

BREAST CANCER

Supervivencia a 5 años del diagnóstico

Por Luisa Beccar Varela

Indice

1. Contexto y audiencia
2. Hipótesis y preguntas de interés
3. Metadata
4. Análisis Exploratorio
5. Insights y sugerencias



1- Contexto y audiencia

Contexto:

Dado que el cáncer de mama es el cáncer más diagnosticado en mujeres argentinas (19,6%) y su mayor causa de mortalidad (16,4/100000), y teniendo personalmente varias familiares que padecieron la enfermedad, algunas que sobrevivieron muchos años tras el diagnóstico y otras en quienes fue fatal, busco con este análisis responder a la pregunta: ¿Sobrevivirá una mujer más de 5 años tras el diagnóstico de esta enfermedad?. Determine el lapso de 5 años de manera totalmente arbitraria, pensando que es una buena cantidad de tiempo que un médico puede proveer a las pacientes como pronóstico para que ellas tengan de perspectiva y poder organizar el tiempo de vida que les queda.

Audiencia:

Para el médico, ginecologo mastologo, responsable de informar los resultados a la pacientes y proponer un tratamiento, me parecería fundamental conocer el pronóstico a 5 años. Las pacientes recibiendo la información podrían beneficiarse al conocer el pronóstico y poder tomar decisiones acerca de cómo seguir con sus vidas en el tiempo informado.



2- Hipótesis y preguntas de interés

Pregunta principal: ¿Sobrevivirá la paciente más de 5 años tras el diagnóstico de cáncer de mama, dadas sus características personales y del tumor?

Preguntas secundarias:

1. ¿Cómo se relaciona el 6th Stage con los meses de supervivencia?
2. ¿Cómo influyen los receptores hormonales en los meses de supervivencia?
3. ¿Tienen los tumores de menor diferenciación menor supervivencia?
4. ¿Existe relación entre la edad al diagnóstico y su sobrevida?

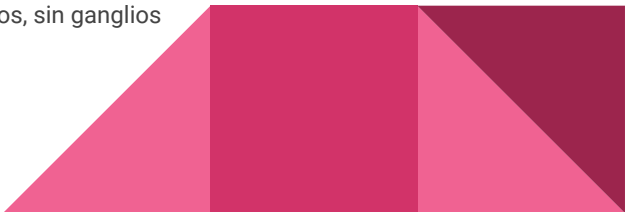
Limitaciones:

La [base de datos](#) se tomó de mujeres en Estados Unidos de entre 30 y 69 años de edad, con carcinoma ductal infiltrante y carcinoma lobular, entre 2006 y 2010, excluyendo pacientes con tamaño tumoral desconocido, sin ganglios regionales examinados, sin ganglios examinados positivos y que sobrevivieron menos de 1 mes del diagnóstico.

La base no cuenta con datos sobre el marcador Her2, que se debe tener en cuenta para la clasificación en 6th Stage.

Tampoco se menciona si las pacientes diagnosticadas recibieron algún tratamiento o no, ni cual.

Las razas de EEUU quizás no sean las mismas que la población argentina. Ignoraremos ese tema ya que tampoco contamos con una descripción de dicho tema para las mujeres argentinas .



3- Metadata

Base de Datos:

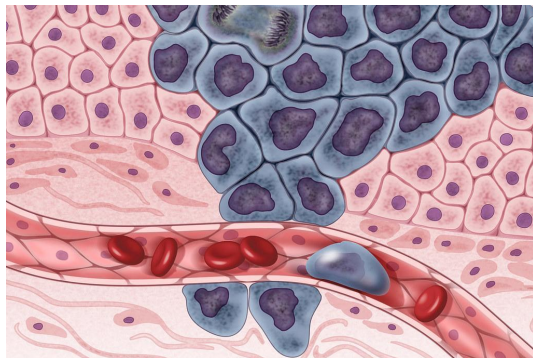
<https://www.kaggle.com/datasets/reihanenamdari/breast-cancer>

Programa SEER del National Cancer Institute en noviembre 2017: Diagnósticos de mujeres con carcinoma ductal infiltrante y carcinoma lobular de mama, recolectados entre el 2006 y 2010.

17 Atributos

4024
registros

0 nulls



Tumor Size

T Stage

N Stage

A Stage

6th Stage

Regional Node Examined

Regional Node Positive

Grade

Differentiate

Estrogen Status

Progesterone Status

Age

Marital Status

Race

Survival Months

Status

Sobrevida

60 meses (5 años)



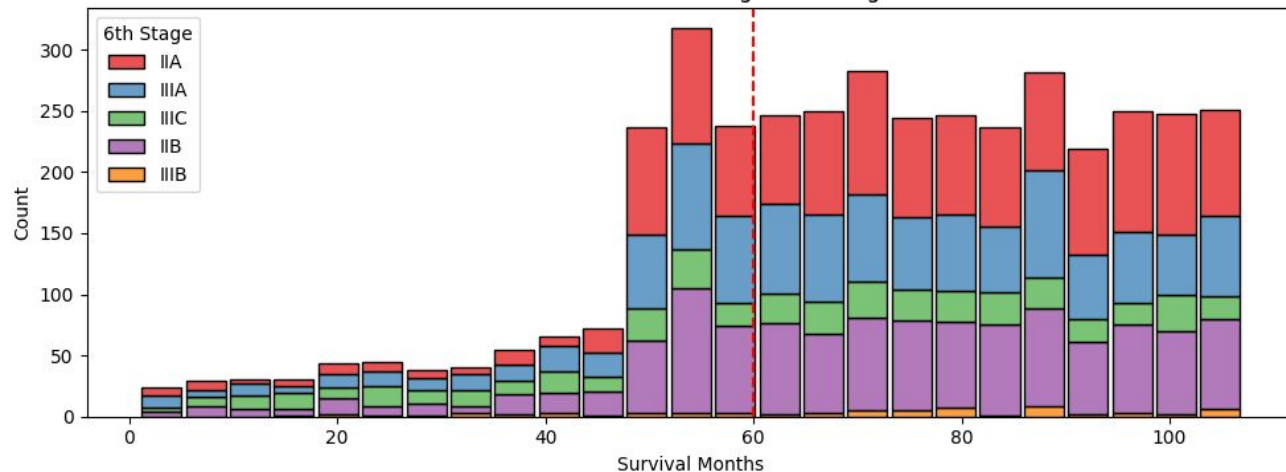
4- Análisis Exploratorio



Survival months

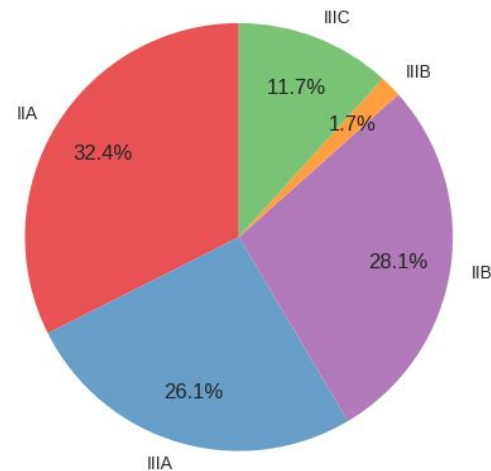
Se observa que a partir de los 44 meses se quintuplica aproximadamente la cantidad de registros en todos los meses de sobrevida.

Meses de sobrevida segun 6th Stage



6th Stage

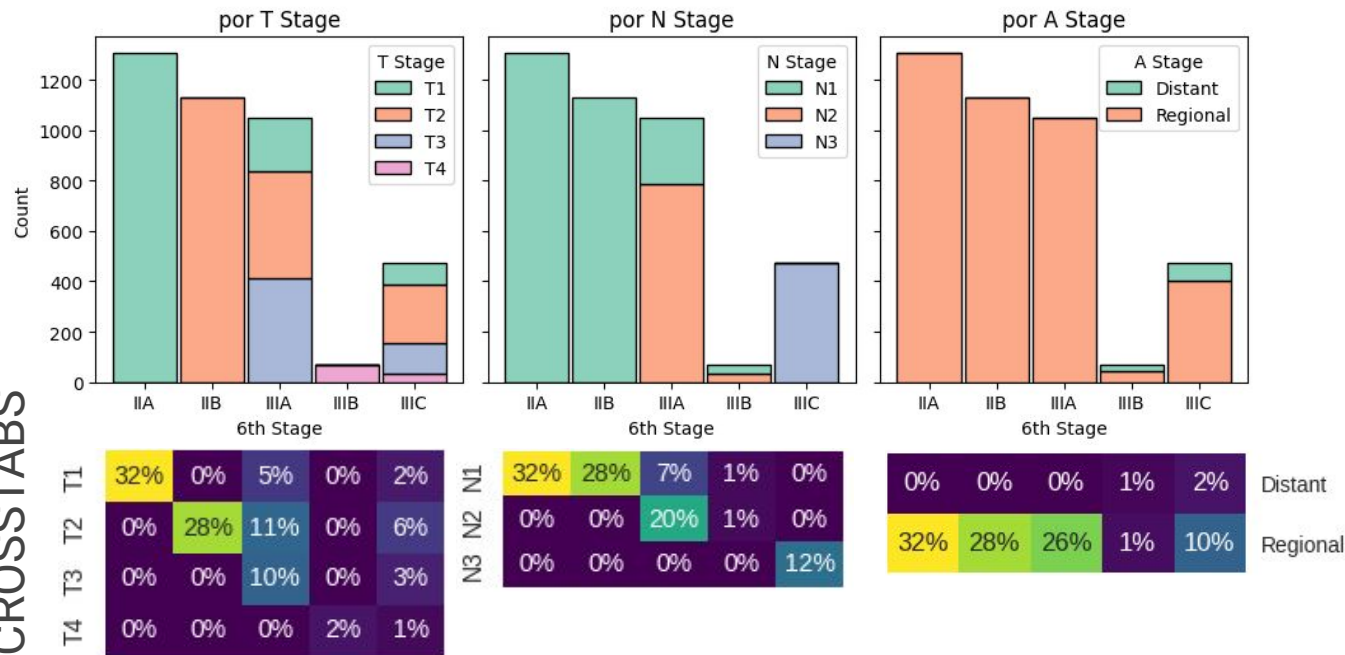
Distribucion de 6th Stage



4- Análisis Exploratorio

T Stage - N Stage - A Stage

Distribucion de 6th Stage



¿Cómo se distribuyen las categorías de las variables T, N y A Stage en las de 6th Stage?

Las categorías de mayor gravedad corresponden a los 6th Stages mayores, algunos no tan estrictamente

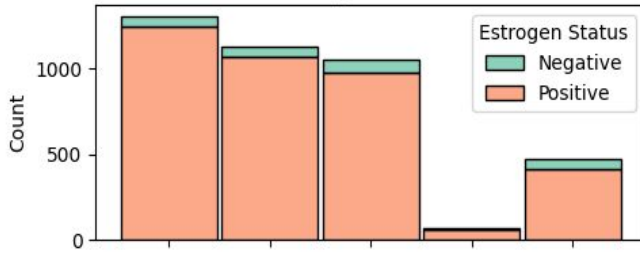
Las tablas de contingencia muestran el peso de cada categoría de variables T, N, A Stage para cada categoría de 6th Stage

4- Análisis Exploratorio

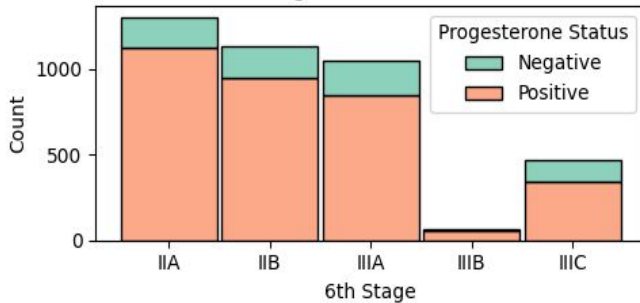
Estrogen Status - Progesterone Status

Distribucion de 6th Stage

Estrogen Status



Progesterone Status



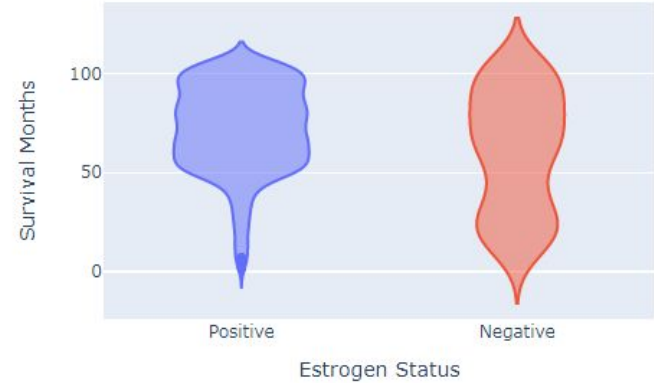
		Progesterone Status	
		Negative	Positive
Estrogen Status	Negative	242	27
	Positive	456	3299

Matriz de contingencia

En todos los 6th Stage, predominan los registros con receptores hormonales positivos.

La proporción de los receptores negativos es mayor en los 6th Stage más graves IIIB y IIIC.

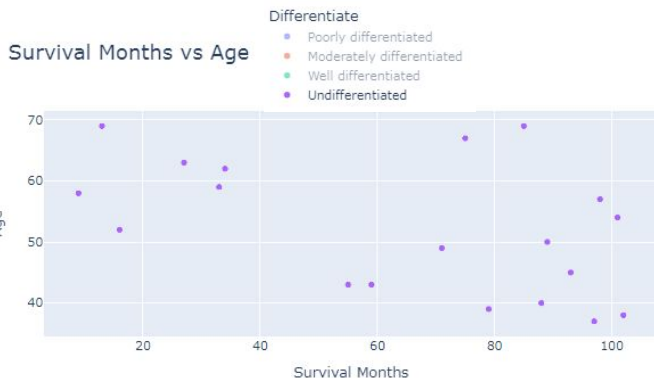
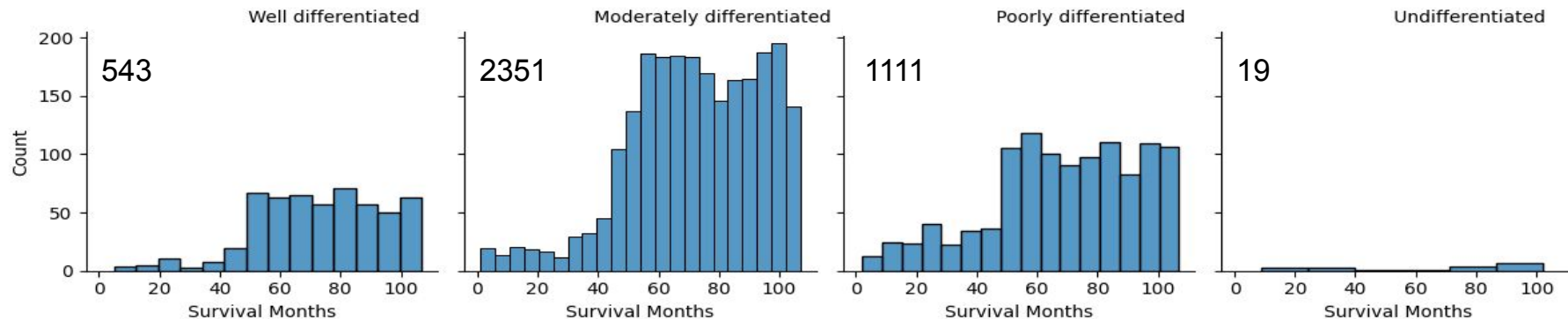
Survival Months - Estrogen Status



Los casos con receptores hormonales negativos, especialmente los estrogénicos, se relacionan más con sobrevividas menores a 50 meses.

4- Análisis Exploratorio

Differentiate



Al filtrar por Undifferentiated el scatterplot de “Survival Months vs Age”, llama la atención como los pocos casos se distribuyen pareciera ser hacia las esquinas de mayor edad menor sobrevida y menor edad mayor sobrevida.

Para el resto de las categorías de Differentiate, los casos no parecen seguir ningún patron con respecto a la edad.

5- Insights y sugerencias

- Los meses de supervivencia van de 1 a 107 meses. A partir de los 46 meses aumentan la cantidad de registros. Se ven 1203 registros de supervivencia menor a 60 meses (29.9% de la muestra), y 2821 registros que sobrevivieron más de 60 meses: 70.1% de la muestra. La variable objetivo "**Sobrevida**" se encuentra desbalanceada.
- Los registros con **6th stage** IIA, IIB y IIA, en conjunto representan el 86.6% de la muestra.
- Los 6th Stage de mayor gravedad (IIIB y IIIC), 13.4% de la muestra, tienen mayor proporción de registros con menores sobrevidas.
- Los registros con **receptores estrogénicos negativos** representan una proporción en supervivencias menores a 50 meses mayor que los positivos. Dicha tendencia se repite en el caso de los receptores de progesterona aunque aquí menos marcado. La proporción de receptores hormonales negativos es mayor en 6th Stages mayores (IIIB y IIIC).
- 82% de los registros tienen ambos **receptores hormonales positivos**.
- La gran mayoría de los casos son de grado de diferenciación moderada (2351).
- La **edad** de las pacientes parece ser mayor en los casos de receptores progéstágenos negativos y menor en los de receptores estrogénicos negativos

to be continued...