AIC and BIC values. Apart from the overall best performing Gompertz AFT accelerated failure-time model, Generalised gamma and Weibull AFT models were respectively the next alternatives.

Results of the Gompertz AFT model revealed that Age at diagnosis, Progesterone receptor (PR), Molecular Subtype, Grade, Stage Metastasis, number of Lymph node involved and BRCA1 or 2 statuses were the significant factors that have an effect on the survival time of breast cancer patients in Ghana. Our findings are corroborated by the studies [34].

Moreover, covariates of recurrent status, Human epidermal receptor2 (HER2) status, Oestrogen receptor status (ER), Menopause status, Ethnicity background, and hospitalization status did not have any significant effect on patients' survival in any of the studied models.

The estimated log time to death to BC with younger patient in comparison with older patients was 1.0. The accelerated factor is exp [1] which is 2.718; an indication that the younger population accelerates the time to die of BC by a factor of 2.718. This also means that the younger population has a shorter time (by a factor of 2.718 to die by BC). The estimated hazard ratio (HR) comparing death by breast cancer by young and old is exp (-1.0) which is 0.368, means younger patients are 0.368 higher risks to die from breast cancer than those at higher ages.

Similarly, PR negative status of BC patients accelerates the time to die of BC by a factor of 4.711; and have 0.212 higher risks to die from BC than those who are PR positive. Molecular Subtype, Grade and Stage at diagnosis accelerate the time to die of BC by factors of 1.648, 36.324 and 2.484 respectively. This means that higher Grade of BC at diagnosis accelerate the time to die of BC by a factor of 36.324; whereas higher stage of BC at diagnosis accelerated the time to die of the disease by 2.484.Again, BC patients with Triple negative (TN) molecular subtype have higher accelerated risk of dying by a factor of 1.648 . The high accelerated factor of 36.324 for high grade BC makes sense since high grade tumors often grow and spread faster and more likely to be invasive in nature [35].

The estimated log time to death to BC patients with tumor that has spread to 10 or more lymph nodes is 0.94. In effect, a BC patient with tumor that has spread to 10 or more lymph nodes accelerates the time to die of BC by a factor of 2.460. Finally, patients whose cause of BC was due to genetic or heredity causes (BRCA 1 or 2) accelerate the time to die from BC by a factor of 64.072. Our high accelerated factor estimates for *BRCA1 and* BRCA2 are meaningful since for example; [34] has established that about 73% of *BRCA1* carriers have triple-negative breast cancer; whereas most (72%) *BRCA2* carriers have hormone receptor-positive tumors which are known for their poor prognosis and are significantly associated with metastasis. Again our finding is reasonable since triple negative BC is the most prevalent molecular subtype among particularly sub-Saharan African countries and has been found to be the major risk factor for BRCA 1 or 2 [33].

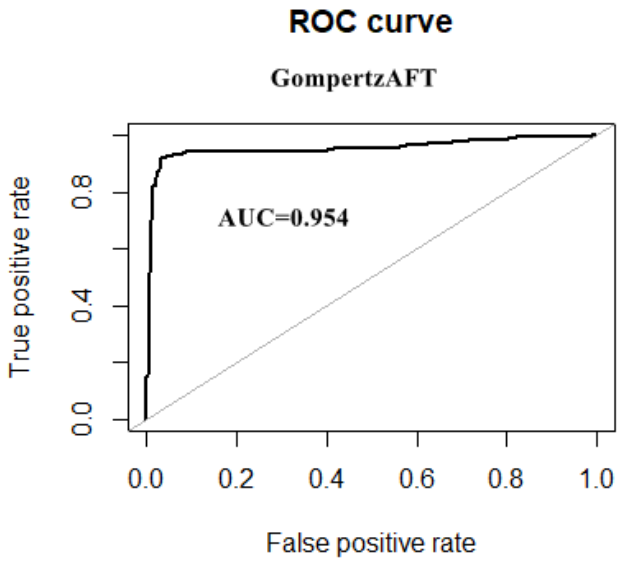

Fig. 3. Predictive Efficiency of the Gompertz AFT Model

Table 6. Significant Covariates Predicted by each Model

	Weibull AFT	**Gen Gamma AFT**	**Gompertz AFT**
Predictors	P-Value	P-Value	P-value
Age	**0.02**	**0.003**	**0.037**
Recurrent	**0.165**	**0.339**	**0.233**
HER2	0.549	0.782	0.546
ER	0.653	0.308	0.615
PR	**0.005**	**0.022**	**0.003**
MSubtype	**0.02**	**0.024**	**0.022**
Grade	**0.024**	**0.026**	**0.024**
Stage	**<0.001**	**<0.001**	**<0.001**
Metastasis	**<0.001**	**<0.001**	**<0.001**
LymphNode	**0.001**	**0.002**	**0.001**
Menopause	0.486	0.531	0.564
Ethnicity	0.632	0.777	0.602
Hospitalization	0.589	0.805	0.628
Genetics	**0.016**	**0.022**	**0.019**

The dataset was partitioned into training (holding 70%) and validation set (30%).The Reciever Operation Characteristic (ROC) curve shows the predictive capacity of the Gompertz Accelerated failure model is very high, evidenced by the high area under the curve (AUC=0.945). The AUC=0.945 shows an outstanding performance of the Gompertz AFT model to discriminate the true disease status of patients.

The table 6 results show that the three best performing AFT models considered in this study consistently predicted the same significant covariates of Age at diagnosis, Progesterone receptor (PR), Molecular Subtype, Grade, Stage Metastasis, number of Lymph node involved and BRCA1 or 2 statuses.

4. PARAMETERIZATION OF GOMPERTZ REGRESSION MODEL AS ACCELERATED FAILURE- TIME MODEL

The *Gompertz* distribution is special in that it can be fit into both the AFT and the PH framework [14, 36,37].

Consider the two parameter Gompertz survival function given by

$$S(t) = \exp\left(-\frac{e^{\ell}}{\varphi}(e^{\ell}-1)\right) \quad (1)$$

Where the corresponding hazard function is given by

$$\psi(t) = \exp(\ell)\exp(\varphi t) = \exp(\ell + \varphi t) \quad (2)$$

The product of (1) and (2) gives the probability density function

$$f(t) = \exp\left(-\frac{e^{\ell}}{\varphi}(e^{\ell} - 1) + \ell + \varphi t \right)$$

The first moment is modelled by the parameter ℓ , where covariates enter the model exponentially as

$$\ell_i = \exp(\omega X_i)$$

The AFT survival model for a group of patients with covariate $x_1, x_2, x_3, .., x_p$ is mathematically represented by

$$S(t|x) = S_0\left(\frac{t}{\eta(x)} \right) \quad (3)$$

$$S(t|x) = tS_0\left(e^{\mu x}\right), \quad t > 0,$$

Where time, t is a factor of the accelerated factor $e^{\mu x}$

Where $\eta(x) = \exp(\mu_1 x_1 + \mu_2 x_2 + ... + \mu_p x_p)$

Given that the lifetime T of a patient with covariates X has a survivor function given by (1)

We can derive the distribution for Y = log (T) as:

$$P(Y \geq y) = P(\log(T) \geq y) = P(T \geq e^{y}) = S(e^{y}; x) = S_0\left(e^{y+\beta x}\right) \quad (4)$$

It can be deduced from (3) that β X is a location parameter in the family of distributions of Y , and has the log-linear model for:

$$Y = \log(T) = \beta x + \varepsilon = \beta_0 + \beta_1 x_1 + ... + \beta_p x_p + \varepsilon \quad (5)$$

Where exp (ε) has the distribution S_0 and ε serves as the "error term".

The Gompertz distribution has exponentially growing hazard function, and can be parameterized as:

$$\psi_a\left(t|(\ell,\varphi)\right) = \ell e^{\varphi t}, \quad \ell > 0, \varphi \geq 0; t > 0. \quad (6)$$

Where proportionality constant λ ,

We now transform the proportionality constant to AFT

With the *transformation* $(\ell, \gamma) \rightarrow \left(\frac{\ell}{\varphi}, \frac{1}{\varphi} \right)$

Leading to the Gompertz AFT model

$$\psi_b\left(t|(\ell,\varphi)\right) = \frac{\ell}{\varphi}\exp\left(\frac{t}{\varphi}\right), \ell, \varphi > 0; t > 0, \quad (7)$$

and now ℓ is the "PH parameter" (as before) and φ is the "AFT parameter

The transition from ψ_a to ψ_b implies that the rate in the canonical form must be strictly positive.

Hence the canonical parameterisation form of the survival function for (7) becomes:

$$S\left(t;(\ell,\varphi)\right) = \exp\left\{ -\ell\left(e^{\frac{t}{\varphi}} - 1 \right) \right\}, t > 0. \quad (8)$$

The conclusion of all this is that the AFT Gompertz model is suitable in situations where the intensity of an event is clearly increasing with time [36,37].

5. DISCUSSION

Breast cancer studies have often neglected the fact that breast cancer covariates could be time-dependent in which case the Cox proportional hazard assumption is violated. Neglecting this violation leads to erroneous and misleading estimates of probabilities associated with survival or hazard [38]. Under such situations where the proportional hazard assumption is violated, accelerated failure-time modelling can be considered. A recent review of cancer related works employing Cox PH model in the past decade revealed that 81% of publications did not

account for the proportional hazard assumption [39]. Very little attention has been paid to violation to Cox PH assumption; as 95% of all studies using the Cox PH model without testing checking if the assumption is violated or not, leading to biased, unreliable and erroneous conclusions [40]. To this end, AFT models such as Loglogistic, Exponential, Lognormal, Gompertz, Gamma, Weibull, Generalized Gamma, Gaussian, LogGaussian , extreme and Raleigh can be better choices in such circumstance. Expressed in another way, AFT models are flexible such that they allow one to make different distributional assumptions that best models the covariates hence suitable alternative to the Cox PH model. In this study, the results of accelerated failure-time models were compared to analyze the survival of patients with BC in Ghana.

To compare these models, AIC, BIC were used. Among accelerated failure-time models, Gompertz, Extreme and Gaussian models were more efficient than other AFT models and hence the best alternative to the Cox PH model. Results from the three best fitted AFT models showed that covariates such as Age at diagnosis, Progesterone receptor (PR), Molecular Subtype, Grade, Stage, Metastasis, number of Lymph node involved and BRCA1 or 2 statuses were the significant factors that have an effect on the survival time of breast cancer patients in Ghana (P<0.05).These results are consistent with the results of many studies in this field [41-44]. Moreover, covariates of recurrent status, Human epidermal receptor2 (HER2) status, Oestrogen receptor status (ER), Menopause status, Ethnicity background, and hospitalization status did not have any significant effect on patients' survival in any of the studied models. This issue is consistent with most studies conducted on patients with breast cancer [13,45].

6. CONCLUSION

Although Cox PH Model remains for the last four decades the most robust in comparison with parametric and nonparametric models, AFT models which do not assume the constant hazards in the survival data provide a more valid, reliable and applicable results in the event that the PH assumption is violated. Based on our results, the Gompertz (AIC=2322, BIC=2391) as the best performing AFT model among all the AFT models considered. This finding is supported by the works of [13,37]. Generalised Gamma (AIC=2378, BIC=2451) and Weibull (AIC=2382, BIC=2452) were the two other alternative models; which are also corroborated by the studies [40,46].

CONSENT

As per international standard or university standard, Participants' written consent has been collected and preserved by the author(s).

ETHICAL CONSIDERATIONS

As per international standard or university standard written ethical approval has been collected and preserved by the author(s).

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Ferlay J, Soerjomataram I, Dikshit R. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012, International Journal of Cancer, Int. J. Cancer. 2014;136:E359–E386 .
2. Ademuyiwa FO, Cyr A, Ivanovich J, Thomas MA. Managing breast cancer in younger women: challenges and solutions. Breast cancer (Dove Medical Press). 2015;8:1–12. Available: https://doi.org/10.2147/BCTT.S68848.
3. Walker ARP, Adam FI, Walker BF. Breast cancer in black African women: a changing situation. Journal of the Royal Society for the Promotion of Health. 2004;124(2):81-85. DOI:10.1177/146642400412400212
4. Rambau PF, Chalya Pl, Manyama MM and Jackson KJ. Pathological features of Breast Cancer seen in Northwestern Tanzania: a nine years retrospective study. BMC Re-search Notes. 2011;4:214:1-6.
5. Adesunkanmi ARK, Lawal OO ,Adelusola KA,and Durosimi MA. The severity, out-come and challenges of breast cancer in Nigeria,Breast. 2006;15 (3);399–409.
6. Seshie B, Adu-Aryee NA, Dedey F, et al: A retrospective analysis of breast cancer subtype based on ER/PR and HER2 status in Ghanaian patients at the Korle Bu Teaching Hospital, Ghana. BMC Clin Pathol. 2015;15:14.

7. Boyle P: Triple-negative breast cancer: Epidemiological considerations and recommendations. Ann Oncol 23:vi7- vi12, 2012, (suppl 6)
8. Stark A, Kleer CG, Martin I, et al: African ancestry and higher prevalence of triple-negative breast cancer: Findings from an international study. Cancer. 2010;116:4926-4932.
9. Lee ET, Oscar TG. Survival Analysis in Public Health Research, Annu. Rev. Public Health. Copyright c 1997 by Annual Reviews Inc. All rights reserved. 1997;18:105–34.
10. Prabhash K, Patil VM, Noronha V, Joshi A, Ramaswamy A, Dhumal S, Juvekar S. Neoadjuvant chemotherapy in geriatric head and neck cancers. Head & Neck, 2017;39(5):886–892. DOI:10.1002/hed.24694
11. Broström A et al., ().Academic engagement and commercialisation: A review of the literature on university–industry relations,Research Policy. 2013;42(2): 423-442. ISSN 0048-7333. Available:https://doi.org/10.1016/j.respol.2012.09.007.
12. Broström G. Event History Analysis with R ISBN 9781439831649 Published May 10, 2012 by CRC Press 236 Pages 75 B/W Illustrations
13. Mahmoodi M, Hooseini M, Zare A, et al.. A Comparison between Accelerated Failure-time and Cox Pro-portional Hazard Models in Analyzing the Survival of Gastric Cancer Patients, Iran J Public Health. 2015;44(8):1095-1102
14. Andersson F, Orsini N. Gompertz regression parameterized as accelerated failure time model, 2017 Nordic and Baltic Stata Meeting; 2017.
15. Khinaifis F.A ,Molecular classification of breast cancer: A retrospective cohort study, Annals of Medicine and Surgery. 2020;49: 44-48,ISSN 2049-0801
16. Berrar D, Flach P. Caveats and pitfalls of ROC analysis in clinical microarray research (and how to avoid them). Brief Bioinform. 2012;13(1):83-97. DOI: 10.1093/bib/bbr008. Epub 2011 Mar 21. PMID: 21422066.
17. Cook NR. Use and Misuse of the Receiver Operating Characteristic Curve in Risk Prediction. Circulation, 2007;115(7): 928–935. DOI:10.1161/circulationaha.106.67
18. DeLong ER, DeLong DM, Clarke-Pearson DL. (). Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. Biometrics, 1988;44(3):837. DOI:10.2307/2531595
19. Pepe MS. The statistical evaluation of medical tests for classification and prediction. Oxford university press, Oxford; 2003.
20. Metz CE. Basic principles of ROC analysis. Semin Nucl Med. 1978;8(4):283-98. DOI: 10.1016/s0001-2998(78)80014-2. PMID: 112681.
21. Ohene-Yeboah M, Adjei E. Breast cancer in Kumasi, Ghana. Ghana Medical Journal. 2012;46(1):8–13.
22. Kwawukume EY, Ghosh TS, Wilson JB. Menopausal age of Ghanaian women. Int J Gynaecol Obstet. 1993;40(2):151-5. DOI: 10.1016/0020-7292(93)90376-8. PMID: 8094686
23. Ganz AP. Improving Outcomes for Breast Cancer Survivors, Perspectives on Research Challenges Advances in Experimental Medicine and Biology 862, Breast Cancer Research Foundation; 2015.
24. Okifo FO, Tuoyire AD, et al. Breast cancer treatment and outcomes at Cape Coast Teaching Hospital, Ghana, Ghana Med J. 2021;55(3):190-197 Available: http://dx.doi.org/10.4314/gmj.v55i3.3
25. Ohene-Yeboah M, Adjei E: Breast cancer in Kumasi, Ghana. Ghana Med J. 2012;46:8-13.
26. Adeloye D, Sowunmi OY, Jacobs W, David RA, Adeosun A, Amuta AO. Estimating the incidence of breast cancer in Africa: a systematic review and meta-analysis. J Glob Health. 2018;8(1):1-21. DOI:10.7189/jogh.08.010419
27. Seshie B, Adu-Aryee NA, Dedey F, et al: A retrospective analysis of breast cancer subtype based on ER/PR and HER2 status in Ghanaian patients at the Korle Bu Teaching Hospital, Ghana. BMC Clin Pathol. 2015;15:14.
28. Boyle P: Triple-negative breast cancer: Epidemiological considerations and recommendations. Ann Oncol 23:vi7- vi12, 2012:(suppl 6).
29. Stark A, Kleer CG, Martin I, et al: African ancestry and higher prevalence of triple-negative breast cancer: Findings from an

international study. Cancer 116:4926-4932, 2010

30. Thomas AS, Kidwell MK, Oppong JK et al. Breast Cancer in Ghana: Demonstrating the Need for Population-Based Cancer Registries in Low- and Middle-Income Countries, Journal of Global Oncology. 2017;3(6).
31. Mensah YB et al. A review of computed tomography patterns of metastatic breast cancer patients undergoing treatment at a private oncology centre in Ghana. Pan African Medical Journal. 2021; 38(50).
DOI:10.11604/pamj.2021.38.50.21945
32. Quayson S.E., E.K. Wiredu DN, Adjei and J.T. Anim (2014). Breast cancer in Accra, Ghana Journal of Medical and Biomedical Sciences. 2014; 3(3):21-26
33. Amankwaa-Frempong E, Yeboah FA, Nguah SB, Newman LA. Breast Cancer Genetic Testing Among African Patients With Breast Cancer: Deoxyribonucleic Acid Extraction From Tumor Tissue and International Multidisciplinary Partnerships. JAMA surgery. 2017;152(8): 800–801.
Avaialble:https://doi.org/10.1001/jamasurg.2017.1090
34. Song Y, Barry WT, Seah DS, Tung NM, Garber JE, Lin NU. Patterns of recurrence and metastasis in BRCA1/BRCA2-associated breast cancers. Cancer. 2020;126(2):271–280.
Available:https://doi.org/10.1002/cncr.32540
35. Duffy MJ, Harbeck N, Nap M, Molina R, Nicolini A, Senkus E, Cardoso F. Clinical use of biomarkers in breast cancer: Updated guidelines from the European Group on Tumor Markers (EGTM). Eur J Cancer. 2017;75:284-298.
36. Göran Broström 2021. Vignette. Gompertz Distribution eha , vignettes/gompertz.Rmd
Available:https://cran.rproject.org/web/packages/eha/vignettes/gompertz.html
37. Göran Broström and Sören Edvinsson. A parametric model for old age mortality in mediation analysis, Paper presented at the XXVII IUSSP International Population Conference, 26–31 August 2013, Busan, Republic of Korea; 2013.
38. Meier-Hirmer C. Schumacher M. Multi-state model for studying an intermediate event using timedependent covariates: application to breast cancer. BMC Medical Research Methodology. 2013;13:80.
39. Rulli E, Ghilotti F, Biagioli E, Porcu L, Marabese M, D'Incalci M, Bellocco R, Torri V. Assessment of proportional hazard assumption in aggregate data: a systematic review on statistical methodology in clinical trials using time-to-event endpoint. Br J Cancer. 2018; 119(12):1456-1463.
DOI: 10.1038/s41416-018-0302-8. Epub 2018 Nov 13. PMID: 30420618; PMCID: PMC6288087.
40. Abadi A, Amanpour F, Bajdik C, Yavari P. Breast cancer survival analysis: Applying the generalized gamma distribution under different conditions of the proportional hazards and accelerated failure time assumptions. Int J Prev Med. 2012;3: 644-51.
41. Karimi A, Delpisheh A, Sayehmiri K. Application of accelerated failure time models for breast cancer patients' survival in Kurdistan Province of Iran. J Cancer Res Ther. 2016;12(3): 1184-1188.
DOI: 10.4103/0973-1482.168966. PMID: 28054533.
42. Sharma MK, Endalkachew Abebe. Determinants of Survival Time of Women with Breast Cancer, Archives of Oncology and Cancer Therapy ISSN: 2638-5074. 2019;2(2):26-39,
43. Perera M, Tsokos C. A Statistical Model with Non-Linear Effects and Non-Proportional Hazards for Breast Cancer Survival Analysis, Advances in Breast Cancer Research, 2018;7(1):65-89.
44. Natarajan L, Pu M, Barbara A. Parker, Cynthia A. Time-Varying Effects of Prognostic Factors Associated With Disease-Free Survival in Breast Cancer, American Journal of Epidemiology. 2009;169(12):1436-1470.
45. Iraji Z, Jafari Koshki T, Dolatkhah R, Asghari Jafarabadi M. Parametric survival model to identify the predictors of breast cancer mortality: An accelerated failure time approach. Journal of research in medical sciences : the official journal of Isfahan University of Medical Sciences. 2020;25: 38.
Available:https://doi.org/10.4103/jrms.JRMS_743_19
46. Bakhshi E, Ali Akbari Khoei R, Azarkeivan A, Kooshesh M, Biglarian A (2017).

Survival analysis of thalassemia major patients using Cox, Gompertz proportional hazard and Weibull accelerated failure time models. Med J Islam Repub Iran. 2017;31: 97.
Available:https://doi.org/10.14196/mjiri.31.97.

© 2022 Bosson-Amedenu et al.; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Peer-review history:
The peer review history for this paper can be accessed here:
https://www.sdiarticle5.com/review-history/83754