

UNIVERSIDAD NACIONAL DE ROSARIO  
Facultad de Ciencias Económicas y Estadística



# Análisis de datos de duración en pacientes con cáncer de mama

Rotterdam tumor bank - 1978-1985

*Alumnas:* Agustina Mac Kay y Rocio Canteros

Año 2024

## Introducción

El cáncer de mama es un tipo de cáncer primario, que se origina en la mama y puede propagarse a otros tejidos u órganos del cuerpo. Es el tipo de cáncer más frecuente y la causa más común de muerte por cáncer en mujeres a nivel mundial.<sup>1</sup>

En este estudio se trabajará con información acerca de 583 mujeres que fueron sometidas, entre 1978 y 1985, a una cirugía primaria para extirpar el tumor.

Los datos fueron obtenidos de la base *rotterdam* del paquete *survival* de R. La misma cuenta con el tiempo desde la cirugía hasta la muerte o pérdida de seguimiento de las pacientes, junto a otras covariables basales que se detallan a continuación:

- **Age:** edad al momento de la cirugía (en años).
- **Meno:** estado menopáusico, donde 0 = premenopáusico y 1 = postmenopáusico.
- **Hormon:** variable indicadora de haber recibido un tratamiento hormonal.
- **Chemo:** variable indicadora de haber recibido quimioterapia.
- **Pgr:** receptores de progesterona (en fmol/l).
- **Er:** receptores de estrógeno (en fmol/l).
- **Grade:** grado de diferenciación del tumor, con valores de 1 a 3.
- **Size:** tamaño del tumor, con niveles: menos de 20mm, entre 20 y 50mm, 50mm.

De la totalidad de mujeres en estudio, se cuenta con el tiempo exacto hasta la muerte de 377 de ellas y 206 censuras.

## Análisis descriptivo

Para comenzar, se hará un breve análisis descriptivo de los datos.

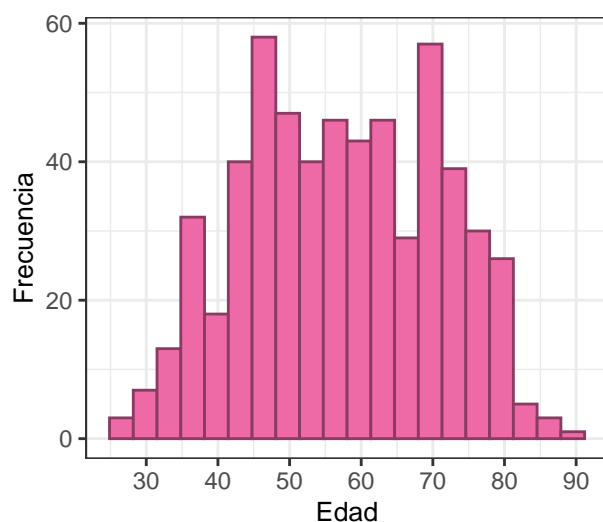


Gráfico 1: Edad de las pacientes

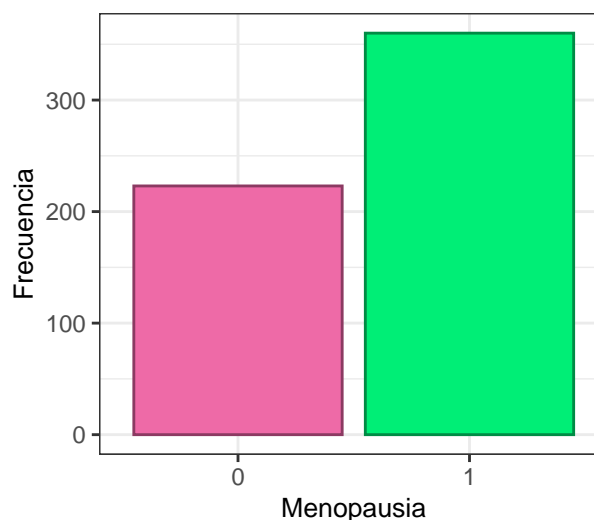


Gráfico 2: Mujeres pre y post menopáusicas

El gráfico 1 muestra que la base de datos cuenta con mujeres entre 27 y 90 años de edad, teniendo la mayoría entre 40 y 80 años.

Para continuar el análisis con esta variable, se la considerará como un factor con las siguientes categorías: mujeres menores de 40, entre 40 y 49, entre 50 y 69, y mayores de 70 años.

<sup>1</sup>Fuente: [Organización Panamericana de la Salud](#)

En el gráfico 2 se ve que la cantidad de mujeres post menopáusicas supera en, aproximadamente, un 50% a las mujeres pre menopáusicas.

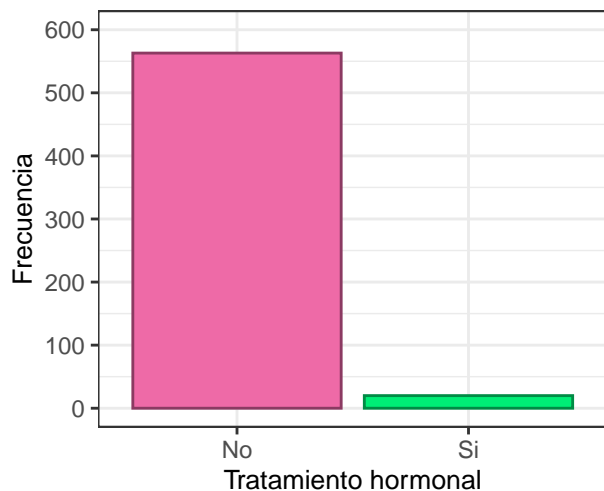


Gráfico 3: cantidad de mujeres que recibieron o no tratamiento hormonal

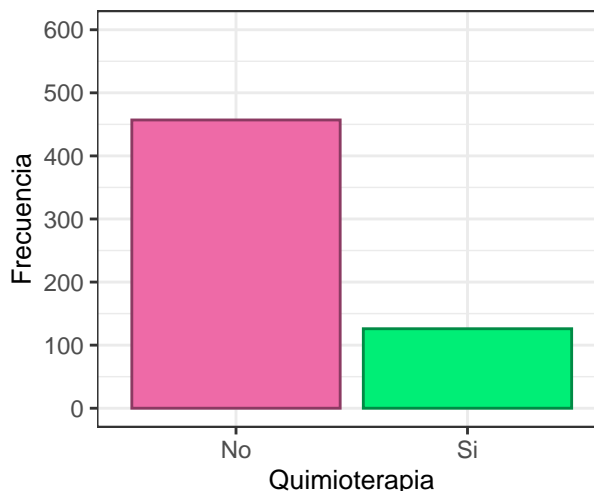


Gráfico 4: Cantidad de mujeres que recibieron o no quimioterapia

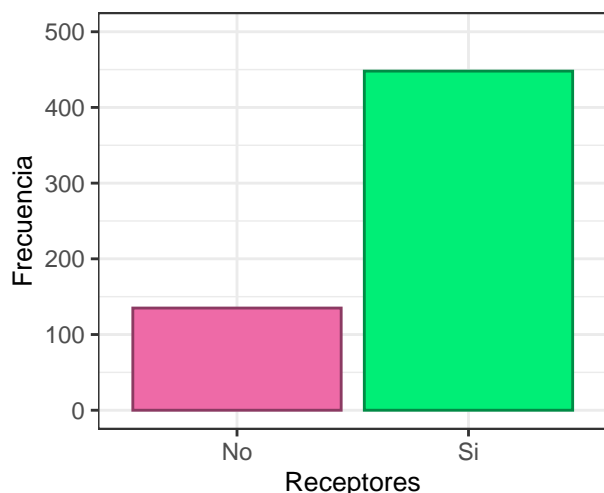


Gráfico 5: Cantidad de mujeres que presentan o no receptores de progesterona

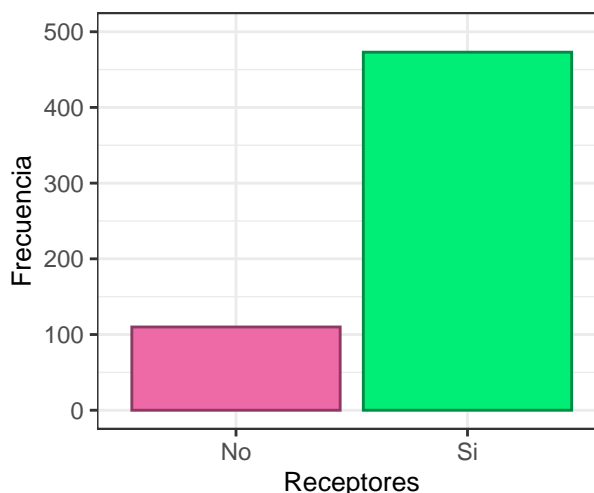


Gráfico 6: Cantidad de mujeres que presentan o no receptores de estrógeno

En los gráficos 5 y 6 se muestra que aproximadamente 150 mujeres no presentan receptores de progesterona, mientras que apenas más de 100 no presentan receptores de estrógenos. Ambos casos se traducen en la no recomendación de un tratamiento hormonal. Resulta interesante notar que si bien muchas mujeres poseen receptores, muy pocas recibieron un tratamiento hormonal (gráfico 3).

Sería interesante evaluar la relación entre la presencia de receptores y el tratamiento hormonal, pero no se cuenta con la cantidad de datos suficientes para obtener conclusiones válidas al respecto.

Además, en el gráfico 4 se observa que del total de mujeres bajo estudio, aproximadamente 455 de ellas no recibieron quimioterapia previo a la realización de la cirugía.

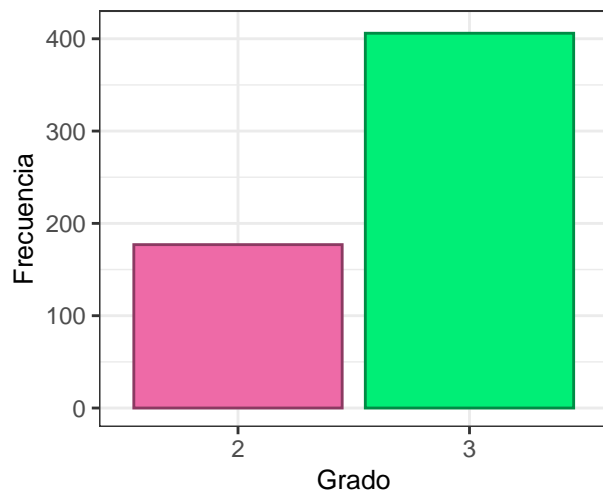


Gráfico 7: Cantidad de mujeres por grado de diferenciación

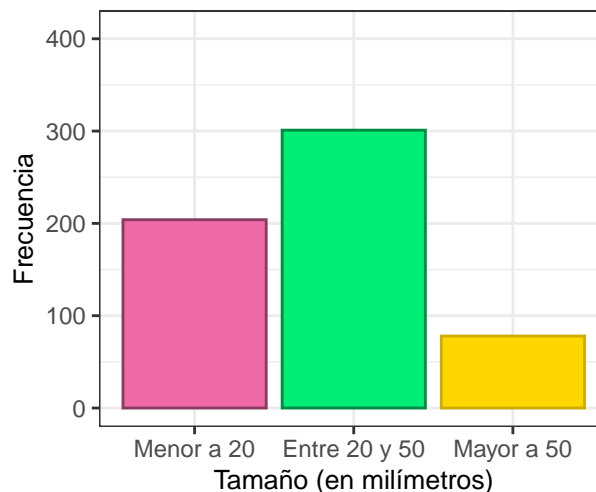


Gráfico 8: Cantidad de mujeres por tamaño del tumor

Se observa, en el gráfico 7, que 400 mujeres de las enroladas presentaron un grado de diferenciación de nivel 3 y ninguna presentó un tumor de grado 1.

El grado del tumor describe qué tan normal o anormal se ven las células cancerosas cuando se observan con un microscopio. Mientras más alto es el grado, más anormal se ven las células y más alta es la probabilidad de que el tumor crezca y se disemine rápido.<sup>2</sup>

La mayoría de las mujeres presentaron, al momento de la cirugía, un tumor de entre 20 y 50mm. Alrededor de 200 mujeres tenían un tumor con menos de 20mm, y aproximadamente 80 tenían un tumor con más de 50mm. (gráfico 8)

A continuación se muestra la curva de supervivencia estimada por Kaplan-Meier para todos los datos:

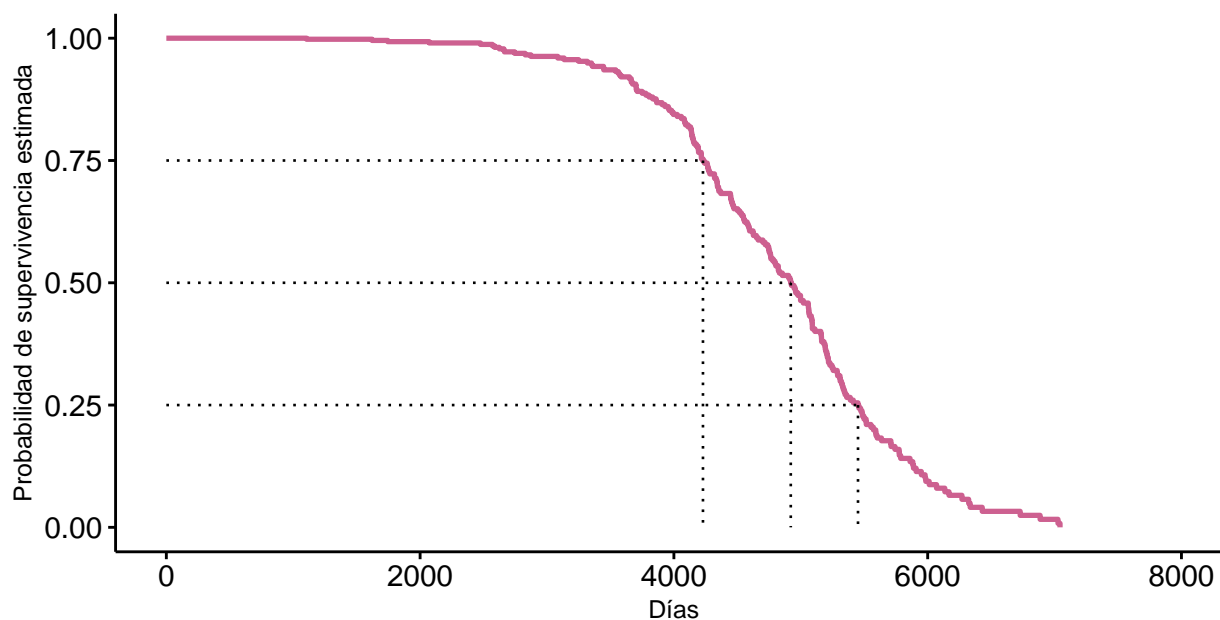


Gráfico 9: Función de supervivencia estimada por KM

<sup>2</sup>Fuente: [National Cancer Institute of the United States](https://www.seer.cancer.gov/)

Se puede ver que el 50% de las mujeres que están bajo estudio sobrevivirían aproximadamente hasta los 5000 días (los intervalos de días van de 2000 en 2000 y se puede considerar que la línea de la mediana estaría cercana a la mitad del intervalo [4000, 6000]), mientras que al menos 25% de ellas sobreviviría más de 4000 días y al menos el 75% de ellas sobreviviría aproximadamente 6000 días.

## Test de hipótesis

A continuación se realizan test de hipótesis que permiten evaluar si existen diferencias en la experiencia de supervivencia de las categorías de cada covariable.

Variable	Test	p_value
Rango de edad	Wilcoxon generalizado	< 0.0001
Menopausia	Wilcoxon generalizado	0.01
Tratamiento hormonal	Peto-Peto	0.40
Quimioterapia	Wilcoxon generalizado	0.22
Receptores de progesterona	Log-Rank	0.77
Receptores de estrógeno	Log-Rank	0.06
Grado de diferenciación	Log-Rank	0.44
Tamaño	Log-Rank	0.62

- Para Tamaño, Grado de diferenciación y ambos receptores se usaron test Log-Rank
- Rango de edad, Menopausia y Quimioterapia: Se aplicó el test de Wilcoxon generalizado, variante del test Log-Rank cuyos pesos a asignar son definidos como  $n_i$  ya que se tienen grupos desbalanceados.
- Tratamiento hormonal: En el gráfico 14 se observa que las funciones de supervivencia no son similares y se sabe que son grupos demasiado desbalanceados, se decide aplicar el test de Peto-Peto modificado que utiliza en su estadística dichos valores

## Curvas de supervivencia

Considerando los test de hipótesis realizados, se presentan las curvas de supervivencia de aquellos cuyo resultado condujo a \**“Rechazar”  $H_0$* \*.

- Edad:

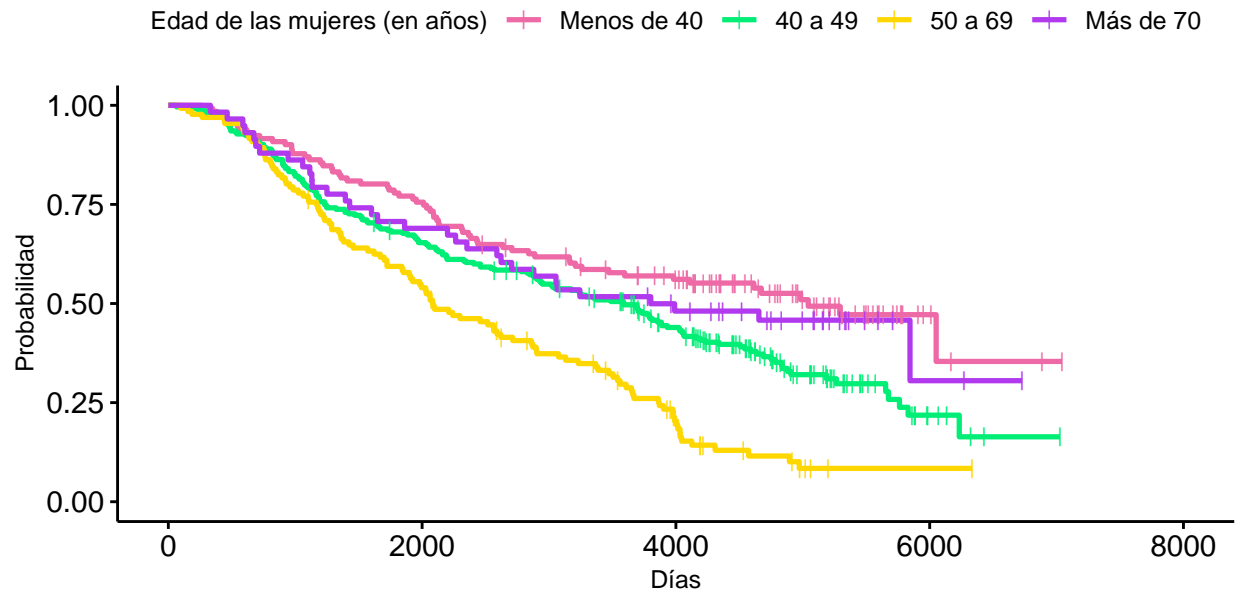


Gráfico 10: Probabilidad de supervivencia estimada según edad

Se puede concluir que la experiencia de supervivencia se ve influenciada por el grupo etario al cuál pertenecen las mujeres.

- Menopausia:

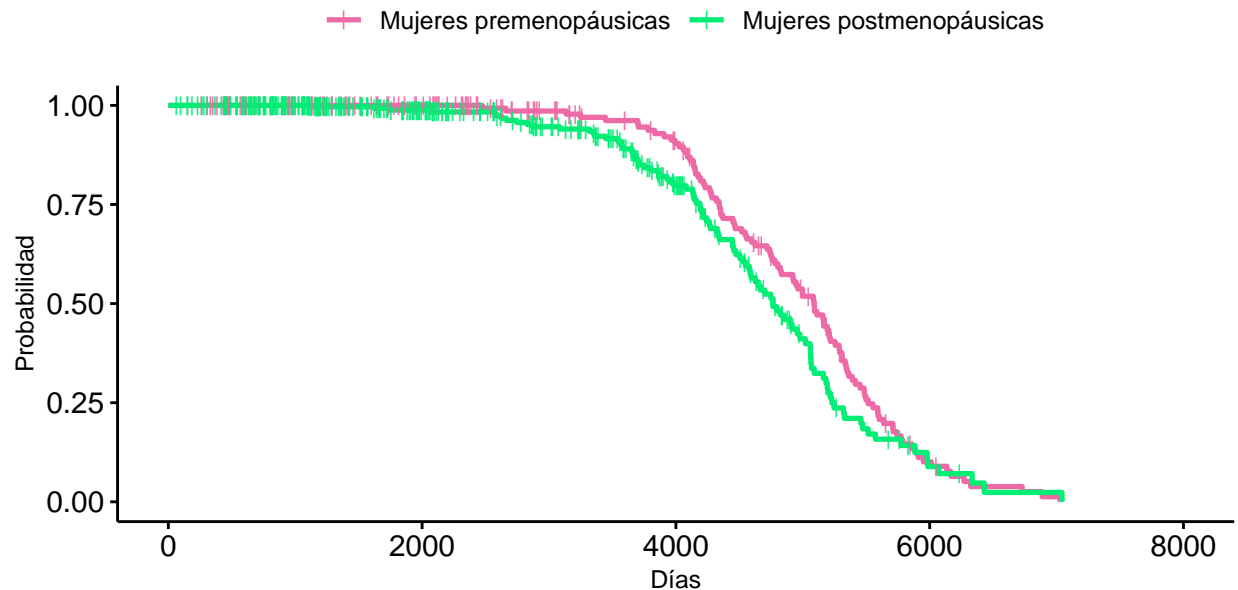


Gráfico 11: Probabilidad de supervivencia estimada según menopausia

Se puede decir que las experiencias de supervivencia son diferentes entre ambos grupos.

- Receptores de estrógeno:

Si bien la probabilidad asociada a la estadística del test resultó ser mayor al nivel de significación con el que se está trabanando (5%), se muestra igual su curva de supervivencia ya que la difernecia entre ambos valores no se la considera como importante.

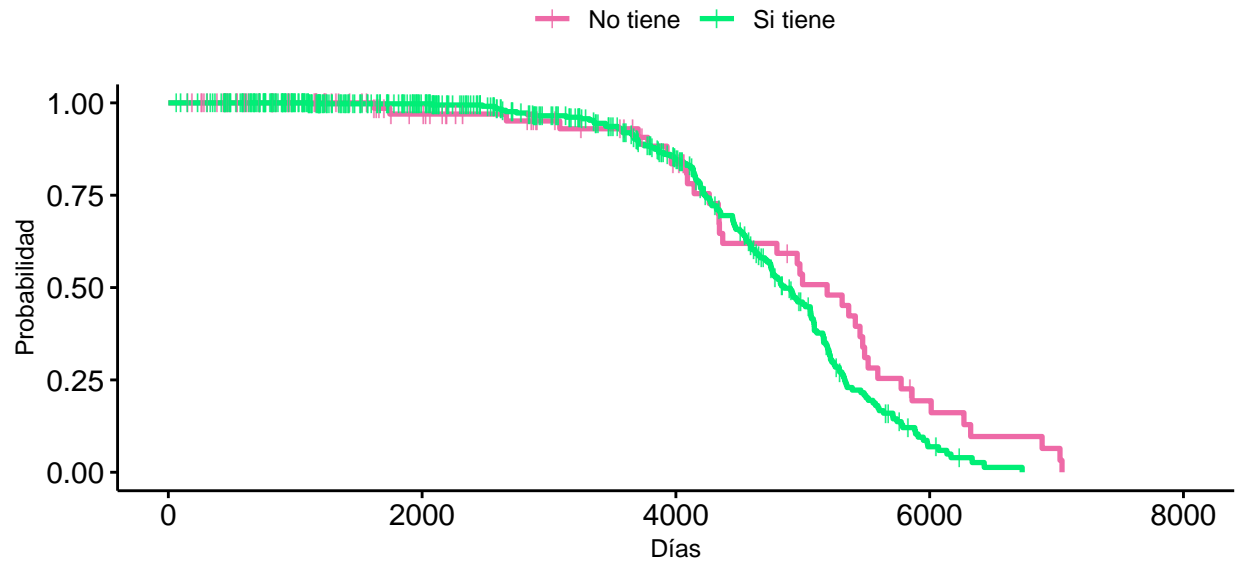


Gráfico 12: Probabilidad de supervivencia estimada según presencia o ausencia receptores de estrógeno

Se puede considerar que la experiencia de supervivencia no es la misma para ambos grupos.

Finalmente, resultó de interés ver la relación entre el tamaño del tumor de la paciente y si recibió o no quimioterapia, por ello se decidió realizar el Test Log-Rank Estratificado. El mismo condujo a “*No rechazar  $H_0$* ” con lo que se puede concluir que el efecto de la quimioterapia es el mismo para cualquier tamaño de tumor. Pese a ello, se muestra la curva de supervivencia para tener una mejor idea de la decisión tomada.

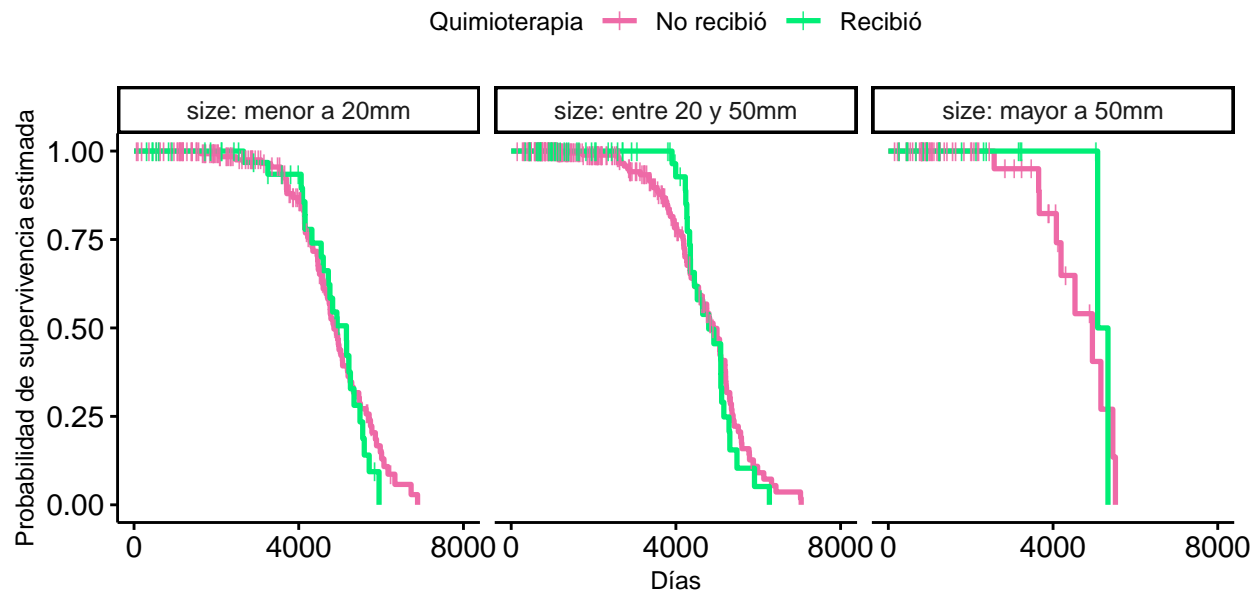


Gráfico 13