

STAT 306 Project (Group 10)

Nicholas Tam (45695970), Ivy Cheung (97777726), Kaichi Nakajima (74175712), Kevin Liu (94200474)

Introduction

Breast invasive ductal carcinoma (IDC) is the most common type of breast cancer, with about 80% of all forms of breast cancer being IDC, according to the American Cancer Society (DePolo, 2024). There are numerous nonsurgical treatments of IDC, such as radiotherapy, chemotherapy, and hormone therapy (Wright, 2023), and each of their effectiveness is partly determined by the patient's condition, such as age and tumor stage. For instance, due to interactions between other treatments or conditions as a consequence of aging (e.g. Diabetes, liver disease, metabolism), more optimal doses of chemotherapy are generally discouraged for older patients due to potentially toxic side effects, implying a smaller difference in survival between older patients with or without chemotherapy (Given, Given, 2008). However, it is unclear how combinations of treatments can interact in a model to predict a patient's survival until death.

For our research project, we have selected a dataset of approximately 1900 primary breast cancer samples, obtained from the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) database through cBioPortal for Cancer Genomics.

Our project question is: **"How do radiotherapy, chemotherapy, and hormone therapy influence the length of time a patient with IDC will survive, given control variables age, surgery type, tumor stage, and their present survival status?"**

Our analysis will involve the inference of covariates within linear models, as we seek to determine the interaction of cancer therapies on allowing patients to survive longer from IDC.

- Variables from columns 32 to 693 consist of genetic attributes containing m-RNA levels z-score for 331 genes, and mutation for 175 genes; they have been omitted due to being difficult to interpret.
- Due to the distribution of "cancer_type_detailed" categories and for ease of computation, we will be filtering the dataset for IDC patients that are either alive or dead from disease, as IDC consists of the majority of the dataset, and the patients that have died from other causes are irrelevant to the project question and not specific enough (e.g. Accident, non-cancer diseases). As such, we have narrowed down the data to the following attributes:

| Variable | Definition | Unit | Categories |
|------------------------|---|---------|---|
| age_at_diagnosis | Age of the patient at diagnosis time | Years | |
| type_of_breast_surgery | Breast cancer surgery type | N/A | "MASTECTOMY", "BREAST CONSERVING" |
| cancer_type_detailed | Detailed breast cancer types | N/A | "Breast Invasive Ductal Carcinoma", "Breast Mixed Ductal and Lobular Carcinoma", "Breast Invasive Lobular Carcinoma", "Breast Invasive Mixed Mucinous Carcinoma", "Metaplastic Breast Cancer" |
| chemotherapy | Boolean on whether or not patient had chemotherapy as a treatment | Boolean | |
| hormone_therapy | Whether or not the patient had hormonal therapy as a treatment | Boolean | |

| | | | |
|-------------------------|--|---------|---|
| overall_survival_months | Duration from the time of the intervention to death | Months | |
| death_from_cancer | Whether the patient's death was due to cancer or not | N/A | “Living”, “Died of Disease”, “Died of Other Causes” |
| radio_therapy | Whether or not the patient had radiotherapy as a treatment | Boolean | |
| tumor_stage | Stage of cancer based on involvement of surrounding structures, lymph nodes and distant spread | N/A | 0, 1, 2, 3, 4, N/A |

Analysis

Uploading relevant table and cleaning data

- All rows with "N/A" or "" for any relevant columns have been removed to ensure no data is missing.
- All categorical variables are transformed to be treated as categorical by the table.
- After initial analysis, "tumor_stage==0" and "tumor_stage==4" have been removed due to lack of sufficient amounts of data points (1 case of stage 0 tumors, 5 cases of stage 4 tumors).

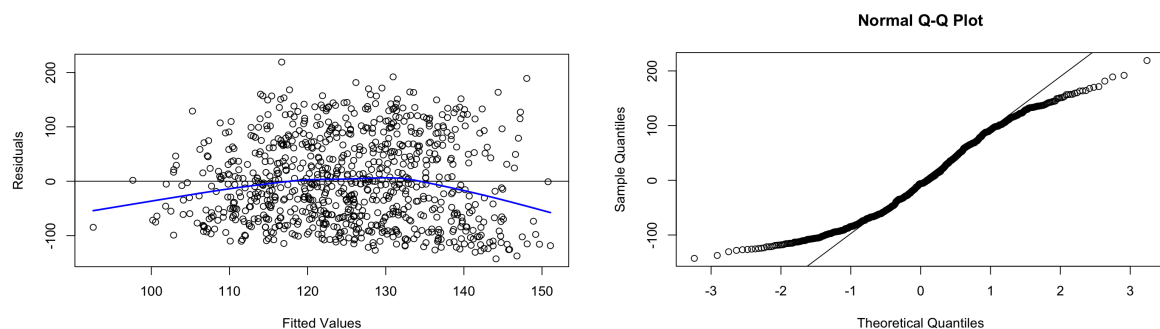
| type_of_breast_surgery | tumor_stage | chemotherapy | hormone_therapy | radio_therapy | death_from_cancer |
|--|---------------------------|------------------|------------------|------------------|-------------------------------------|
| BREAST CONSERVING: 372 MASTECTOMY: 464 | 1: 277 2: 485 3: 74 | 0: 586 1: 250 | 0: 349 1: 487 | 0: 246 1: 590 | Living: 468 Died of Disease: 368 |

| | Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. |
|-------------------------|-------|---------|--------|--------|---------|--------|
| age_at_diagnosis | 26.36 | 48.24 | 57.33 | 57.31 | 67.05 | 96.29 |
| overall_survival_months | 0.10 | 57.67 | 115.10 | 125.18 | 186.71 | 337.03 |

Exploratory data analysis

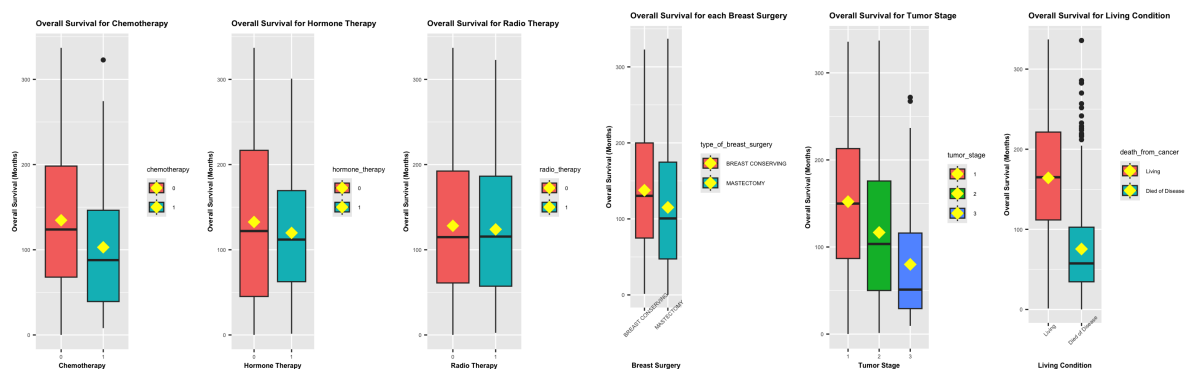
Residual plot and QQ-plot (overall_survival_months against age_at_diagnosis)

- There does not appear to be any observable patterns in residual values beyond being centered around 0, though the slight variation in residual spread as fitted values increase may indicate heteroscedasticity.
- QQ-plot indicates distribution of residuals is light-tailed normal.



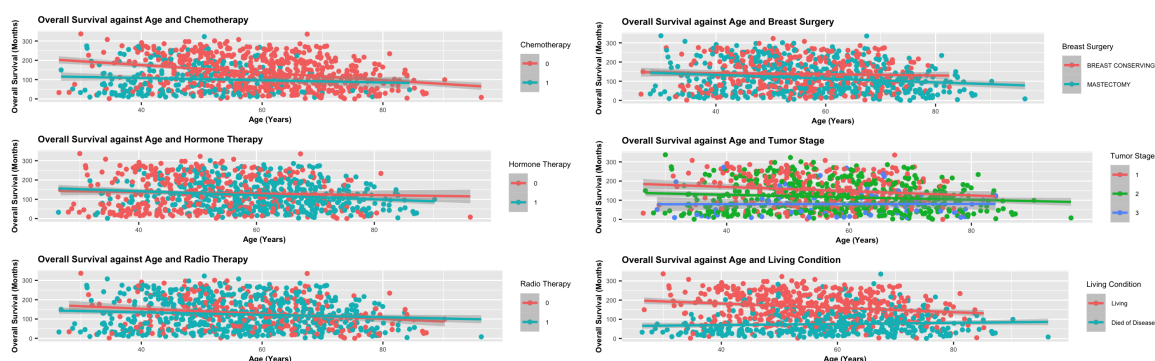
Boxplots of survival against interactions between age and categorical variables

- overall_survival_months for those that took chemotherapy is generally lower than for those without in terms of mean and median, and the range of overall_survival_months is smaller.
- Mean and median overall_survival_months for patients with hormone therapy are lower than those without, while the interquartile range for those with hormone therapy is smaller than for those without.
- Boxplots for overall_survival_months against each radio_therapy class are extremely similar, including the mean and median values.
- overall_survival_months for those with mastectomy is lower than for those with breast conserving surgery in terms of mean and median, and the values are more skewed toward lower values.
- Increasing levels of tumor stages provide decreasing mean, median and percentile values.
- Patients that died from disease have significantly lower overall_survival_months values than those who lived and the most extreme outliers.



Scatterplots of survival against interactions between age and categorical variables

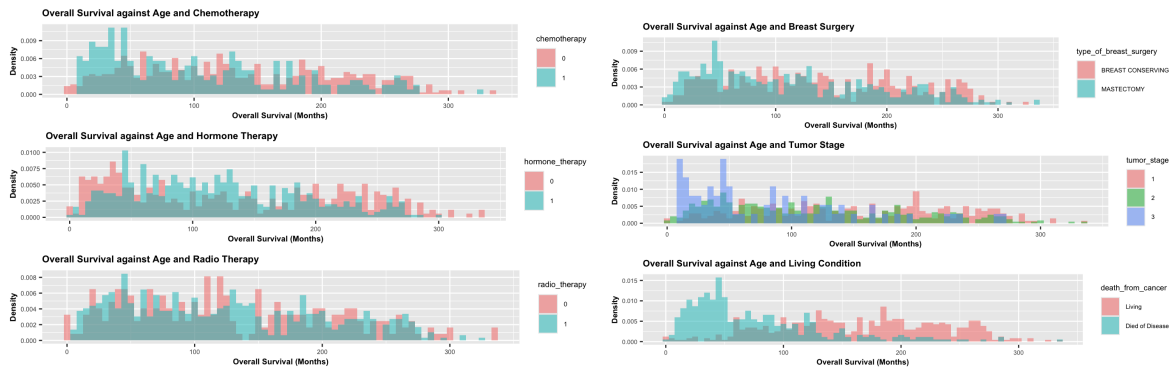
- overall_survival_months varies significantly by chemotherapy, tumor_stage, and death_from_cancer, and the interaction terms for age with chemotherapy, type_of_breast_surgery, tumor_stage, and death_from_cancer appear to be statistically significant.
- Confidence intervals of linear models heavily overlap for hormone_therapy and radio_therapy.
- Confidence intervals of linear models for chemotherapy, type_of_breast_surgery and tumor_stage only overlap for more extreme age values.



Histograms of survival against interactions between age and categorical variables

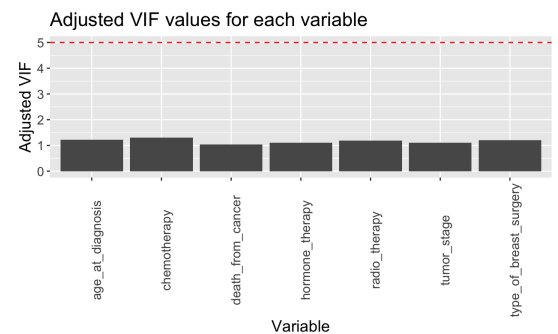
- The histograms for those that took chemotherapy, mastectomy or had died from disease compared to those without chemotherapy, with breast preserving surgery or are alive respectively are significantly more skewed to the right.
- Histogram peaks for those that took hormone therapy are generally greater than for those without hormone therapy within the range of overall_survival_months between 45 and 195.

- The histogram distributions between radio therapy categories are relatively similar.
- The histogram for those with stage 3 tumors compared to that for the other tumor stages is significantly more skewed to the right.
- Given our exploratory analysis on overall_survival_months within boxplots and histograms, we will assume that the overall distribution of the variable is skewed to the right, and thus may have heteroscedasticity; as such, a log transformation would be beneficial for model fitting.



Variance Inflation Factor (VIF) on covariates

- VIF for each variable is calculated to test for multicollinearity, which could make it difficult to interpret coefficients, and it reduces the power of the model to identify statistically significant independent variables.
- $GVIF_{.1,.2,Df..}$ is used for comparison due to different degrees of freedom for some variables.
- Horizontal line of $VIF = 5$ to indicate severe correlation of a variable with other variables; any variables with $VIF > 5$ are to be removed to reduce multicollinearity in the model.
- $GVIF_{.1,.2,Df..}$ for all variables are significantly lower than 5; multicollinearity between all variables is relatively low, thus no variables need to be removed.



Model selection 1

- As our goal is to determine how the various forms of therapy influence a patient's survival until death, interaction terms regarding all combinations of therapies are considered, since it is unclear if the effect of one therapy will influence the effect of another (e.g. Chemotherapy and hormone therapy in similar timeframe).
- Interaction terms for age and each type of therapy are included, as prior studies have indicated varying degrees of influence between age and treatment method (Given, Given 2008; Cleveland Clinic, 2024; U.S. National Library of Medicine; Steinfeld, Diamond, Hanks, Coia, Kramer, 1989).
- Interaction terms for tumor_stage and each type of therapy are also included since prior studies have indicated that the type of treatment a patient receives is influenced by the stage, size of tumor and the spread of cancer cells. (“Invasive Ductal Carcinoma”, 2024; “Treatment of breast cancer”, 2024)
- Due to not being significant parts of the question of interest and the indeterminate interaction between them and the therapies, interaction terms for type_of_breast_surgery and death_from_cancer are ignored for the full model.
- We had attempted to fit linear models without the log transformation, and the resulting residual plots had demonstrated heteroscedasticity.
- Full model: $\log(\text{overall_survival_months}) \sim \text{age_at_diagnosis} + \text{chemotherapy} * \text{hormone_therapy} * \text{radio_therapy} + \text{age_at_diagnosis} : \text{chemotherapy} + \text{age_at_diagnosis} : \text{hormone_therapy} +$

age_at_diagnosis : radio_therapy + tumor_stage + tumor_stage : chemotherapy + tumor_stage : hormone_therapy + tumor_stage : radio_therapy + type_of_breast_surgery + death_from_cancer

- Forward and backward AIC selection with the same full model was attempted, but only backward selection produced the same model, while forward selection produced a less optimal model.
- Model finalmod_both has age_at_diagnosis:chemotherapy1 and hormone_therapy1:tumor_stage2 to be not statistically significant on the 5% significance level.

| | Estimate | Std. Error | t value | Pr(> t) |
|-----------------------------------|-----------|------------|---------|----------|
| (Intercept) | 5.947000 | 0.214458 | 27.730 | < 2e-16 |
| age_at_diagnosis | -0.014682 | 0.003806 | -3.857 | 0.000124 |
| chemotherapy1 | -0.701659 | 0.294312 | -2.384 | 0.017348 |
| hormone_therapy1 | -0.622362 | 0.266825 | -2.332 | 0.019915 |
| tumor_stage2 | -0.238917 | 0.093835 | -2.546 | 0.011073 |
| tumor_stage3 | -0.815091 | 0.166351 | -4.900 | 1.15e-06 |
| death_from_cancerDied of Disease | -0.857453 | 0.054094 | -15.851 | < 2e-16 |
| age_at_diagnosis:chemotherapy1 | 0.009678 | 0.005384 | 1.797 | 0.072624 |
| age_at_diagnosis:hormone_therapy1 | 0.009458 | 0.004510 | 2.097 | 0.036273 |
| hormone_therapy1:tumor_stage2 | 0.216084 | 0.117380 | 1.841 | 0.065997 |
| hormone_therapy1:tumor_stage3 | 0.711055 | 0.203734 | 3.490 | 0.000508 |

Model selection 2

- finalmod_both was compared to a version of the final model without interaction term age_at_diagnosis:chemotherapy1 (finalmod_both_1), as indicated below.
- Model finalmod_both_1 has all terms to be statistically significant on the 5% level.
- This model uses living, stage 1 cancer patients without chemo or hormone therapy as the baseline.

| | Estimate | Std. Error | t value | Pr(> t) |
|----------------------------------|-----------|------------|---------|----------|
| (Intercept) | 5.773782 | 0.191845 | 30.096 | < 2e-16 |
| age_at_diagnosis | -0.011469 | 0.003365 | -3.408 | 0.000685 |
| chemotherapy1 | -0.189144 | 0.073044 | -2.589 | 0.009782 |
| hormone_therapy1 | -0.570923 | 0.265644 | -2.149 | 0.031908 |
| tumor_stage2 | -0.271223 | 0.092222 | -2.941 | 0.003363 |
| tumor_stage3 | -0.830862 | 0.166344 | -4.995 | 7.19e-07 |
| death_from_cancerDied of Disease | -0.857409 | 0.054167 | -15.829 | < 2e-16 |

| | | | | |
|-----------------------------------|----------|----------|-------|----------|
| age_at_diagnosis:hormone_therapy1 | 0.008276 | 0.004467 | 1.852 | 0.064316 |
| hormone_therapy1:tumor_stage2 | 0.240445 | 0.116753 | 2.059 | 0.039766 |
| hormone_therapy1:tumor_stage3 | 0.726981 | 0.203816 | 3.567 | 0.000382 |

Model comparisons

- regsubsets() applied to the original full model for selection of best subsets (see Appendix section 1)
- For the selected best subsets with 11 and 12 parameters, both contain the parameters age_at_diagnosis, chemotherapy1, hormone_therapy1, tumor_stage2, death_from_cancerDied of Disease, age_at_diagnosis:chemotherapy, age_at_diagnosis:hormone_therapy1, hormone_therapy1:tumor_stage3 and radio_therapy1:tumor_stage3; the interaction term radio_therapy1:tumor_stage3 was not present in finalmod_both or finalmod_both_1.
- The parameters tumor_stage3 and hormone_therapy1:tumor_stage2 are included for the selected best subset with 11 parameters but not the selected best subset with 12 parameters, though both were present in finalmod_both and finalmod_both_1.
- The parameters chemotherapy1:hormone_therapy1, hormone_therapy1:radio_therapy1 and age_at_diagnosis:radio_therapy1 are included for the selected best subset with 12 parameters but not the selected best subset with 11 parameters, but the interaction terms involving radiotherapy are not present in finalmod_both or finalmod_both_1.

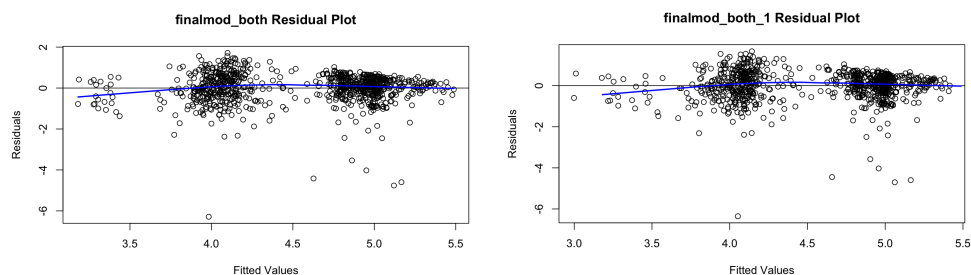
Model statistics

- Compared to finalmod_both, the model finalmod_both_1 has an AIC value increased by 1.21, and adjusted R² decreased by 0.0018577 (7 d.p.).

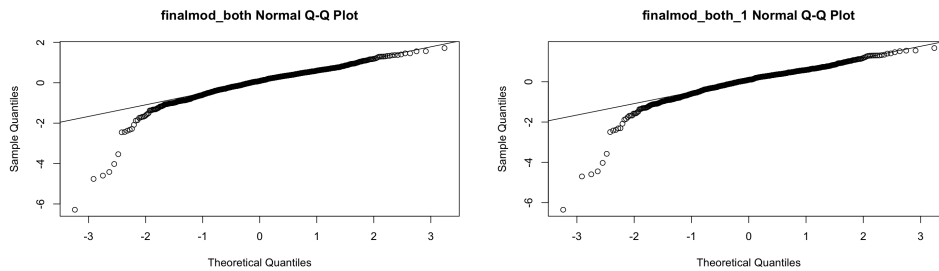
| Statistics | finalmod_both | finalmod_both_1 |
|--------------------------------|---------------|-----------------|
| Number of estimated parameters | 12 | 11 |
| Residual standard error | 0.7441687 | 0.7451730 |
| Multiple R-squared | 0.3204380 | 0.3177766 |
| Adjusted R-squared | 0.3122009 | 0.3103432 |
| AIC | 1891.72 | 1892.93 |

Residual plot and QQ-plot

- For both models, there does not appear to be any observable patterns in residual values beyond being centered around 0, though there are outliers present for larger fitted values.



- The QQ-plots indicate the distribution of residuals are normal distributions that are slightly left-skewed for both models; given the outliers in the residual plot, the left-skew may be due to outliers as well.



- By the parsimony principle, and due to our project question's focus on inference, the model `finalmod_both_1` was selected due to having fewer variables, and thus being easier to interpret, while producing similar results and a marginally worse fit to the data.

Discussion

Analysis

The final regression model includes the following key variables: age, chemotherapy, hormone therapy, and tumor stage (specifically stages 2 and 3). It also accounts for the outcome of death from cancer (death from cancer and death from other causes) and incorporates interaction terms between age and hormone therapy, as well as hormone therapy and tumor stage (for stages 2 and 3).

The results from bidirectional selection and backward selection consistently provided the same set of covariates, but the results from forward selection did not, possibly due to it missing covariates that are only significant in combination with other covariates. The covariates of the final regression model are all statistically significant on the 5% significance level, with the exception of the interaction term between age and hormone therapy. This consistency between bidirectional and backward selection enhances the credibility of the findings, indicating that the observed relationships are relatively robust and not due to random chance.

In the model selection, we discarded radiotherapy, indicating that it doesn't provide a significant influence on the survival period of patients. For chemotherapy, regardless of tumor stage, the survival period is reduced by a factor of $e^{-0.189144}$. This indicates that chemotherapy consistently reduces survival periods, and its negative effects on survival do not change at higher tumor stages.

For hormone therapy, it initially reduces the survival period by multiplying it by $e^{-0.570923}$ at tumor stage 1, but its effect decreases to a multiplicative reduction by a factor of $e^{-0.330478}$ in stage 2, and it even leads to a multiplicative increase in survival by a factor of $e^{0.156058}$ in stage 3. From tumor stage 1 to stage 3, the impact of hormone therapy changes from harmful to beneficial, ultimately improving survival chances. This highlights the importance of carefully considering the use of hormone therapy based on tumor stage.

The model shows a negative correlation between `age_at_diagnosis` and `overall_survival_months`, indicating that for each year of age (at the time of diagnosis), the patient's months of survival is multiplied by $e^{-0.011469}$. This indicates that younger cancer patients will have a higher chance of recovering from IDC (higher overall survival months). The presence of hormone therapy increases the factor of multiplication of age to $e^{-0.003193}$, implying that the impact of age on survival months becomes less severe with hormone therapy.

The final model shows that the intercept for patients who have died from disease is multiplied by a factor of $e^{-0.857409}$ compared to living patients. This is consistent with the scatterplot between age/living condition from the exploratory analysis, where the overall survival was significantly lower for deceased patients compared to living

ones. The scatterplot does not show the two slopes to be parallel, which would have been accounted for by the interaction term between `death_from_cancer` and `age_at_diagnosis`; however, this term was removed during model selection due to low significance.

To further validate the reliability of the model, the plot of fitted values versus residuals shows a good fit with residuals centered around zero, demonstrating homoscedasticity. Additionally, the residuals are randomly scattered without any discernible pattern beyond being centered around 0, suggesting that the errors are independent and that the relationship between the independent variables and the dependent variable is assumed to be linear. Furthermore, the points in the QQ plot align closely along the diagonal line, indicating that the residuals are approximately normally distributed. These observations confirm that the regression model does not violate the key assumptions of regression analysis, thereby minimizing the risk of drawing incorrect conclusions.

Conclusion

To address our initial research question, our model suggests that chemotherapy and hormone therapy have varying influences on survival time before death depending on the initial tumor stage while radiotherapy does not have a significant association with survival time before death. In general, our model suggests that chemotherapy has a negative association with survival time before death regardless of tumor stage level while hormone therapy has a negative association for patients with an initial tumor stage of 1 and 2 while it has a positive association when the patient's original tumor stage was 3. Our model does not align with prior research as prior research suggests that chemotherapy is generally effective and increases the survival time before death. Similarly prior research suggests that hormone therapy and radiotherapy are generally effective at treating patients with IDC. However, our model aligns with prior research suggesting chemotherapy is most effective for later stages of cancer (Penn Medicine, n.d.).

Our model likely differs from prior research due to several limitations in our model. One limitation is our variable selection method, a bidirectional stepwise selection starting with a full model. While bidirectional model selection allows for better flexibility and model fit onto the data, there is also the risk of the final model being overfitted to the existing data. Our response variable being the log of `overall_survival_months` in order to maintain homoscedasticity also makes our model estimates to be rather difficult to interpret. Moreover, our results may have been skewed from using tumor stage 1 as our baseline, since we did not have enough data points for tumor stage 0. Lastly, our model could not account for the context of the patients undergoing such treatments, such as the patients that took the treatment likely already being in poorer health than those that did not, or the side effects of such treatments impacting health, such as organ damage from chemotherapy ("Side effects of chemotherapy", Canadian Cancer Society, 2024), osteoporosis from hormone therapy ("Side effects of hormone therapy", Canadian Cancer Society, 2017) and low blood cell counts from radiation therapy ("Side effects of radiation therapy", Canadian Cancer Society, 2017). Overall, the limitations of our model are reflected by the relatively low adjusted R^2 of 0.3103432 (7 s.f.), which indicates our model only accounts for around 31.03% (2 d.p.) of the variability in survival time before death for IDC patients.

The R code script was submitted by Nicholas Tam.

Sources

- *Breast cancer gene expression profiles (METABRIC)*. Kaggle. (2016, May 10). <https://www.kaggle.com/datasets/raghadalharbi/breast-cancer-gene-expression-profiles-metabric>
- CBioPortal for Cancer Genomics. (n.d.). https://www.cbioportal.org/study/summary?id=brca_metabric
- DePolo, J. (2024, October 2). *Invasive ductal carcinoma (IDC)*. Breastcancer.org - Breast Cancer Information and Support. <https://www.breastcancer.org/types/invasive-ductal-carcinoma>
- Wright, P. (2023, March 21). *Invasive ductal carcinoma (IDC)*. Johns Hopkins Medicine. <https://www.hopkinsmedicine.org/health/conditions-and-diseases/breast-cancer/invasive-ductal-carcino>

[ma-idx#:~:text=Radiation%20therapy%20might%20be%20part,lymph%20nodes%2C%E2%80%9D%20Wright%20says.](#)

- *Tumor size and staging*. Susan G. Komen®. (2024, May 2).
<https://www.komen.org/breast-cancer/diagnosis/stages-staging/tumor-size/#:~:text=Tumor%20size%20is%20related%20to,the%20size%20of%20the%20tumor.>
- Given, B., & Given, C. W. (2008, December 15). *Older adults and cancer treatment*. Cancer.
<https://pmc.ncbi.nlm.nih.gov/articles/PMC2606910/#S10>
- *Hormone therapy for cancer*. Cleveland Clinic. (2024, May 1).
<https://my.clevelandclinic.org/health/treatments/17108-hormone-therapy-to-treat-cancer>
- U.S. National Library of Medicine. (n.d.). *Aging changes in hormone production: Medlineplus medical encyclopedia*. MedlinePlus.
<https://medlineplus.gov/ency/article/004000.htm#:~:text=In%20women%2C%20estrogen%20and%20p,rolactin,Cortisol>
- Steinfeld, A. D., Diamond, J. J., Hanks, G. E., Coia, L. R., & Kramer, S. (1989). Patient age as a factor in radiotherapy. Data from the patterns of care study. *Journal of the American Geriatrics Society*, 37(4), 335–338. <https://doi.org/10.1111/j.1532-5415.1989.tb05501.x>
- *Invasive Ductal Carcinoma*. Cleveland Clinic (June 27, 2024)
<https://my.clevelandclinic.org/health/diseases/22117-invasive-ductal-carcinoma-idx>
- Penn Medicine. (n.d.). *Invasive Ductal Carcinoma (IDC)*. Pennmedicine.org.
<https://www.pennmedicine.org/cancer/types-of-cancer/breast-cancer/types-of-breast-cancer/invasive-ductal-carcinoma#:~:text=The%20IDC%20treatment%20your%20physician, focuses%20only%20on%20breast%20cancer>
- “Treatment of breast cancer stages I-III”. (2024, September 22). *Treatment of breast cancer stages I-III*. American Cancer Society.
<https://www.cancer.org/cancer/types/breast-cancer/treatment/treatment-of-breast-cancer-by-stage/treatment-of-breast-cancer-stages-i-iii.html#:~:text=The%20stage%20of%20your%20breast,gone%20through%20menopause%20or%20not>
- *Side effects of chemotherapy*. Canadian Cancer Society. (2024, May).
<https://cancer.ca/en/treatments/treatment-types/chemotherapy/side-effects-of-chemotherapy>
- *Side effects of hormone therapy*. Canadian Cancer Society. (2017).
<https://cancer.ca/en/treatments/treatment-types/hormone-therapy/side-effects-of-hormone-therapy>
- *Side effects of radiation therapy*. Canadian Cancer Society. (2017b).
<https://cancer.ca/en/treatments/treatment-types/radiation-therapy/side-effects-of-radiation-therapy>

Appendix

Section 1

| | age_at_diagnosis | chemotherapy1 | hormone_therapy1 | radio_therapy1 | tumor_stage2 | tumor_stage3 |
|-------|------------------|---------------|------------------|----------------|--------------|--------------|
| 1 (1) | " " | " " | " " | " " | " " | " " |
| 2 (1) | " " | "*" | " " | " " | " " | " " |
| 3 (1) | "*" | "*" | " " | " " | " " | " " |
| 4 (1) | "*" | "*" | " " | " " | " " | " " |
| 5 (1) | "*" | "*" | " " | " " | " " | " " |
| 6 (1) | "*" | "*" | " " | " " | " " | " " |
| 7 (1) | "*" | "*" | " " | " " | "*" | " " |
| 8 (1) | "*" | "*" | " " | " " | "*" | " " |

| | | | | | | |
|--------|-----|-----|-----|-----|-----|-----|
| 9 (1) | "*" | "*" | "*" | " " | "*" | " " |
| 10 (1) | "*" | "*" | "*" | " " | "*" | "*" |
| 11 (1) | "*" | "*" | "*" | " " | "*" | "*" |
| 12 (1) | "*" | "*" | "*" | " " | "*" | "*" |
| 13 (1) | "*" | "*" | "*" | " " | "*" | "*" |

```

age_at_diagnosis      chemotherapy1      hormone_therapy1      radio_therapy1      tumor_stage2
tumor_stage3
1  ( 1 )  " "              " "              " "              " "              " "
" "
2  ( 1 )  " "              "* "              " "              " "              " "
" "
3  ( 1 )  "* "              "* "              " "              " "              " "
" "
4  ( 1 )  "* "              "* "              " "              " "              " "
" "
5  ( 1 )  "* "              "* "              " "              " "              " "
" "
6  ( 1 )  "* "              "* "              " "              " "              " "
" "
7  ( 1 )  "* "              "* "              " "              " "              "* "
" "
8  ( 1 )  "* "              "* "              " "              " "              "* "
" "
9  ( 1 )  "* "              "* "              "* "              " "              " "
" "
10 ( 1 )  "* "              "* "              "* "              " "              "* "
"* "
11 ( 1 )  "* "              "* "              "* "              " "              "* "
"* "
12 ( 1 )  "* "              "* "              "* "              " "              "* "
" "
13 ( 1 )  "* "              "* "              "* "              " "              "* "
"* "

```

| | type_of_breast_surgeryMASTECTOMY | death_from_cancerDiedof Disease |
|-------|----------------------------------|---------------------------------|
| 1 (1) | " " | "*" |
| 2 (1) | " " | "*" |
| 3 (1) | " " | "*" |
| 4 (1) | " " | "*" |
| 5 (1) | " " | "*" |
| 6 (1) | " " | "*" |

| | | |
|--------|-----|-----|
| 7 (1) | " " | "*" |
| 8 (1) | " " | "*" |
| 9 (1) | " " | "*" |
| 10 (1) | " " | "*" |
| 11 (1) | " " | "*" |
| 12 (1) | " " | "*" |
| 13 (1) | " " | "*" |

| | type_of_breast_surgery | MASTECTOMY | death_from_cancer | Died of Disease |
|----------|------------------------|------------|-------------------|-----------------|
| 1 (1) | " " | | "*" | |
| 2 (1) | " " | | "*" | |
| 3 (1) | " " | | "*" | |
| 4 (1) | " " | | "*" | |
| 5 (1) | " " | | "*" | |
| 6 (1) | " " | | "*" | |
| 7 (1) | " " | | "*" | |
| 8 (1) | " " | | "*" | |
| 9 (1) | " " | | "*" | |
| 10 (1) | " " | | "*" | |
| 11 (1) | " " | | "*" | |
| 12 (1) | " " | | "*" | |
| 13 (1) | " " | | "*" | |

| | chemotherapy1: hormone_therapy1 | chemotherapy1: radio_therapy1 | hormone_therapy1:radio_thera py1 |
|--------|------------------------------------|----------------------------------|-------------------------------------|
| 1 (1) | " " | " " | " " |
| 2 (1) | " " | " " | " " |
| 3 (1) | " " | " " | " " |
| 4 (1) | " " | " " | " " |
| 5 (1) | " " | " " | " " |
| 6 (1) | " " | " " | " " |
| 7 (1) | " " | " " | " " |
| 8 (1) | " " | " " | " " |
| 9 (1) | "*" | " " | " " |
| 10 (1) | " " | " " | " " |
| 11 (1) | " " | " " | " " |

| | | | |
|--------|-----|-----|-----|
| 12 (1) | "*" | " " | "*" |
| 13 (1) | " " | " " | " " |

chemotherapy1:hormone_therapy1 chemotherapy1:radio_therapy1
hormone_therapy1:radio_therapy1

| | | | | |
|----|-------|-----|-----|-----|
| 1 | (1) | " " | " " | " " |
| 2 | (1) | " " | " " | " " |
| 3 | (1) | " " | " " | " " |
| 4 | (1) | " " | " " | " " |
| 5 | (1) | " " | " " | " " |
| 6 | (1) | " " | " " | " " |
| 7 | (1) | " " | " " | " " |
| 8 | (1) | " " | " " | " " |
| 9 | (1) | "*" | " " | " " |
| 10 | (1) | " " | " " | " " |
| 11 | (1) | " " | " " | " " |
| 12 | (1) | "*" | " " | "*" |
| 13 | (1) | " " | " " | " " |

| | age_at_diagnosis:chemotherapy1 | age_at_diagnosis:hormone_therapy1 | age_at_diagnosis:radio_therapy1 |
|--------|--------------------------------|-----------------------------------|---------------------------------|
| 1 (1) | " " | " " | " " |
| 2 (1) | " " | " " | " " |
| 3 (1) | " " | " " | " " |
| 4 (1) | " " | " " | " " |
| 5 (1) | " " | " " | " " |
| 6 (1) | "*" | " " | " " |
| 7 (1) | "*" | " " | " " |
| 8 (1) | "*" | " " | " " |
| 9 (1) | "*" | "*" | " " |
| 10 (1) | "*" | "*" | " " |
| 11 (1) | "*" | "*" | " " |
| 12 (1) | "*" | "*" | "*" |
| 13 (1) | "*" | "*" | " " |

age_at_diagnosis:chemotherapy1 age_at_diagnosis:hormone_therapy1

| | | | |
|---|-------|-----|-----|
| 1 | (1) | " " | " " |
| 2 | (1) | " " | " " |
| 3 | (1) | " " | " " |

| | | | |
|----|-------|-----|-----|
| 4 | (1) | " " | " " |
| 5 | (1) | " " | " " |
| 6 | (1) | "*" | " " |
| 7 | (1) | "*" | " " |
| 8 | (1) | "*" | " " |
| 9 | (1) | "*" | "*" |
| 10 | (1) | "*" | "*" |
| 11 | (1) | "*" | "*" |
| 12 | (1) | "*" | "*" |
| 13 | (1) | "*" | "*" |

| | chemotherapy1:tumor_stage2 | chemotherapy1:tumor_stage3 |
|--------|----------------------------|----------------------------|
| 1 (1) | " " | " " |
| 2 (1) | " " | " " |
| 3 (1) | " " | " " |
| 4 (1) | " " | " " |
| 5 (1) | " " | " " |
| 6 (1) | " " | " " |
| 7 (1) | " " | " " |
| 8 (1) | " " | " " |
| 9 (1) | " " | " " |
| 10 (1) | " " | " " |
| 11 (1) | " " | " " |
| 12 (1) | " " | " " |
| 13 (1) | "*" | "*" |

| | age_at_diagnosis:radio_therapy1 | chemotherapy1:tumor_stage2 | chemotherapy1:tumor_stage3 |
|----|---------------------------------|----------------------------|----------------------------|
| 1 | (1) | " " | " " |
| 2 | (1) | " " | " " |
| 3 | (1) | " " | " " |
| 4 | (1) | " " | " " |
| 5 | (1) | " " | " " |
| 6 | (1) | " " | " " |
| 7 | (1) | " " | " " |
| 8 | (1) | " " | " " |
| 9 | (1) | " " | " " |
| 10 | (1) | " " | " " |
| 11 | (1) | " " | " " |
| 12 | (1) | "*" | " " |

13 (1) " "

"*"

"*"

| | hormone_therapy1:tumor_stage2 | hormone_therapy1:tumor_stage3 |
|--------|-------------------------------|-------------------------------|
| 1 (1) | " " | " " |
| 2 (1) | " " | " " |
| 3 (1) | " " | " " |
| 4 (1) | " " | " " |
| 5 (1) | " " | " " |
| 6 (1) | " " | "*" |
| 7 (1) | " " | "*" |
| 8 (1) | "*" | "*" |
| 9 (1) | " " | "*" |
| 10 (1) | "*" | "*" |
| 11 (1) | "*" | "*" |
| 12 (1) | "*" | "*" |
| 13 (1) | "*" | "*" |

hormone_therapy1:tumor_stage2 hormone_therapy1:tumor_stage3
radio_therapy1:tumor_stage2

| | | | |
|----------|-----|-----|-----|
| 1 (1) | " " | " " | " " |
| 2 (1) | " " | " " | " " |
| 3 (1) | " " | " " | " " |
| 4 (1) | " " | " " | " " |
| 5 (1) | " " | "*" | " " |
| 6 (1) | " " | "*" | " " |
| 7 (1) | " " | "*" | " " |
| 8 (1) | "*" | "*" | " " |
| 9 (1) | " " | "*" | " " |
| 10 (1) | "*" | "*" | " " |
| 11 (1) | "*" | "*" | " " |
| 12 (1) | " " | "*" | " " |
| 13 (1) | "*" | "*" | " " |

| | radio_therapy1:tumor_stage2 | radio_therapy1:tumor_stage3 | chemotherapy1: hormone_therapy1: radio_therapy1 |
|-------|-----------------------------|-----------------------------|---|
| 1 (1) | " " | " " | " " |
| 2 (1) | " " | " " | " " |

| | | | |
|--------|-----|-----|-----|
| 3 (1) | " " | " " | " " |
| 4 (1) | " " | "*" | " " |
| 5 (1) | " " | "*" | " " |
| 6 (1) | " " | "*" | " " |
| 7 (1) | " " | "*" | " " |
| 8 (1) | " " | "*" | " " |
| 9 (1) | " " | "*" | " " |
| 10 (1) | " " | " " | " " |
| 11 (1) | " " | "*" | " " |
| 12 (1) | " " | "*" | " " |
| 13 (1) | " " | "*" | " " |

```

radio_therapy1:tumor_stage3 chemotherapy1:hormone_therapy1:radio_therapy1
1  ( 1 )  " "                " "
2  ( 1 )  " "                " "
3  ( 1 )  " "                " "
4  ( 1 )  "*"                " "
5  ( 1 )  "*"                " "
6  ( 1 )  "*"                " "
7  ( 1 )  "*"                " "
8  ( 1 )  "*"                " "
9  ( 1 )  "*"                " "
10 ( 1 )  " "                " "
11 ( 1 )  "*"                " "
12 ( 1 )  "*"                " "
13 ( 1 )  "*"                " "

```