



# Cáncer de páncreas

**Análisis clínico y genómico, y modelización a partir  
de los datos del TGCA**

**Master universitario en ciencia de datos**

**Carlos Gatón Rodriguez**

# Contexto

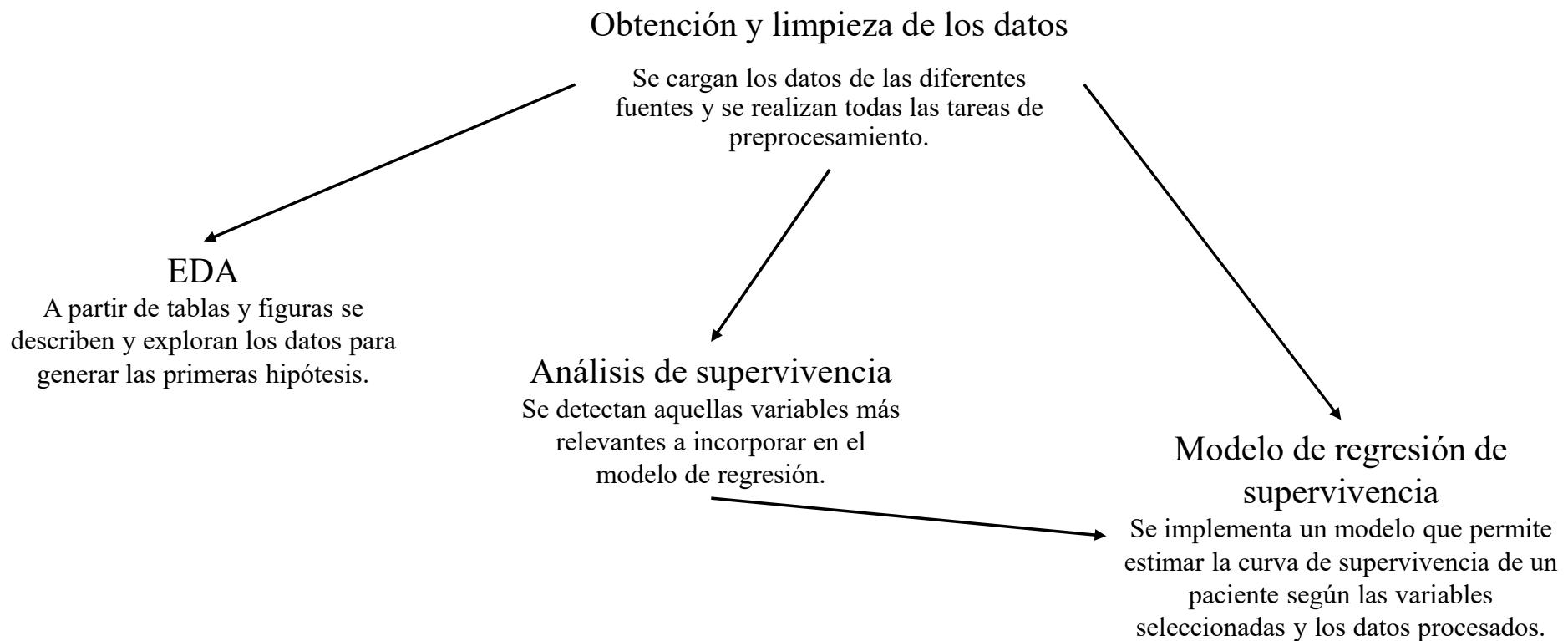
- El cáncer son un conjunto de enfermedades con una incidencia en aumento.
- El cáncer de páncreas tiene una tasa de supervivencia muy baja.
- El TCGA es un proyecto de caracterización molecular. Contiene 20000 muestras de 33 tipos de cáncer.
- Contiene datos genómicos, epigenéticos, transcriptómicos y proteómicos.

# Objetivos

Dos tipologías de objetivos:

- Por un lado, objetivos relacionados con la ciencia de datos. Se pretende extraer y procesar los datos para, a continuación, describirlos y visualizarlos.
- Por otro lado, se tienen objetivos relacionados con la investigación científica. Se realizará un análisis completo de supervivencia y se implementará un modelo de regresión.

# Materiales y métodos

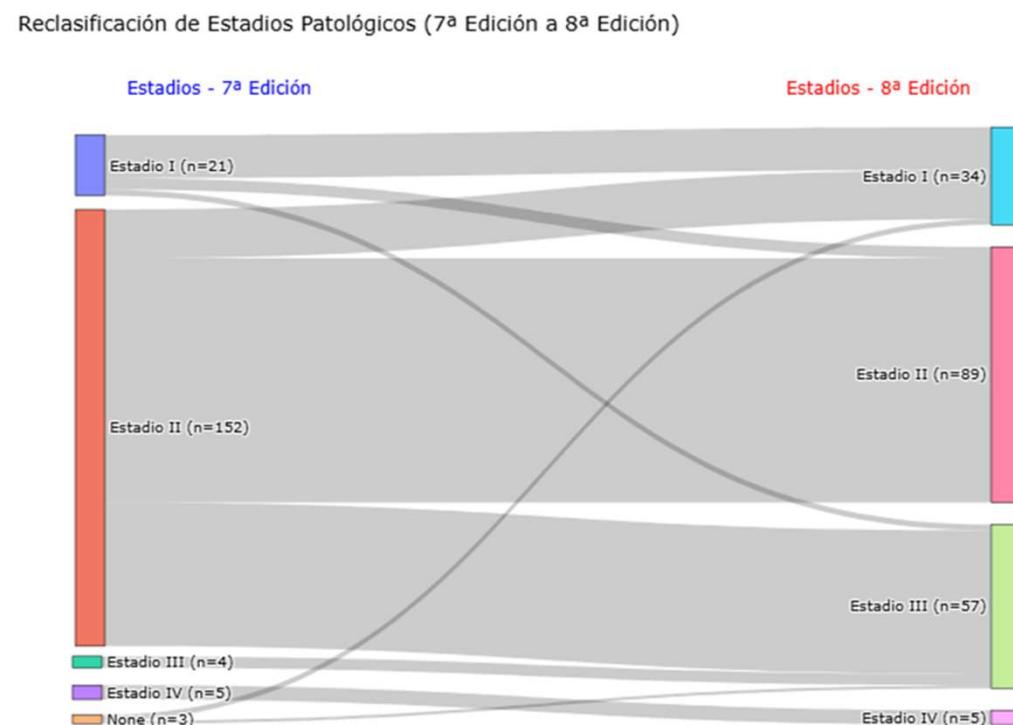


# Datos demográficos

n		185
Edad, mediana [min, max]		65 [35, 88]
Género, n (%)	Masculino	102 (55.1)
	Femenino	83 (44.9)
Etnia, n (%)	No reportado	43 (23.2)
	Hispano o latino	5 (2.7)
	No hispano o latino	137 (74.1)
Raza, n (%)	No reportado	5 (2.7)
	Asiática	11 (5.9)
	Negra o afroamericana	7 (3.8)
	Blanca	162 (87.6)
Estado vital, n (%)	Vivo	113 (61.1)
	Muerto	72 (38.9)

# Reclasificación de los estadios patológicos

- A partir de la clasificación de la 7<sup>a</sup> edición.
- Incorporando datos histopatológicos del número de nódulos positivos y tamaño del tumor.



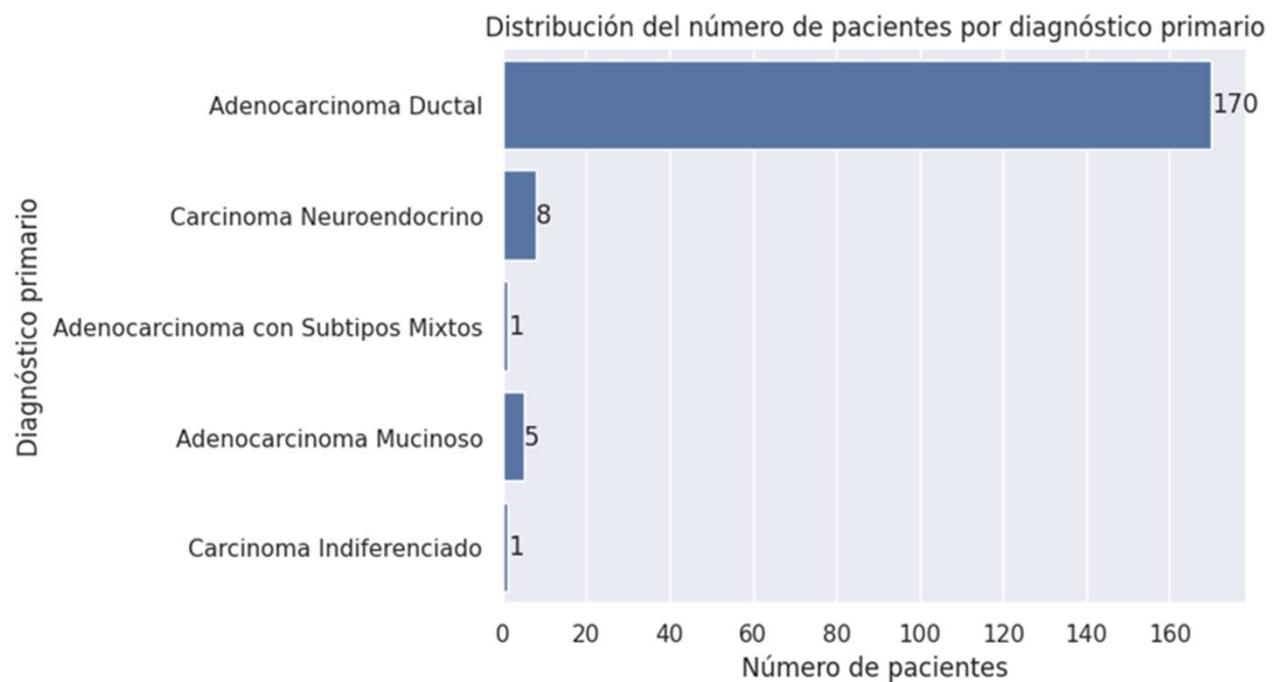
# Margen de resección y grado del tumor

- El margen de resección indica la presencia de células cancerosas tras la extirpación del tumor.
- El grado del tumor describe la apariencia del tumor bajo el microscopio.

Margen de resección, n (%)	R0	111 (60.0)
R1	53 (28.6)	
R2	5 (2.7)	
RX	16 (8.7)	
Grado del tumor, n (%)	G1	32 (17.3)
G2	97 (52.4)	
G3	51 (27.6)	
G4	2 (1.1)	
GX	3 (1.6)	

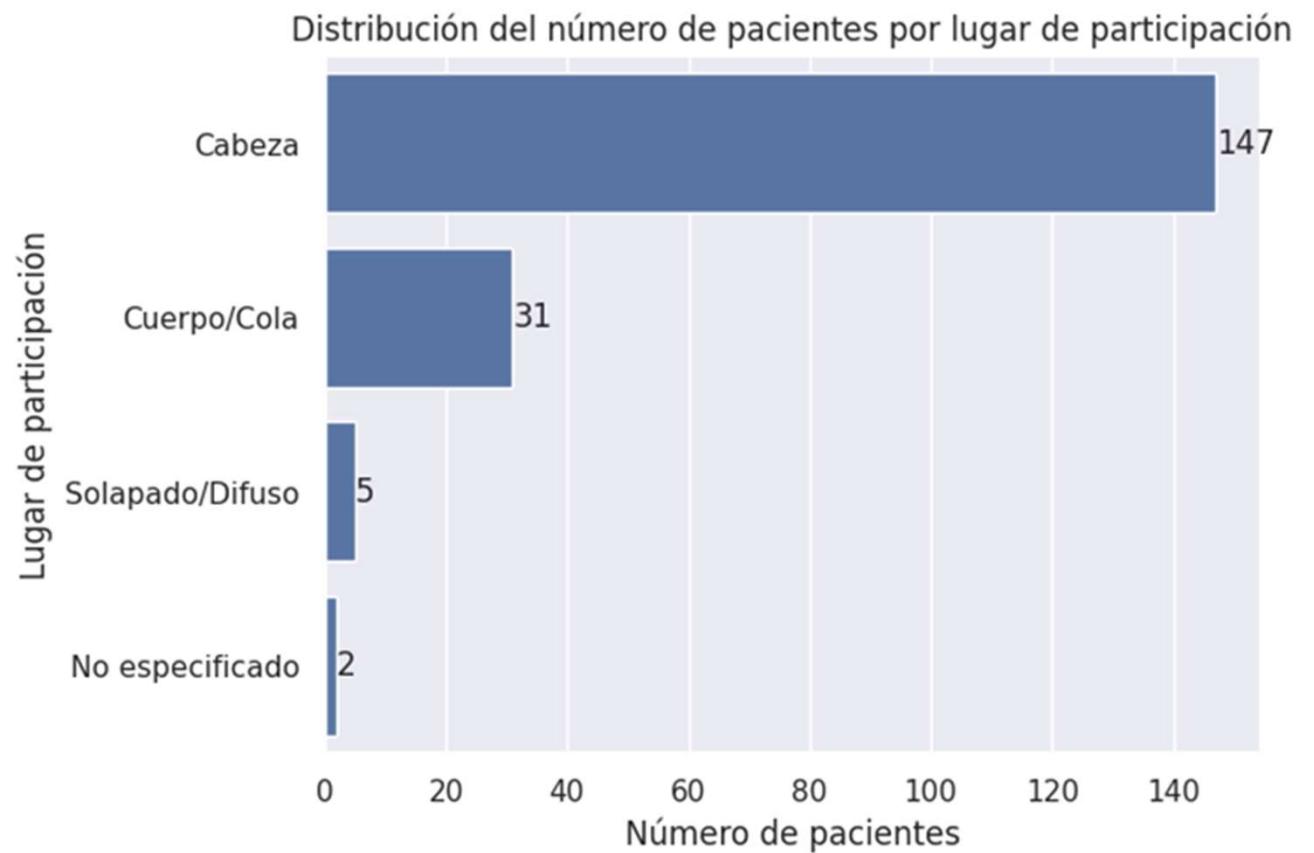
# Diagnóstico primario

- En el análisis de supervivencia solo se incluyen los casos de adenocarcinoma ductal.



# Lugar de participación

- La mayoría de los casos se concentran en la cabeza del páncreas, típicamente con mejor pronóstico.



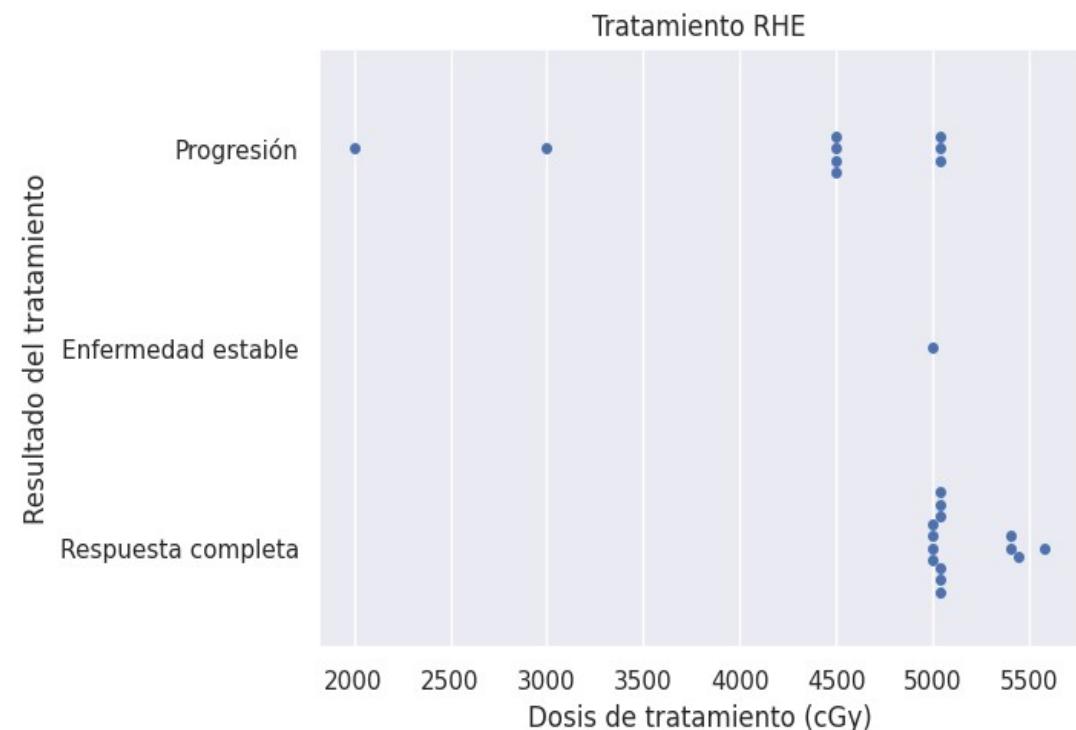
# Metástasis y recurrencia

- Principales lugares de recurrencia son el hígado y tejidos cercanos.
- Principales lugares de metástasis son el hígado y el pulmón.

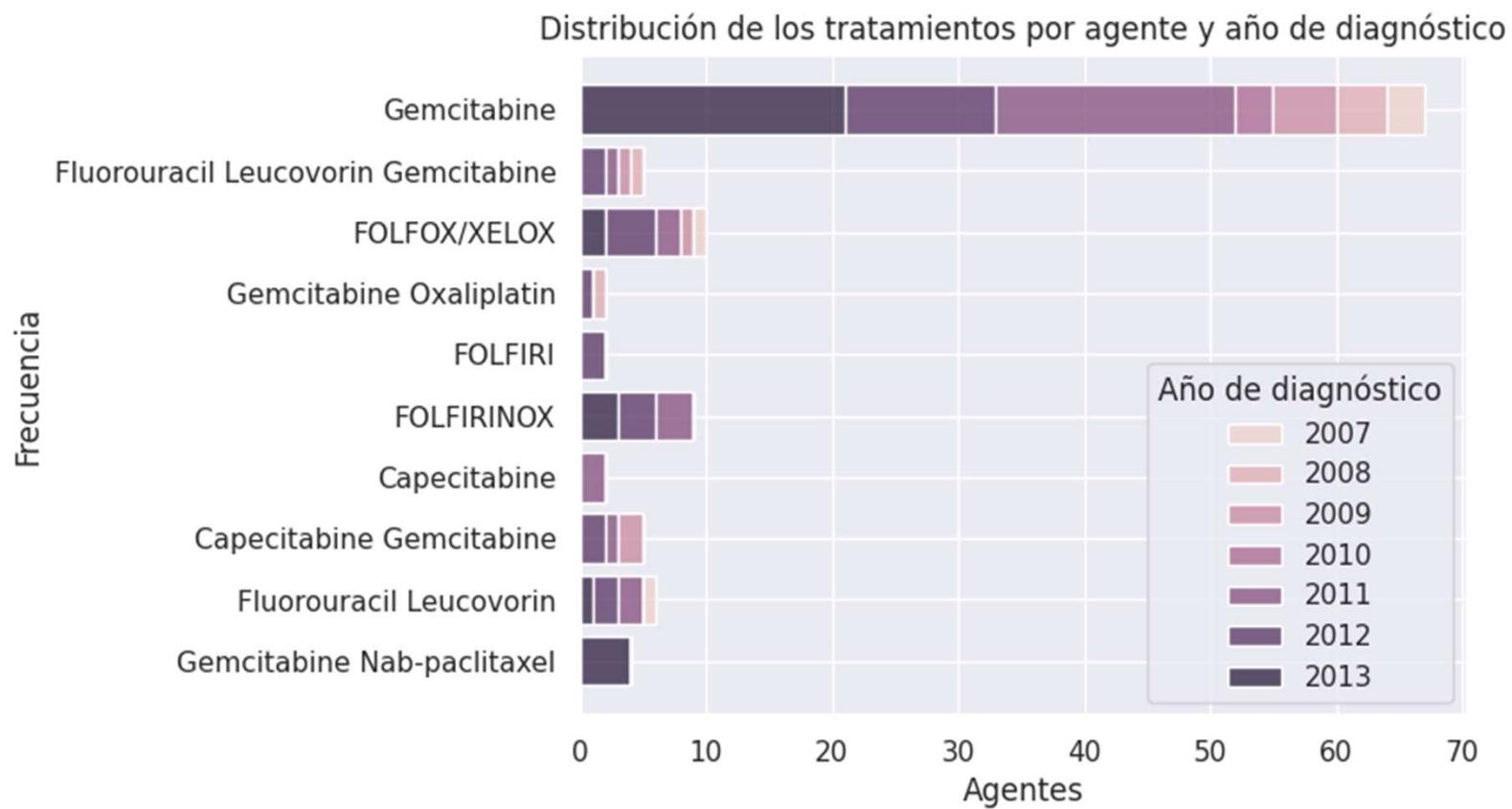
Lugar de cáncer posterior, n (%)	Abdomen	Metástasis	Recurrencia
	Abdomen	1 (1.4)	2 (8.0)
	Glándula suprarrenal	1 (1.4)	0 (0.0)
	Hueso	1 (1.4)	0 (0.0)
	Tejidos blandos	1 (1.4)	0 (0.0)
	Hígado	41 (58.6)	3 (12.0)
	Pulmón	15 (21.4)	0 (0.0)
	Nódulo linfático	3 (4.3)	1 (4.0)
	Páncreas	0 (0.0)	1 (4.0)
	Peritoneo	5 (7.1)	8 (32.0)
	Retroperitoneo	1 (1.4)	1 (4.0)
	No reportado	1 (1.4)	9 (36.0)

# Tratamientos de radioterapia de haz externo

- A través de una prueba de Wilcoxon-Mann-Whitney, se establece que existe una relación estadísticamente significativa entre la dosis total y el resultado del tratamiento.
- Valor de  $p=0.003$ .

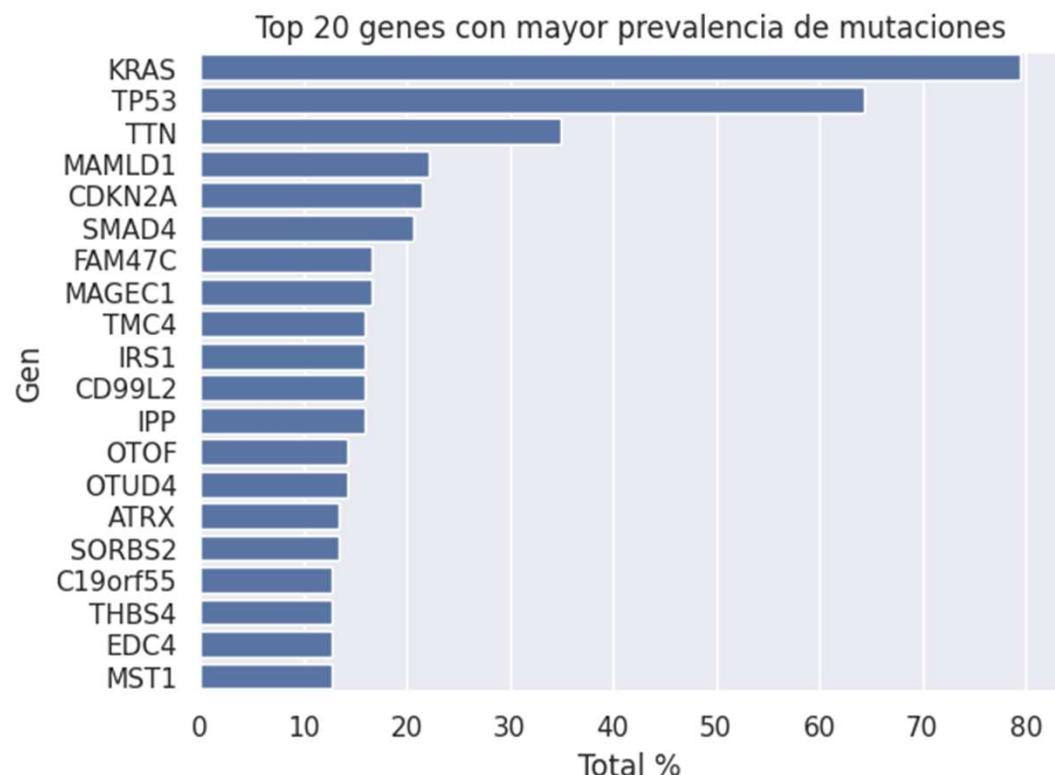


# Tratamientos de quimioterapia



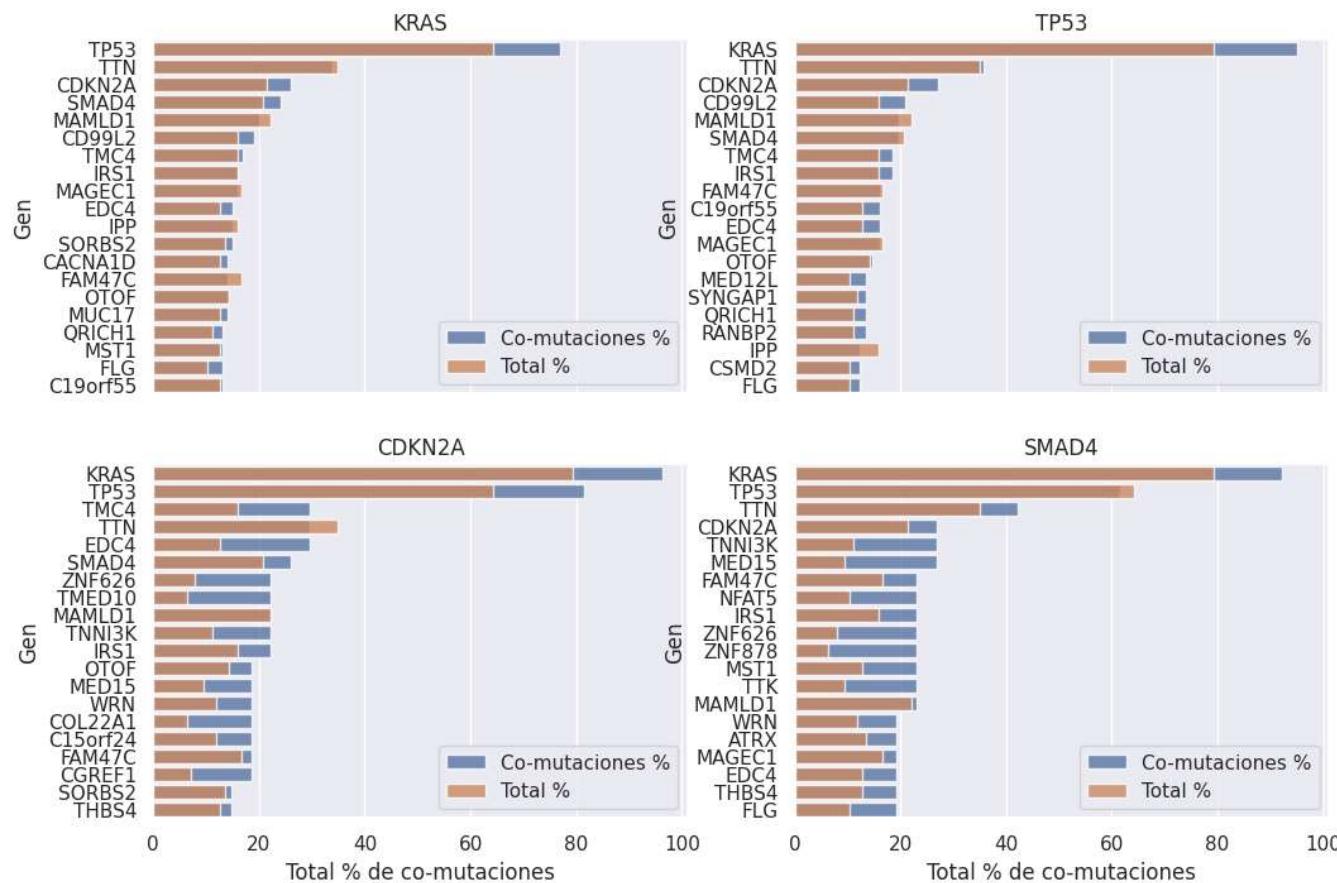
# Mutaciones genéticas

- Se destacan las mutaciones en los genes *KRAS*, *TP53*, *CDKN2A* y *SMAD4*.



# Análisis de las co-mutaciones

Top 20 genes con mayor prevalencia de co-mutaciones

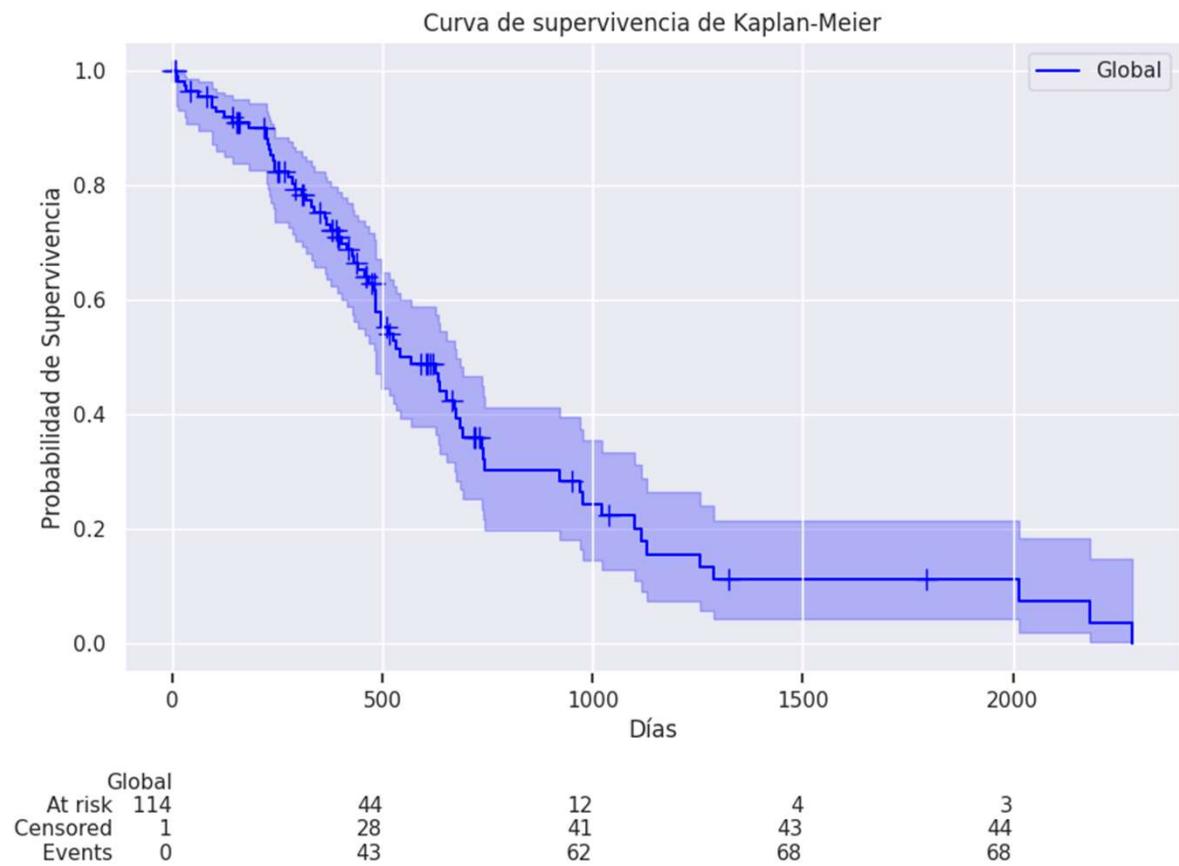


# Análisis de supervivencia (I)

- Se utilizan curvas de supervivencia de Kaplan-Meier para estimar la supervivencia a lo largo del tiempo.
- Se implementan regresiones de Cox para evaluar como las diferentes variables afectan la tasa de riesgo de muerte.

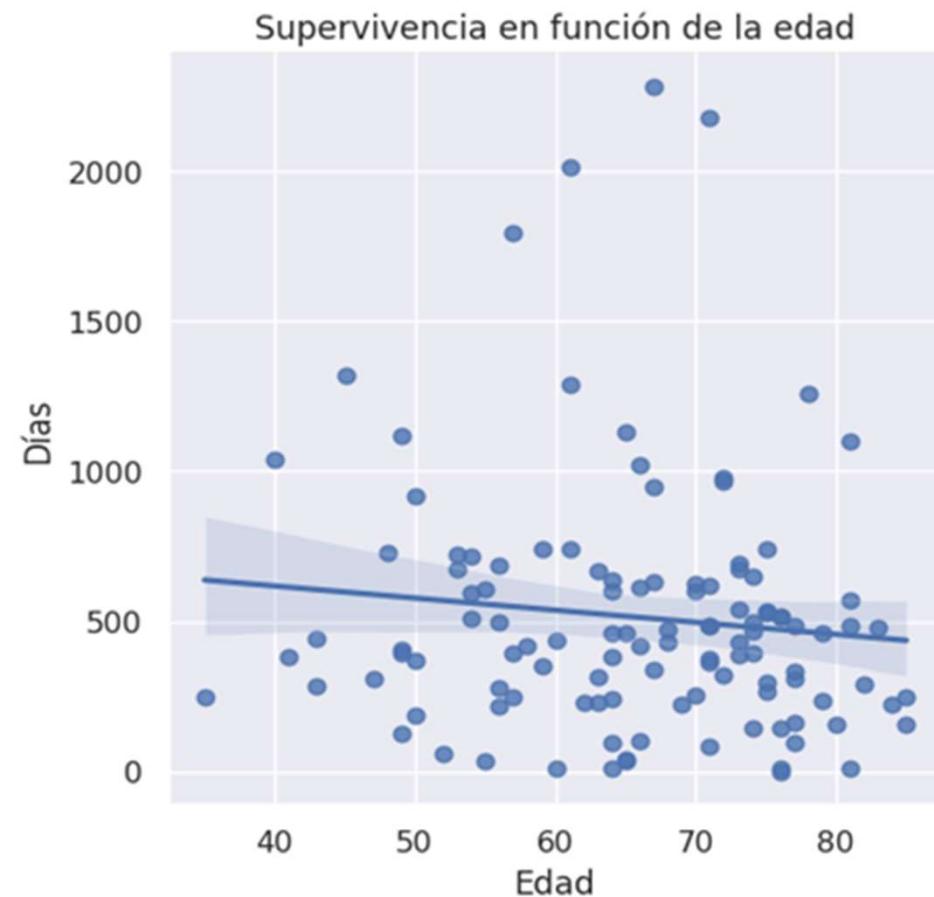
# Análisis de supervivencia (II)

- Se observa una supervivencia media de 568 días.
- Superior a lo obtenido en la población general.



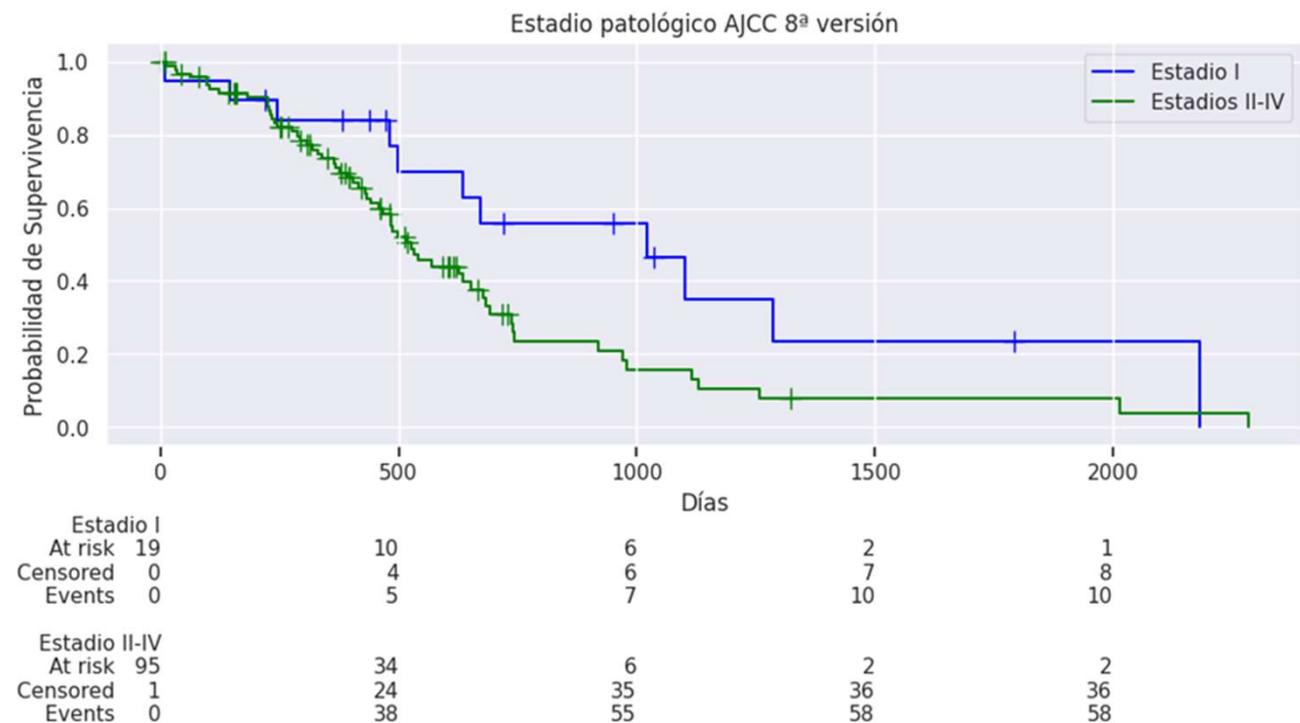
# Análisis de supervivencia (III)

- Se observa una cierta correlación entre la supervivencia y la edad.
- Si se calcula el coeficiente de correlación de Spearman se obtiene un valor de -0,13.



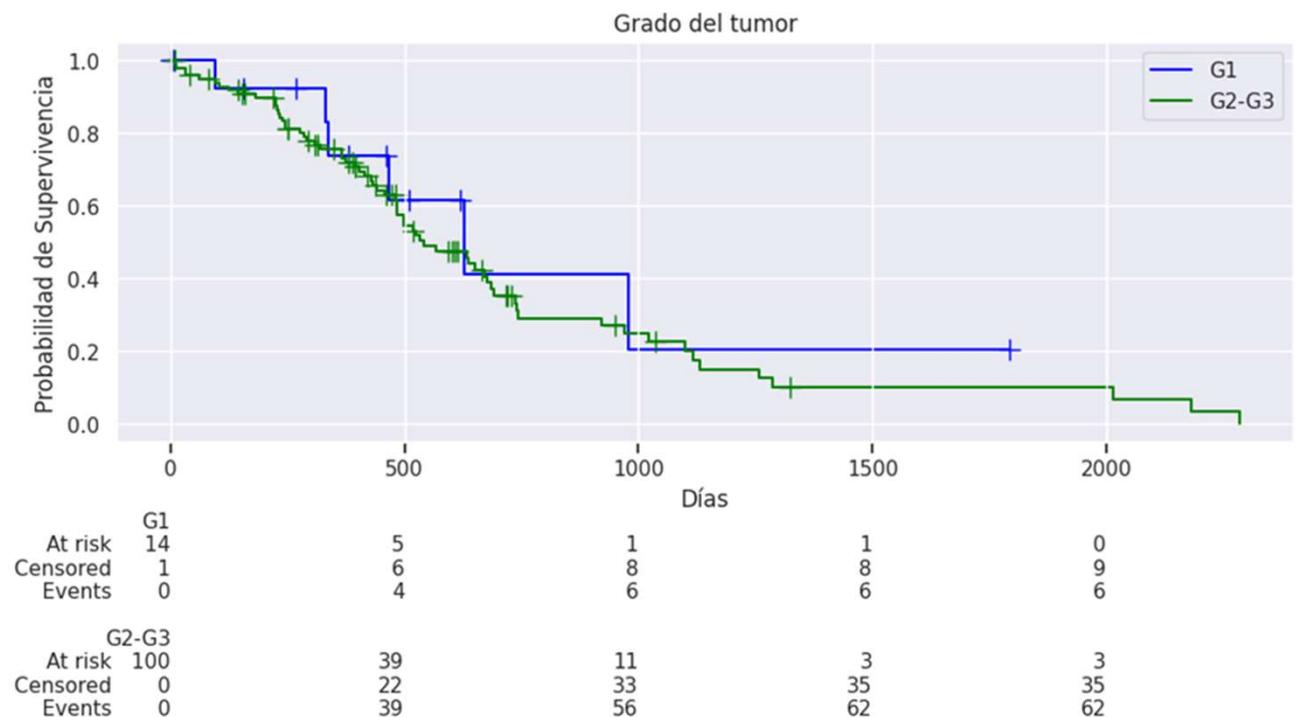
# Análisis de supervivencia (IV)

- Reducción del riesgo del 49%.
- Estadísticamente significativo,  $p=0.05$ .



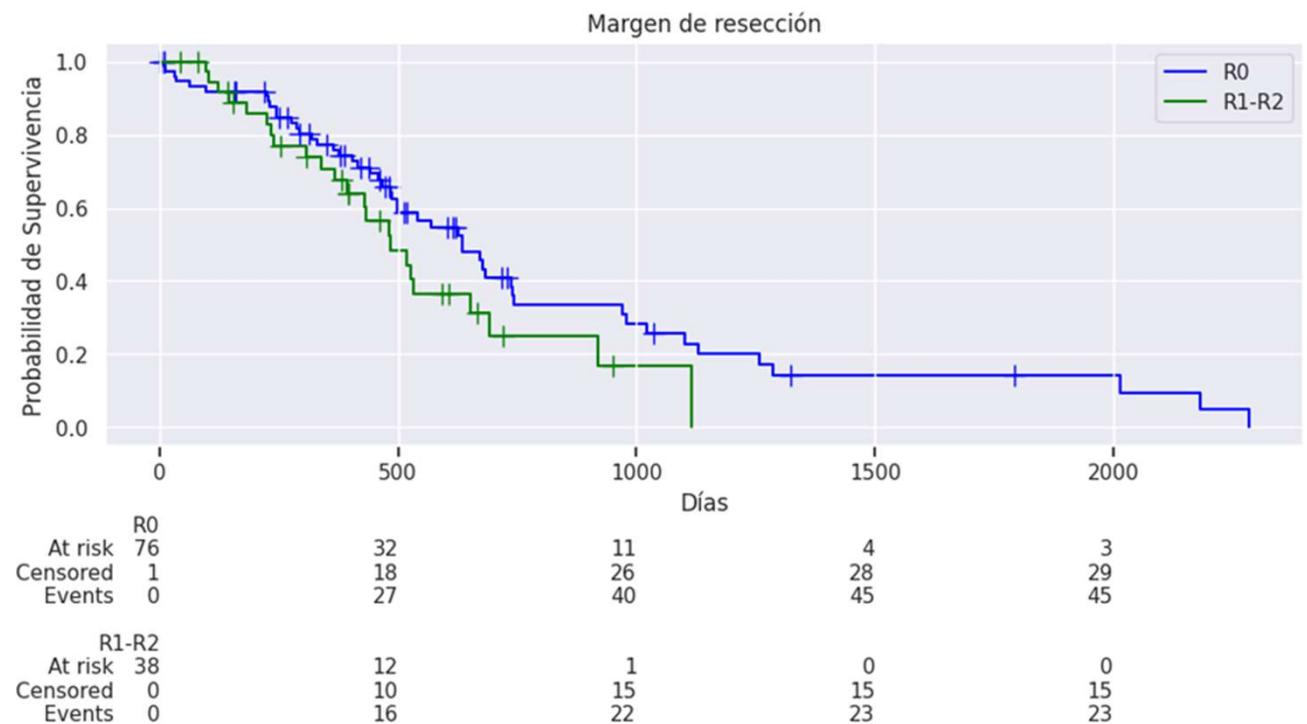
# Análisis de supervivencia (V)

- Reducción del riesgo del 24%.
- No es estadísticamente significativo,  $p=0.52$ .



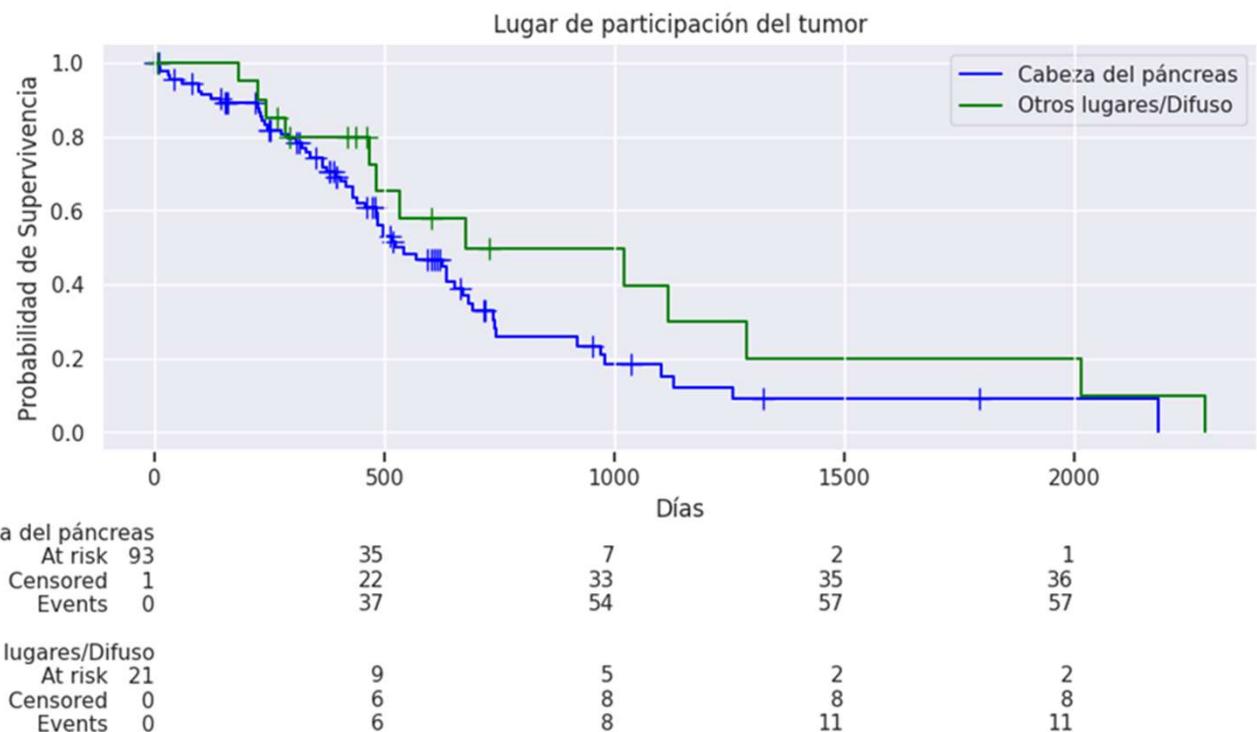
# Análisis de supervivencia (VI)

- Reducción del riesgo del 35%.
- No es estadísticamente significativo,  $p=0.1$ .



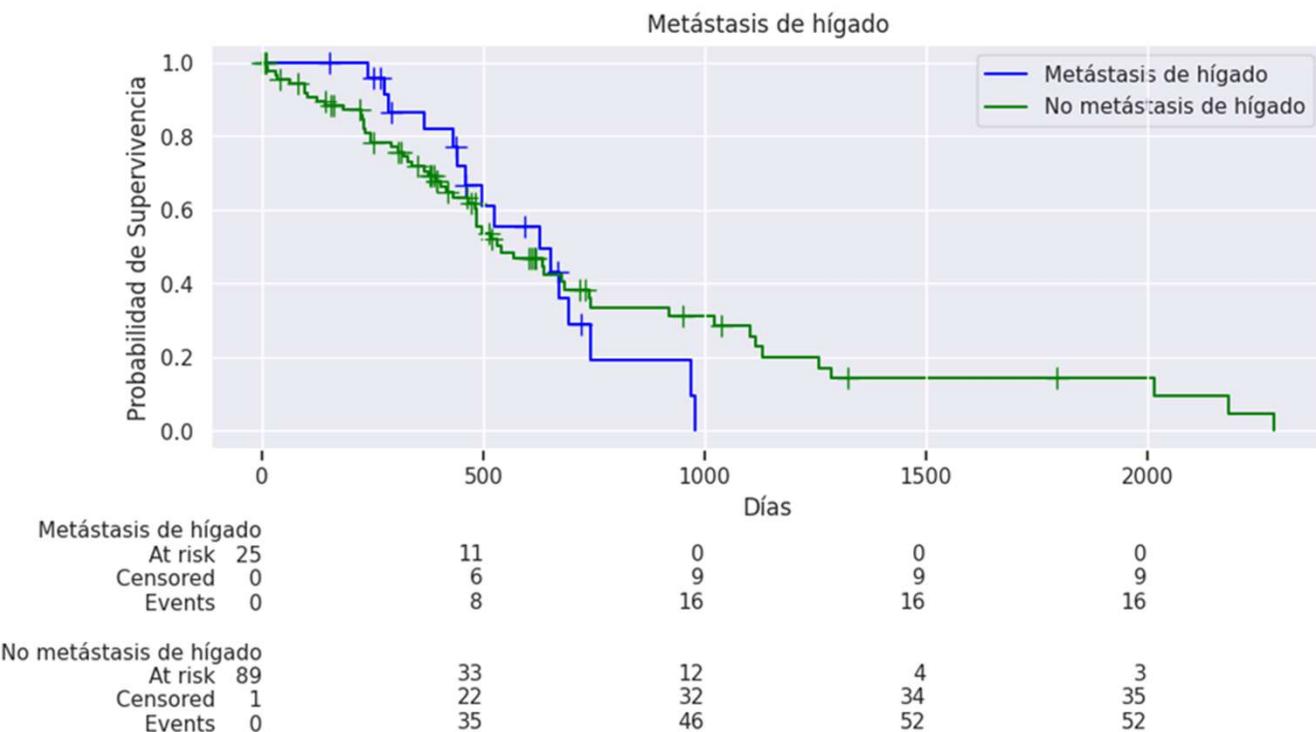
# Análisis de supervivencia (VII)

- Aumento del riesgo del 61%.
- No es estadísticamente significativo,  $p=0.14$ .
- Resultado contrario a lo esperado.



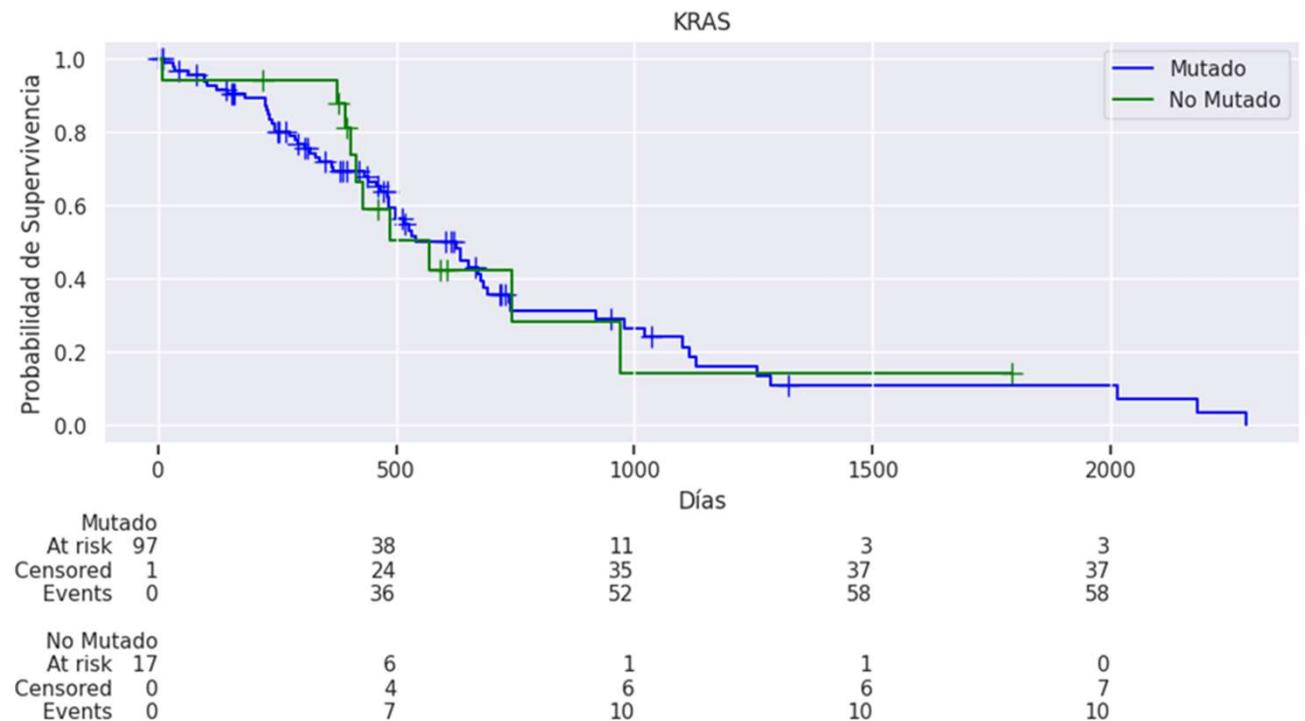
# Análisis de supervivencia (VIII)

- Aumento del riesgo del 11%.
- No es estadísticamente significativo,  $p=0.72$ .



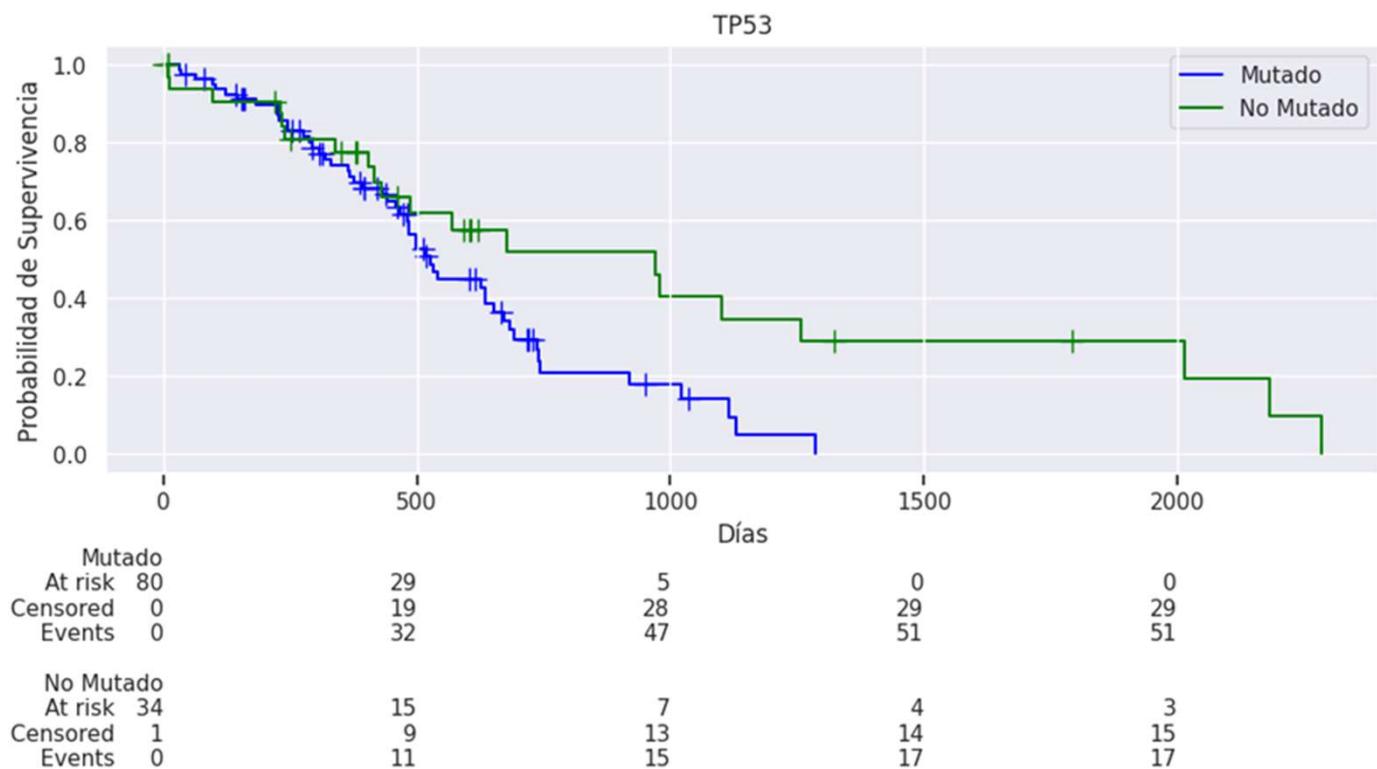
# Análisis de supervivencia (IX)

- Aumento del riesgo del 3%.
- No es estadísticamente significativo,  $p=0.92$ .



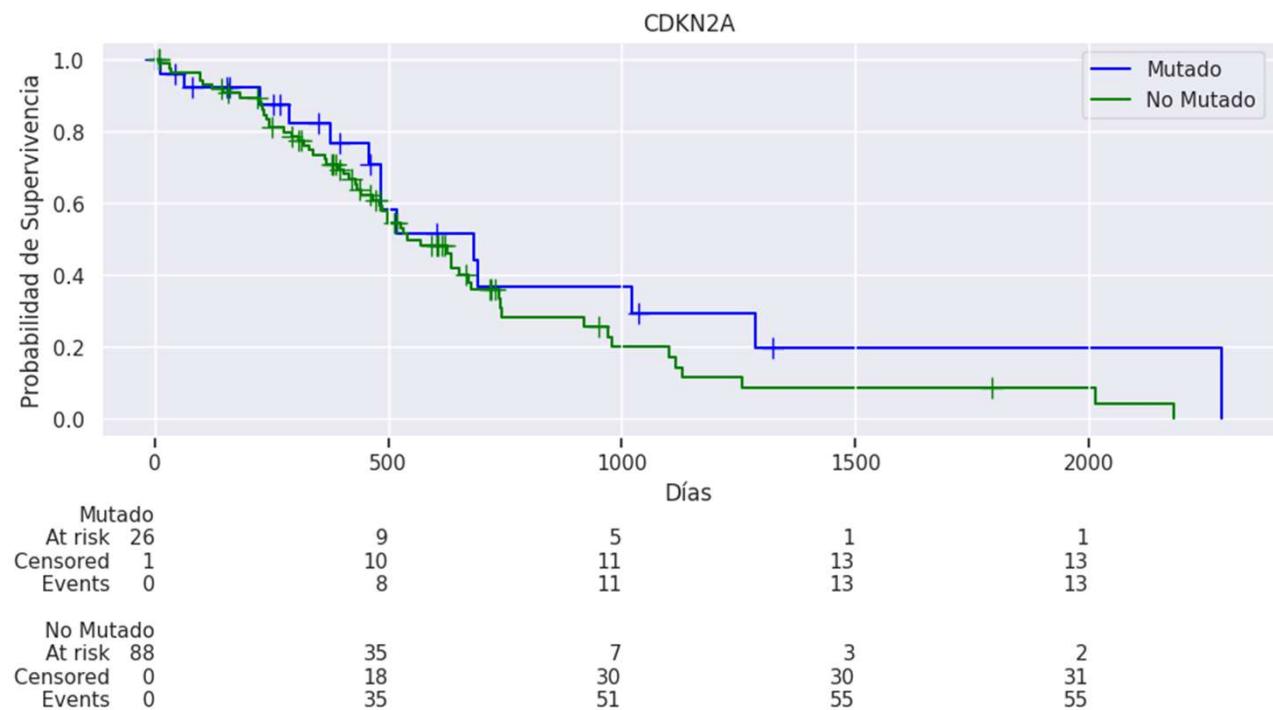
# Análisis de supervivencia (X)

- Aumento del riesgo del 90%.
- Es estadísticamente significativo,  $p=0.03$ .



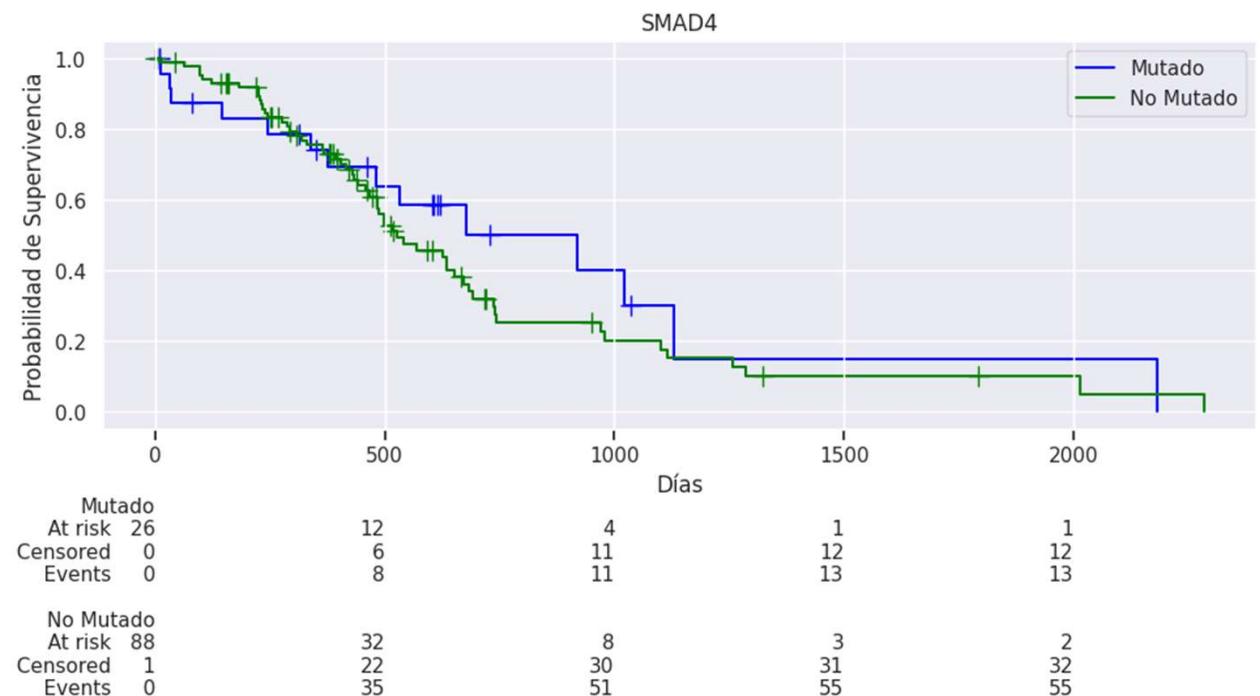
# Análisis de supervivencia (XI)

- Reducción del riesgo del 31%.
- No es estadísticamente significativo,  $p=0.23$ .



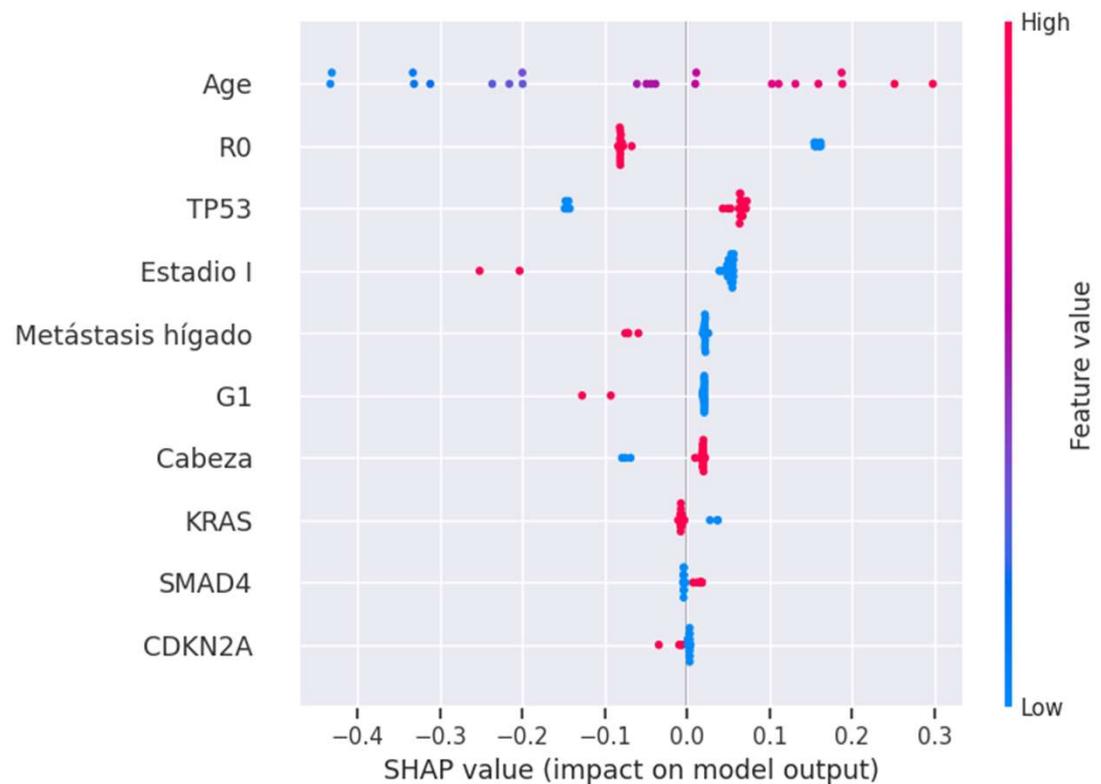
# Análisis de supervivencia (XII)

- Reducción del riesgo del 21%.
- No es estadísticamente significativo,  $p=0.43$ .



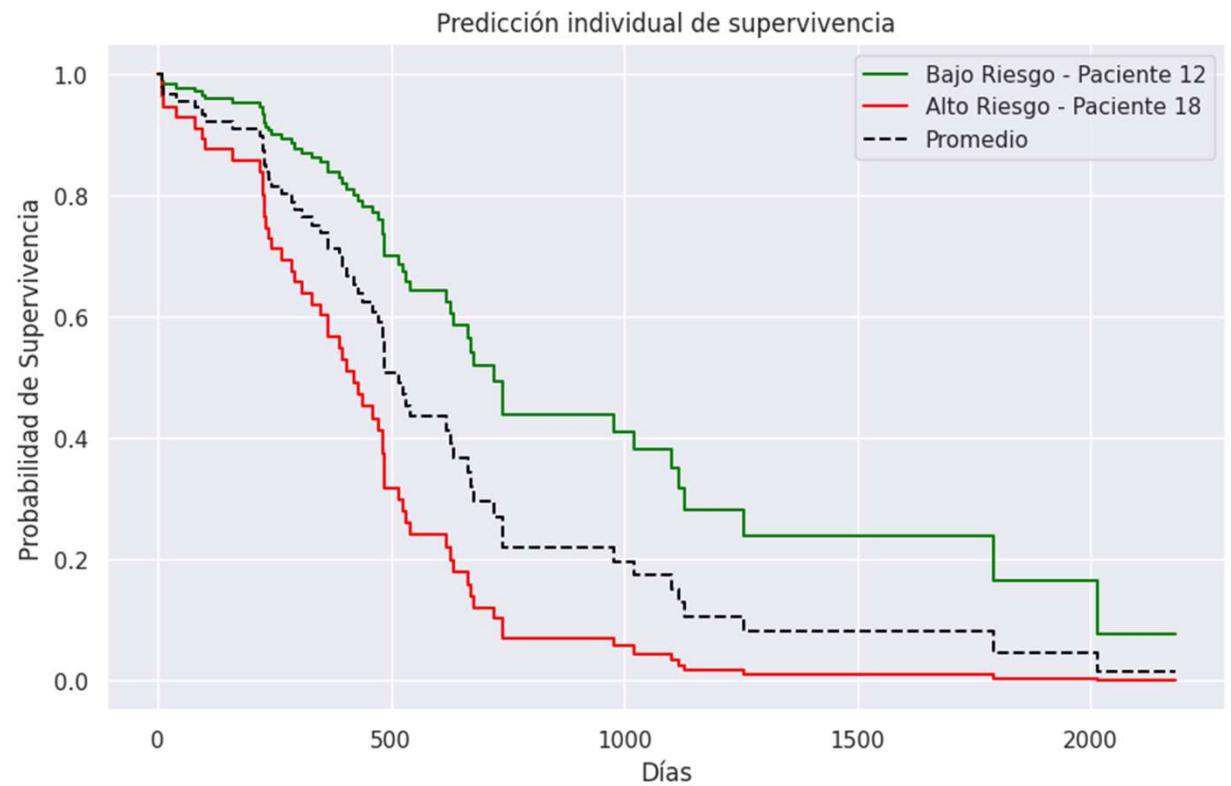
# Modelo de regresión de supervivencia (I)

- Indica la importancia y la dirección de cada variable.
- El comportamiento descrito concuerda con lo observado en el análisis de supervivencia.



# Modelo de regresión de supervivencia (II)

- Predicción de las curvas de supervivencia según el riesgo.
- Mayor amplitud entre las curvas indica mayor capacidad de predicción del modelo.



# Conclusiones

- La mayoría de los resultados obtenidos coinciden con los descritos en la literatura.
- El modelo de regresión de supervivencia, aunque simple y con pocos datos, consigue recrear el comportamiento observado en el análisis de supervivencia.

# Próximos pasos

- Incluir más capas de información disponibles en la muestra con el fin de buscar relaciones más profundas.
- Incorporar información de otras muestras con el fin de aumentar la cantidad de datos disponibles.