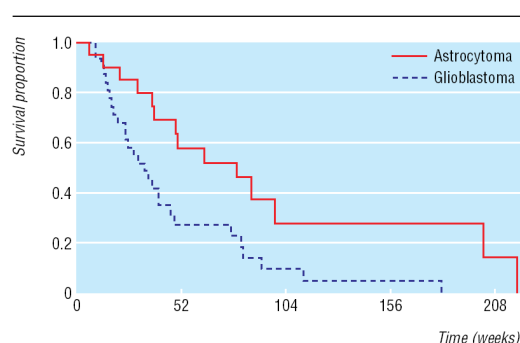


Logrank Test

por **Mauricio Lema Medina MD**

El Logrank test es el método favorito para comparar la sobrevida de grupos, que incluye todo el período de seguimiento y no la sobrevida en un punto arbitrario de las curvas. Otra ventaja es que no requiere del investigador que sepa nada sobre el comportamiento o la forma de las curvas o de la distribución de las mismas.

El logrank test tiene como hipótesis nula que no hay diferencias entre las poblaciones para la ocurrencia de un evento (i.e. sobrevida) en ningún momento del seguimiento. El análisis es basado en el tiempo de los eventos (i.e. muertes) de cada grupo que se comparan con el número esperado de eventos si no hubiera diferencias entre los grupos.



Survival curves for women with glioma by diagnosis

Ejemplo: Comparar la sobrevida de estos dos grupos de pacientes:

A (Astrocitoma): Muertos a las 6, 13, 21, 30, 31*, 37, 38, 47*, 49, 50, 63, 79, 80*, 82*, 82*, 86, 96, 149*, 202, 219 semanas

B (Glioblastoma Multiforme): Muertos a los 10, 10, 12, 13, 14, 15, 16, 17, 18, 20, 24, 24, 25, 28, 30, 33, 34*, 35, 37, 40, 40, 40*, 46, 48, 70*, 76, 81, 82, 91, 112, 181 semanas (asteriscos significan datos CENSORADOS, es decir, que se perdieron para el seguimiento en esa semana pero estaban con vida en ese momento).

Al inicio del estudio había 51 pacientes. La primera muerte ocurrió en la semana 6 del grupo A. El riesgo de muerte a la semana 6 es de $1/51$. Como había 20 pacientes en el grupo A, entonces la hipótesis nula es que el número de muertos esperados en el grupo A en la semana 6 es $20 \times 1/51 = 0.39$. Haciendo lo mismo con el grupo B tendríamos $30 \times 1/51 = 0.61$. Posteriormente hubo 2 muertos en la semana 10. Para ese momento había 19 pacientes en el grupo A y 31 pacientes en el grupo B. La probabilidad del evento en esa semana es $2/50$ (número de eventos, sobre número a riesgo). El riesgo esperado en el grupo A es $19 \times 2/50 =$

0.76 y el riesgo esperado en el grupo B es $31 \times 2/50 = 1.24$, respectivamente. Los cálculos se repiten cada vez que ocurra una muerte. Si un dato es censurado se cuenta como con riesgo de muerte hasta la semana en que está censurado pero no de allí en adelante (tal y como se maneja en las curvas de Kaplan-Meier). La sumatoria de los esperados en cada grupo ($0.39 + 0.76 + 0.40 + 0.81 \dots$ etcétera para el grupo A y $0.61 + 1.24 + 0.6 + 1.19 \dots$ etcétera para el grupo B) nos muestran que se esperan 22.48 y 19.52 muertos en el grupo A y B, respectivamente. El número de muertos observados fue de 14 y 28, respectivamente. Ahora podemos utilizar el test del Chi-cuadrado de la hipótesis nula. El test estadístico es la sumatoria de $(\text{Observado} - \text{Esperado})^2 / \text{Esperado}$. Entonces se calcula $(14 - 22.48)^2 / 22.48 + (28 - 19.52)^2 / 19.52 = 6.88$. Los grados de libertad son el número de grupos $- 1$ ($2 - 1 = 1$). Se busca este resultado en una tabla de Chi-cuadrado y se encuentra que $p < 0.01$, de tal manera que la diferencia entre los dos grupos es estadísticamente significativa. Existen otras formas de calcular este test, pero esta se puede aplicar a más de 2 grupos, y permite mostrar tendencias.

El logrank test asume lo mismo que las curvas de Kaplan-Meier: que la censura NO está relacionada con el pronóstico, que la probabilidad de sobrevida es igual en pacientes reclutados temprano o tarde en el estudio, y que los eventos ocurrieron en los tiempos especificados.

El logrank test tiende a encontrar diferencias cuando las curvas son consistentes en el tiempo y tienden a no encontrar diferencias cuando las curvas se cruzan. Se recomienda que siempre se elaboren las curvas. Igualmente se recomienda el uso de software para el cálculo del estadístico.

El logrank test es un test de significancia pura y no ofrece ninguna información sobre la magnitud de las diferencias entre los grupos o un intervalo de confianza. Para ello se utiliza el *hazard ratio* incluyendo el *cox-proportional hazards model* que no son discutidos en este escrito.

Tomado de:

Bland JM, Altman DG. **The logrank test.** BMJ 2004;328: 1073.