

I CURSO GEICAM

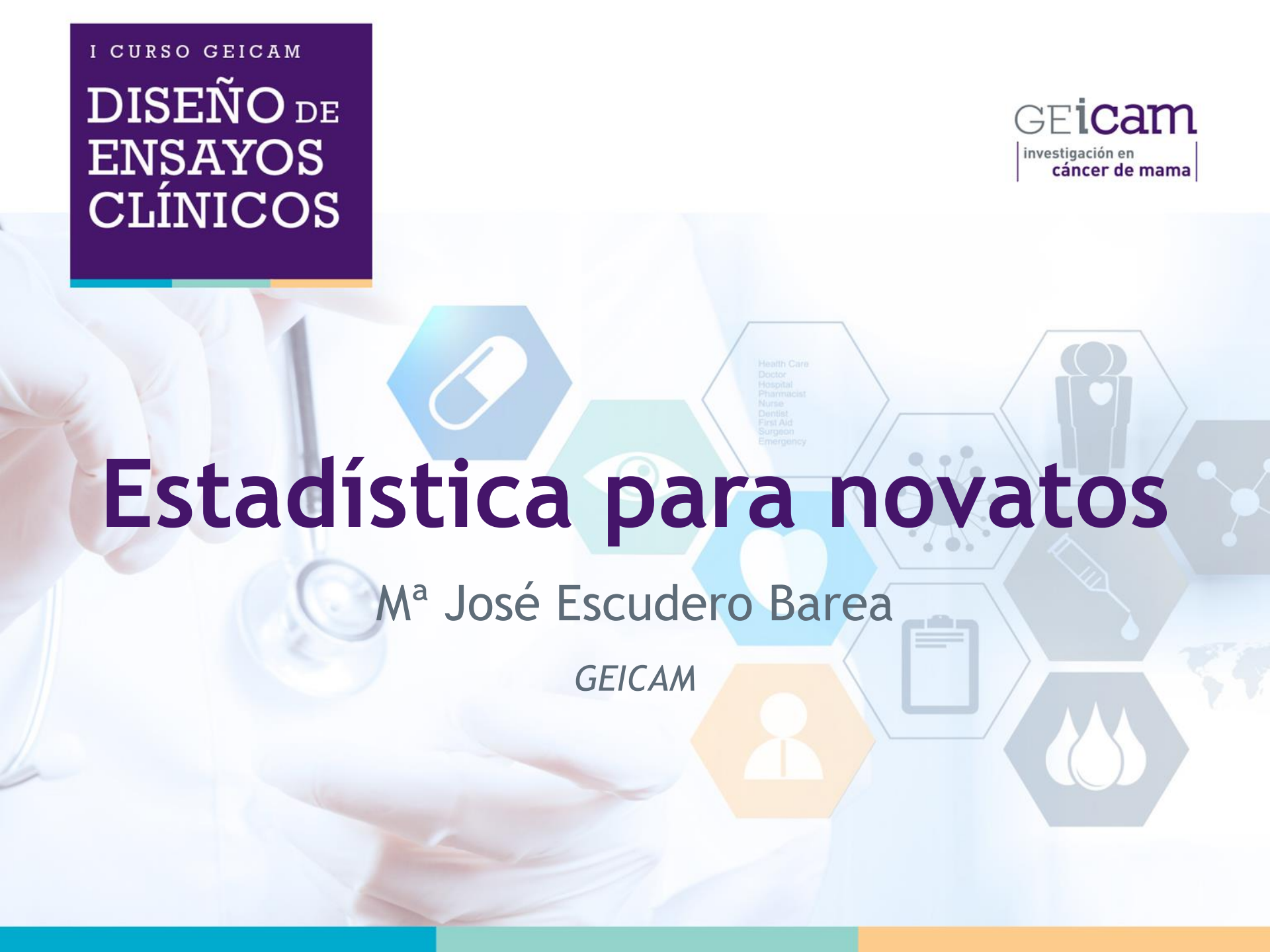
# DISEÑO DE ENSAYOS CLÍNICOS

GEICAM  
investigación en  
cáncer de mama

# Estadística para novatos

M<sup>a</sup> José Escudero Barea

GEICAM



- 1- Tipos de estudio
- 2- Ensayos clínicos
- 3- Cálculos de tamaños muestrales
- 4- Aleatorización/Estratificación
- 5- Clasificación del los EECC según el objetivo estudiado
- 6- Resultados
  - a) Supervivencia
  - b) Análisis Uni/Multivariante
  - c) Forest Plot

# Tipos de Estudios

# 1- Tipos de estudios

Los diseños de investigación se pueden clasificar según diversas características:

1- Según la manipulación de la variable que se quiere investigar:

- Experimentales
- No experimentales (Observación)
  - Descriptivos (No hipótesis)
  - Analíticos (Hipótesis)

# 1- Tipos de estudios

## 2- Según la existencia de seguimiento:

- Transversales (No seguimiento)
- Longitudinales (Si seguimiento)

## 3- Según el momento de ocurrencia de la información respecto al inicio del estudio:

- Prospectivos (El efecto sucede tras el inicio del estudio)
- Retrospectivos (El efecto ya se ha producido cuando se inicia el estudio)

## 4- Según el sentido:

- Estudios transversales (Sin sentido)
- Casos – Control (Efecto → Exposición)
- Cohortes (Exposición → Efecto)

# 1- Tipos de estudios

Ejemplos:

1- Se quiere estudiar la relación entre terapia hormonal sustitutiva en la incidencia de cáncer de mama. Si se seleccionan las personas con cáncer de mama y se estudia cuántas han recibido terapia hormonal sustitutiva.

CASOS Y CONTROLES

2- ¿En qué tipo de estudio se debe evaluar la exposición de manera aleatoria?

ENSAYOS CLÍNICOS

3- Se entrevista a 1000 personas y se documenta que el tabaquismo es el factor de riesgo más importante para el cáncer de pulmón. Este tipo de estudios que en los de menor nivel de evidencia científica.

TRANSVERSAL

# Ensayos clínicos



# 2-Ensayos clínicos

<b>FASE I</b>	<p>OBJETIVOS: Obtener pauta de administración Información preliminar sobre el efecto y la seguridad. Muy pocos pacientes (6-15)</p>
<b>FASE II</b>	<p>OBJETIVOS: Obtener información preliminar de eficacia (Respuesta/ Supervivencia) Información sobre el efecto y la seguridad, ampliar datos del Fase I. Pocos pacientes (25-200)</p>
<b>FASE III</b>	<p>OBJETIVOS: Evaluar la eficacia y seguridad del tratamiento experimental intentando reproducir las condiciones de uso habituales y considerando las alternativas terapéuticas disponibles. Aleatorizados. Muchos pacientes (100-5000)</p>
<b>FASE IV</b>	<p>OBJETIVOS: Se evalúa, en medicamentos comercializados, condiciones de uso distintas a las autorizadas, como por ejemplo una nueva indicación. Muchos pacientes (100-5000)</p>

# 2-Ensayos clínicos

## Fases II

- **Fase II:** Suelen tener como objetivos respuesta, supervivencia libre de progresión o supervivencia global. Su objetivo principal suele ser discriminar entre tratamientos prometedores
- **Población de eficacia: Per Protocol:** Subconjunto de ITT que ha recibido la primera dosis del fármaco y que cumple con los criterios de inclusión y exclusión (y puede tener algún otro requerimiento, según lo que se especifique en el protocolo)

### TRADICIONALES, solo 1 brazo

Se utiliza el siguiente test de hipótesis

$H_0 : p=p_0$  (si no la podemos rechazar, no se debe seguir estudiando el nuevo tratamiento)

$H_1 : p=p_1$

Ver ejemplo en el cálculo de tamaño muestral - **SIMON**

# 2-Ensayos clínicos

## *Fases II-Aleatorizados*

### **FASE II ALEATORIZADO, 2 o más brazos**

Igual que el anterior pero se incluyen pacientes en el brazo control.

El objetivo de aleatorizar es asegurarnos que el porcentaje de respuestas elegido en el brazo control es correcto (poca información histórica, cambios rápidos en el tratamiento standard).

En caso de que el objetivo sea OS o PFS es aconsejable comparar con brazo control en lugar de con datos históricos.

- Aumenta considerablemente el tamaño muestral.
- No se deben comparar los brazos de tratamiento.

# 2-Ensayos clínicos

## Fases II-Aleatorizados

### FASE II ALEATORIZADO, “Pick the winner”

El objetivo es decidir de entre un grupo de tratamientos cual debo seleccionar para continuar con la siguiente fase.

La conclusión en este tipo de estudios es “Con un 90% de probabilidad el mejor tratamiento será X , cuando la diferencia entre ambos brazos en términos de respuesta sea mayor o igual al 15%”.

#### **Problema con este tipo de diseños:**

- ✓ Si la diferencia entre los brazos es menor al 15%, el procedimiento no siempre elegirá el mejor.
- ✓ Si hay mas de uno con una diferencia superior al 15%, solo elegirá uno y los demás se descartarán.
- ✓ Si ningún tratamiento es muy bueno el procedimiento seguirá eligiendo uno.

# 2-Ensayos clínicos

## *Fases III*

- **Fase III:** Estudios aleatorizados. Su objetivo principal suele ser comparar tratamientos.
- Suelen tener como objetivos comparar Supervivencia (SLE, SLP, SG) entre 2 tratamientos. También se suele usar como objetivo comparar la respuesta a los diferentes tratamientos.
- Población de eficacia: ITT

# 2-Ensayos clínicos

## Fases III-Interim Analysis

**Interim Analysis:** Se realizan en estudios muy largos para comprobar la seguridad y eficacia del tratamiento.

- Están planificados en el protocolo del estudio.
- Se debe usar un alfa menor que se calcula por un método ya establecido.
- Ejemplo: **PEARL STUDY** (*Phase III study of Palbociclib (PD-0332991) in combination with Endocrine therapy (exemestane or fulvestrant) versus chemotherapy (capecitabine) in Hormonal Receptor (HR) positive/HER2 negative Metastatic Breast Cancer (MBC) patients with Resistance to non-steroidal Aromatase inhibitors* )

Interim analysis: p-value (2 sided) = 0.0002

Final analysis: p-value (2 sided) = 0.0499

- En estos análisis intermedios se puede planear una modificación del tamaño muestral o una parada del estudio, dependiendo de los datos (diseños adaptativos)

# Cálculos de tamaños muestrales

# 3- Cálculos de tamaños muestrales

## ¿Qué es el tamaño muestral?

El tamaño de la muestra (**TM**) representa el número mínimo de elementos que debería tener una muestra para que proporcione resultados suficientemente fiables.

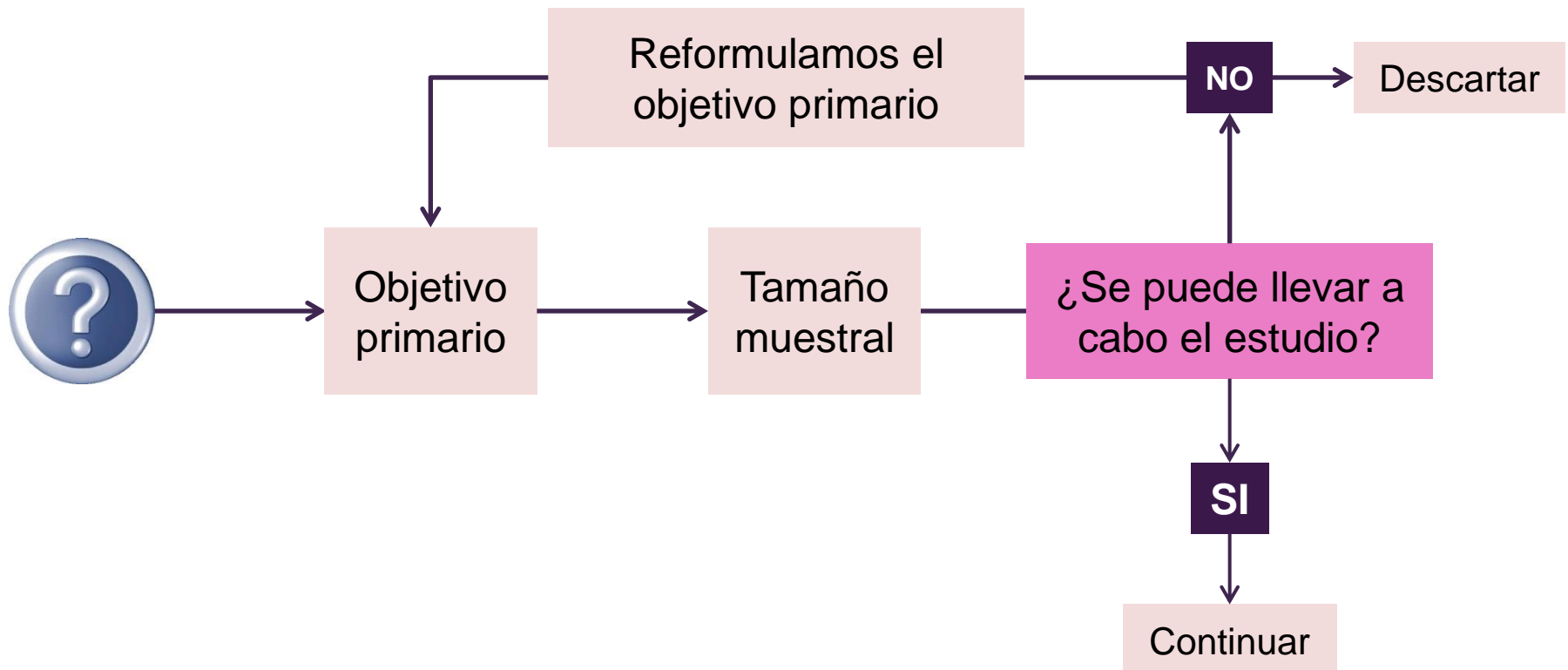
El TM es el número óptimo de sujetos de manera que se obtenga una mayor precisión con el mínimo de sujetos.

- ✓ Cuanto mayor sea el TM, mayor precisión se obtendrá en el análisis.
- ✓ Por motivos éticos / económicos se requiere que el tamaño de muestra sea el más pequeño posible.



# 3- Cálculos de tamaños muestrales

- ✓ El cálculo del tamaño muestral debe hacerse al comienzo del proceso de investigación.
- ✓ El tamaño muestral se debe basar en el objetivo primario.



# 3- Cálculos de tamaños muestrales

CONOCIMIENTOS  
CLÍNICOS



CONOCIMIENTOS  
ESTADÍSTICOS

# 3- Cálculos de tamaños muestrales

## Siempre es más importante la significación clínica que la estadística

- ✓ Relevancia clínica es la diferencia mínima entre 2 tratamientos que consideremos clínicamente relevante para administrar un tratamiento y no otro.

### **Estadísticamente significativo y Clínicamente no significativo**

La significación estadísticamente se puede “comprar” con una N muy grande:

Ej: Aumento de 3 días en la supervivencia del cáncer.

### **Estadísticamente NO significativo y Clínicamente significativo**

Ej: Aparición de eventos adversos muy graves que modifican el balance riesgo / beneficio.

# 3- Cálculos de tamaños muestrales

## ¿Qué necesitamos para calcular el TM?

- ✓ Variable primaria.
- ✓ Test estadístico.
- ✓ Hipótesis nula.
- ✓ Hipótesis de trabajo (con las asunciones del efecto esperado o deseado).
- ✓ Probabilidad de error tipo I (alfa).
- ✓ Probabilidad de error tipo II (beta).
- ✓ Estimación de pérdidas.
- ✓ Un buen software.

# 3- Cálculos de tamaños muestrales

## Contraste de Hipótesis

- ✓ Hipótesis es una suposición.
- ✓ Queremos determinar si existe evidencia estadística a favor de la hipótesis inicial.

$H_0$	$H_1$
Se plantea en primer lugar y se asume como correcta	Negación de la $H_0$
Se mantendrá a no ser que los datos indiquen su falsedad	Es la que queremos verificar
Nunca se acepta	Puede aceptarse o no aceptarse
Contiene = $\leq$ $\geq$	Nunca contiene = $\leq$ $\geq$

# 3- Cálculos de tamaños muestrales

## *Contraste de Hipótesis*

### ✓ **Ejemplo:**

Queremos estudiar si existe evidencia estadística para afirmar que las pacientes que reciben el tratamiento B tienen una respuesta diferente a las que reciben el tratamiento A.

$$H_0: pCR (A) = pCR (B)$$

$$H_1: pCR (A) \neq pCR (B)$$

# 3- Cálculos de tamaños muestrales

## Errores $\alpha$ y $\beta$

- Estamos ante un juicio con jurado.
- Dicho jurado **asume inicialmente** que el acusado es inocente.
- El jurado deberá probar que el acusado es culpable más allá de toda **duda razonable**.
- Sólo entonces el jurado rechazará la hipótesis de inocencia y declarará al acusado culpable.



# 3- Cálculos de tamaños muestrales

## Errores $\alpha$ y $\beta$

NUESTRA DECISIÓN	LA VERDAD	
	<i>Inocente</i>	<i>Culpable</i>
<i>Presunción de inocencia no aceptada (Culpable)</i>	Una persona inocente irá a la cárcel <b>Error Tipo I (<math>\alpha</math>)</b>	Decisión Correcta
<i>Presunción de inocencia aceptada (No Culpable)</i>	Decisión Correcta	El culpable queda libre <b>Error Tipo II (<math>\beta</math>)</b>

¿Qué puede ocurrir en el juicio?

Hipótesis Nula: El acusado es no culpable

Hipótesis Alternativa: El acusado es culpable



# 3- Cálculos de tamaños muestrales

## Errores $\alpha$ y $\beta$

Se realiza el estudio y, basándonos en los datos, se toma una decisión. Al tomar esa decisión podemos cometer 2 errores: **Tipo I y Tipo II**

		Realidad	
		H0 cierta	H0 falsa
Resultado del estudio	Rechazar H0	Error Tipo I ( $\alpha$ )	Correcto
	No rechazar la H0	Correcto	Error Tipo II ( $\beta$ )

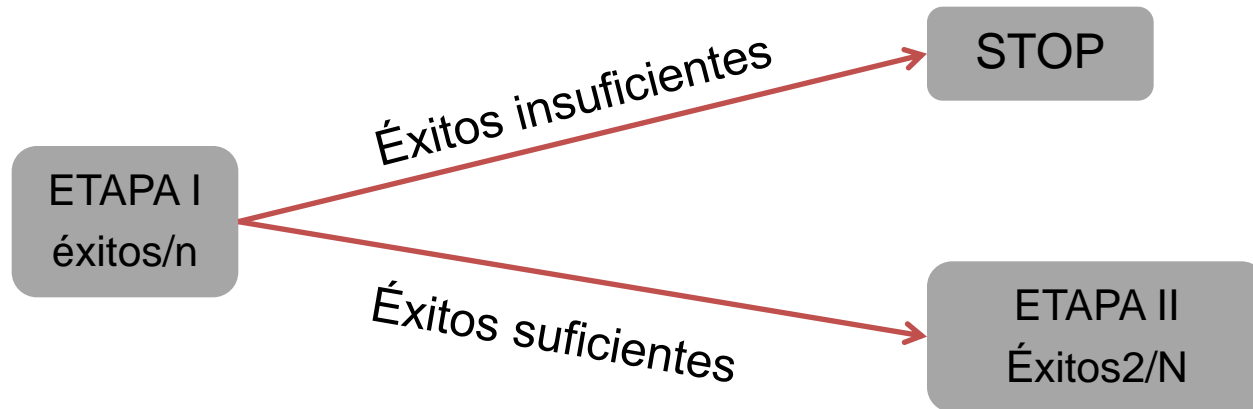
- $\alpha$  : Nivel de significación (error tipo I) (Falso positivo):
  - $\alpha = p$  (rechazar H0 | H0 cierta)
  - Estándar: 0,05
  - Conservador: 0,01. Si el tto. es muy tóxico y se quiere tener más evidencia de efectividad.
- $1-\beta$  : Potencia (error tipo II) (Falso negativo):
  - $\beta = p$  (aceptar H0 | H0 falsa)
  - Potencia =  $1- \beta = p$  (rechazar H0 | H0 falsa)
  - Estándar : 0,80 ( $\beta=0,20$ )

\*  $\alpha$  y  $\beta$  están inversamente relacionados  
\* Solo pueden disminuirse simultáneamente aumentando la n.

# 3- Cálculos de tamaños muestrales

## Fases II

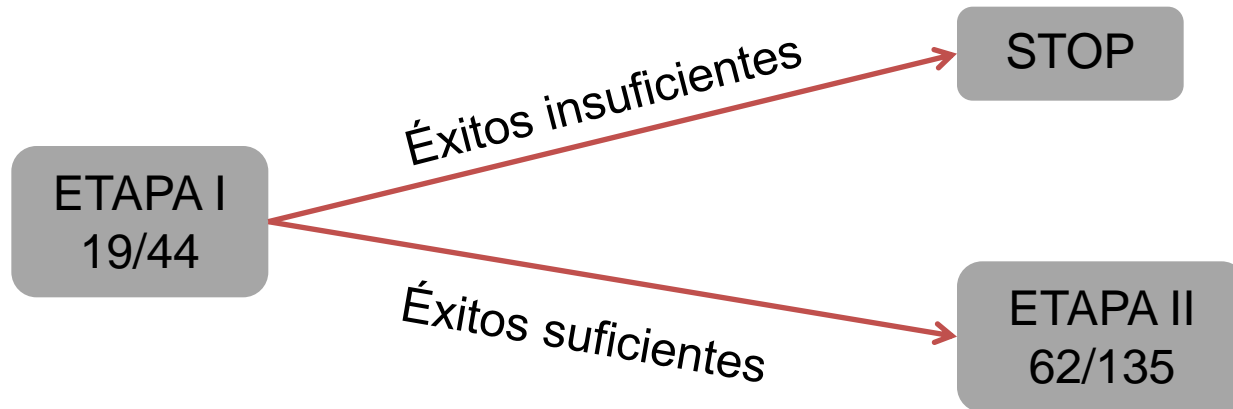
- ✓ Valorar la eficacia de un tto. experimental frente a la eficacia del tto. de referencia.
- ✓ Dicha valoración se lleva a cabo mediante el % de éxitos.
- ✓ Se realiza un análisis intermedio para detener el estudio si no hay indicio del éxito del tratamiento experimental.
- ✓ 1º ETAPA: Se establece un umbral de éxitos a superar



# 3- Cálculos de tamaños muestrales

## Fases II- Ejemplos

- Fase II
- Evaluar la eficacia de un tratamiento experimental en pacientes con cáncer de mama.
- Se considera que el tratamiento experimental ofrecerá una eficacia del 52%.
- Sabemos que con el tratamiento de referencia la eficacia está en torno al 40%.
- $\alpha=0,05$ ;  $\beta=0,80$  (poder= $1-\beta=0,20$ ).



# 3- Cálculos de tamaños muestrales

## Fases II- Ejemplos

- <https://stattools.crab.org>
- ENE

Comparación de dos proporciones / Una muestra frente a un valor teórico / No igualdad, Simon bietápico

Tamaño de Muestra

Parámetros	
Parámetro	Valor
Nivel de Significación ( $\alpha$ )	5,00
Proporción Teórica Grupo de Referencia ( $p_1$ )	40,00
Proporción Grupo Experimental ( $p_2$ )	52,00
Potencia ( $pow = 1 - \beta$ )	80,00

Resultados	
Resultado	Valor
Eventos Fase I ( $rs_1$ )	19
Tamaño de Muestra Fase I ( $ns_1$ )	44
Eventos Totales ( $rs$ )	62
Tamaño de Muestra Total ( $ns$ )	135

**Redactado**

Para conseguir una potencia del 80,00% para detectar diferencias en el contraste de la hipótesis nula  $H_0: p_1=p_2$  mediante una prueba binomial exacta unilateral para una muestra en un estudio bietápico, teniendo en cuenta que el nivel de significación es del 5,00%, y asumiendo que la proporción en el grupo de Referencia es del 40,00% y la proporción en el grupo Experimental es del 52,00%, será necesario emplear el criterio de decisión:  $rs_1 / ns_1$  y  $rs / ns$  respectivamente para ambas fases.

Ejemplo Imprimir... Borrar todo Calcular

Calculate Power/Sample Size

Power	Sample Size	Approx Lower Count Critical Value	Approx Upper Count Critical Value
.80	135	43	65

# 3- Cálculos de tamaños muestrales


## Fases III- Ejemplo

- ✓ Se desea comparar el tiempo de supervivencia entre 2 muestras.
- ✓ Hipótesis:
  - $H_0$ : La supervivencia en ambos brazos es igual
  - $H_A$ : La supervivencia en ambos brazos no es igual
- ✓ Prueba estadística: Log-Rank.
- ✓ Se tiene en cuenta la duración del periodo de inclusión / duración máxima del seguimiento.
- ✓ Se especifica la mediana de tiempo de supervivencia en ambos brazos.

# 3- Cálculos de tamaños muestrales

## Fases III- Ejemplo

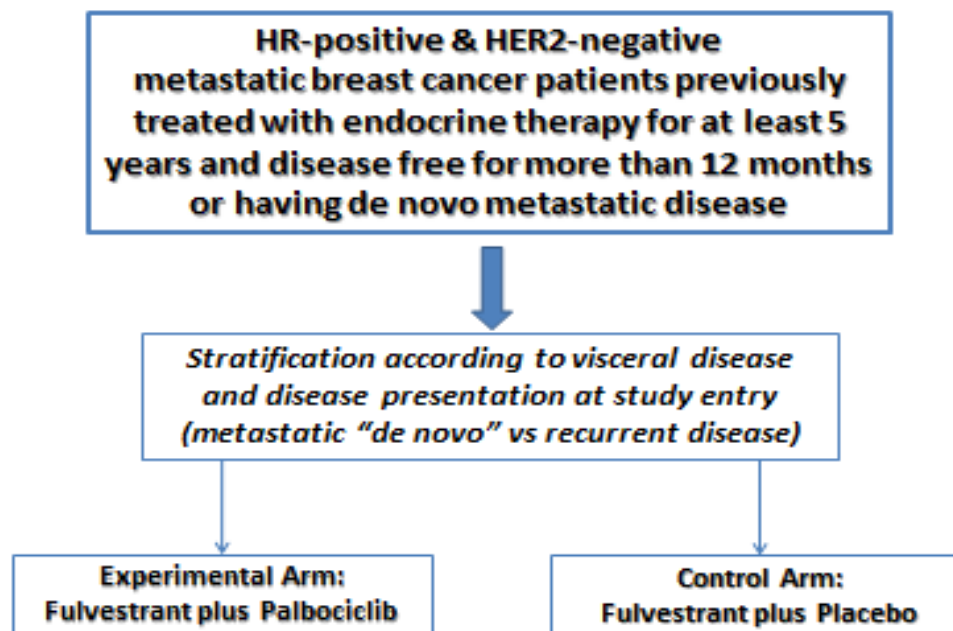
- ✓ Queremos comparar un tratamiento experimental contra el cáncer de mama respecto de un tratamiento de referencia.
- ✓ Periodo de inclusión 12 meses.
- ✓ Periodo de seguimiento 36 meses.
- ✓ Consultando la bibliografía sabemos que la mediana de tiempo de supervivencia de los pacientes en el brazo de referencia es de 30 meses.
- ✓ Se asume que la mediana del tiempo de supervivencia de los pacientes en el brazo experimental es de 40 meses.
- ✓  $\alpha=0.05$ ; potencia del 80%.


$$n_1=n_2= 485$$

# 3- Cálculos de tamaños muestrales

## Fases II- Ejemplo

**FLIPPER STUDY:** A randomized, double-blind, parallel-group, multicentre, phase II study to compare the efficacy and tolerability of fulvestrant (Faslodex™) 500mg with placebo and fulvestrant (Faslodex™) 500mg in combination with PD-0332991 (Palbociclib) as first line treatment for postmenopausal women with hormone receptor-positive metastatic breast cancer, who have completed at least 5 years of adjuvant endocrine therapy and remained disease free for more than 12 months following its completion or have “de novo” metastatic disease.



# 3- Cálculos de tamaños muestrales

## Fases II- Ejemplo

- ✓ Queremos comparar la PFS a 1 año entre 2 brazos de tratamiento.
- ✓ PFS a 1 año en el brazo control=54,5%.
- ✓ HR= 0,6.
- ✓  $\alpha=0.2$ ; potencia del 80%.

$$HR = \frac{\ln(PFS\ expe)}{\ln(PFS\ control)}$$
$$0.6 = \frac{\ln(x)}{\ln(0,545)}$$

X=0,695

NQUERYSRV - 10.10.10.142 - Escritorio remoto  
nQuery Advisor - [STT0-tmp53A7.nqa]

File Edit View Options Assistants Randomize Plot Window Help

Log-rank test of survival in two groups followed for fixed time, constant hazard rate

	1	2	3	4
Test significance level, $\alpha$	0,200	0,200		
1 or 2 sided test?	2	2		
Group 1 proportion $\pi_1$ at time t	0,545	0,695		
Group 2 proportion $\pi_2$ at time t	0,695	0,545		
Hazard ratio, $h = \ln(\pi_1) / \ln(\pi_2)$	1,668	0,599		
Power ( % )	80	80		
n per group	95	95		
Total number of events required, E	69	69		



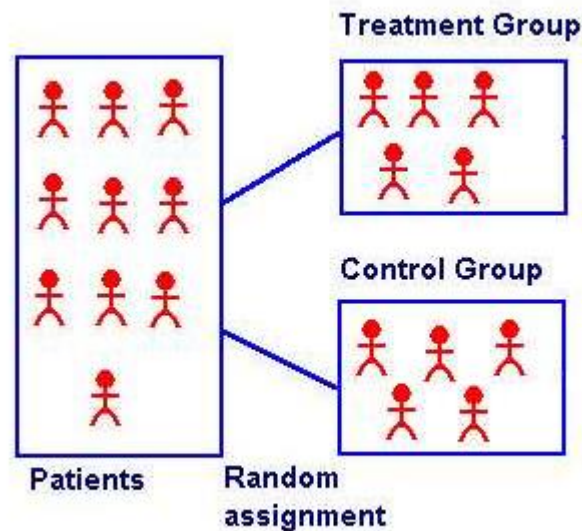
# Aleatorización / Estratificación

# 4- Aleatorización/Estratificación

## ¿Qué es la aleatorización?

La aleatorización significa que los sujetos son distribuidos al azar en cada brazo del estudio. Es decir tienen la misma probabilidad de pertenecer al brazo A (p) y la misma probabilidad de pertenecer al brazo B (1-p).

Los controles siempre deben ser prospectivos.



## PROPIEDADES

Aumenta la validez interna del estudio, al luchar contra los sesgos de selección y confusión.

- ✓ Sesgo de selección: Sesgo que se produce cuando el investigador asigna los tratamientos.
- ✓ Sesgo de confusión: Se produce cuando se incorporan en el modelo factores de riesgo extraños al estudio.

## OBJETIVO

Conseguir que todos los factores extraños se distribuyan por igual en los grupos de estudio, y que la única diferencia entre los grupos que se comparan sea el tratamiento.

## Tipos de aleatorización:

- 1. Aleatorización total:** Todos los participantes tienen la misma probabilidad de pertenecer a un brazo o a otro (Ej: lanzamiento de una moneda).  
**Problemas:** Tiende a producir brazos de distinto tamaño (tanto más cuanto menor es el número de participantes).
- 2. Aleatorización por bloques:** Procedimiento pseudo-aleatorio. Garantiza la igualdad total entre los brazos que se comparan. Se crean bloques de tratamiento que estén completamente balanceados entre los brazos, y se va asignando un tratamiento por orden de aleatorización.  
**Problemas:** Si los tratamientos no están enmascarados se puede llegar a adivinar el tratamiento.

## 4- Aleatorización/Estratificación

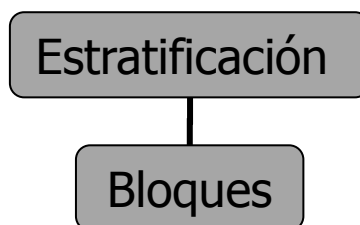
**3. Estratificación:** Distribuir a los pacientes que tengan una determinada característica equitativamente entre el brazo A y el B. Se utiliza cuando se quieren eliminar factores determinantes que influyen en la respuesta a la intervención.

### EJEMPLO

Se quiere comparar un tratamiento experimental para las mujeres con cáncer de mama en estadios I y II.

**¿Podemos dejar al azar la distribución de la afectación axilar entre uno y otro brazo?**

Utilizamos para aleatorizar 2 métodos de manera conjunta:



# 4- Aleatorización/Estratificación

## Ventajas de la estratificación

- ✓ Asegurarse de que los brazos que se comparan son similares con respecto a factores pronósticos conocidos.
- ✓ Mayor protección frente al error tipo I (disminuyen los falsos positivos).
- ✓ Mayor protección frente al error tipo II (disminuyen los falsos negativos).
- ✓ Mayor eficiencia.

## Desventajas de la estratificación

- ✓ Si el número de estratos es pequeño no hay desventajas claras.
- ✓ La estratificación por más de 2 factores con más de 2 niveles en cada uno de ellos, puede conducir a diseños impracticables.
- ✓ La estratificación excesiva es inútil: normalmente conduce a una falta de balance (desequilibrio entre los brazos a comparar).



# Clasificación de los EECC según el objetivo estudiado



# 5- Clasificación de los EECC según objetivo

## Superioridad

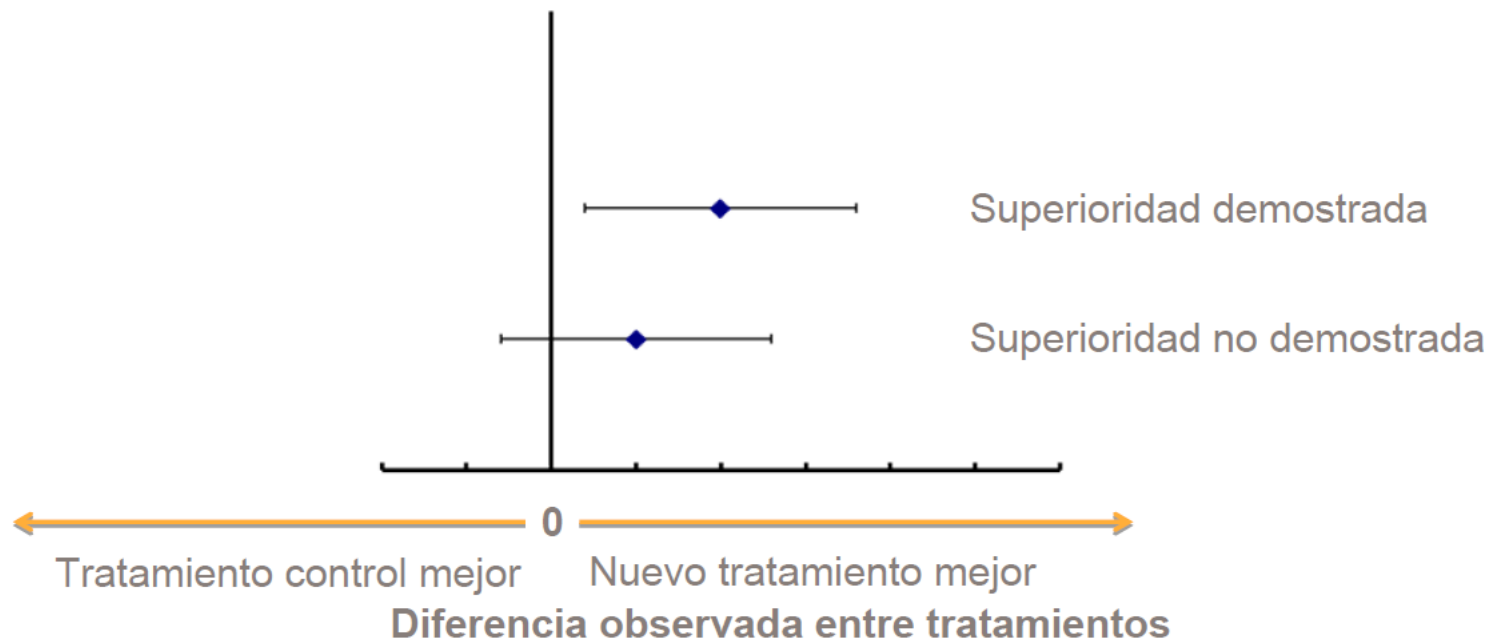
**1. Superioridad:** ¿Es el nuevo tratamiento mejor que el que se usa en la actualidad?

**ICH-E9: A trial with the primary objective of showing that the response to the investigational product is superior to a comparative agent (active or placebo)**

- **Objetivo:** Detectar diferencias entre tratamientos.
- Solo estableceremos superioridad si el grupo control demuestra eficacia en nuestro ensayo
- Si el nuevo tratamiento es mejor, podremos establecer la superioridad, pero utilizando un elevado número de sujetos

# 5- Clasificación de los EECC según objetivo

*Superioridad*



- Hipótesis nula:  $A=B$
- Hipótesis alternativa:  $A \neq B$

NOTA: Si obtenemos un p-valor de 0,853 ¿implica equivalencia?

NO → Nunca podemos aceptar la  $H_0$

“Absence of evidence is not evidence of absence”

# 5- Clasificación de los EECC según objetivo

## Equivalencia

**2. Equivalencia:** ¿Puedo utilizar el nuevo tratamiento o el utilizado en el grupo control indistintamente?

**ICH-E9: A trial with the primary objective of showing that the response to two or more treatments differs by an amount which is clinically unimportant. This is usually by showing that the true treatment difference is likely to lie between a lower and an upper equivalence margin of clinically acceptable differences.**

- **Objetivo:** Demostrar que la respuesta de 2 tratamientos difiere únicamente en una cantidad considerada clínicamente no importante.
- Debemos seleccionar un margen de equivalencia clínico ( $\Delta$ ) que vendrá definido como la máxima diferencia que es clínicamente aceptable entre los tratamientos, y por lo tanto una diferencia mayor a la propuesta tendría inconvenientes en la práctica clínica.

# 5- Clasificación de los EECC según objetivo

## Equivalencia

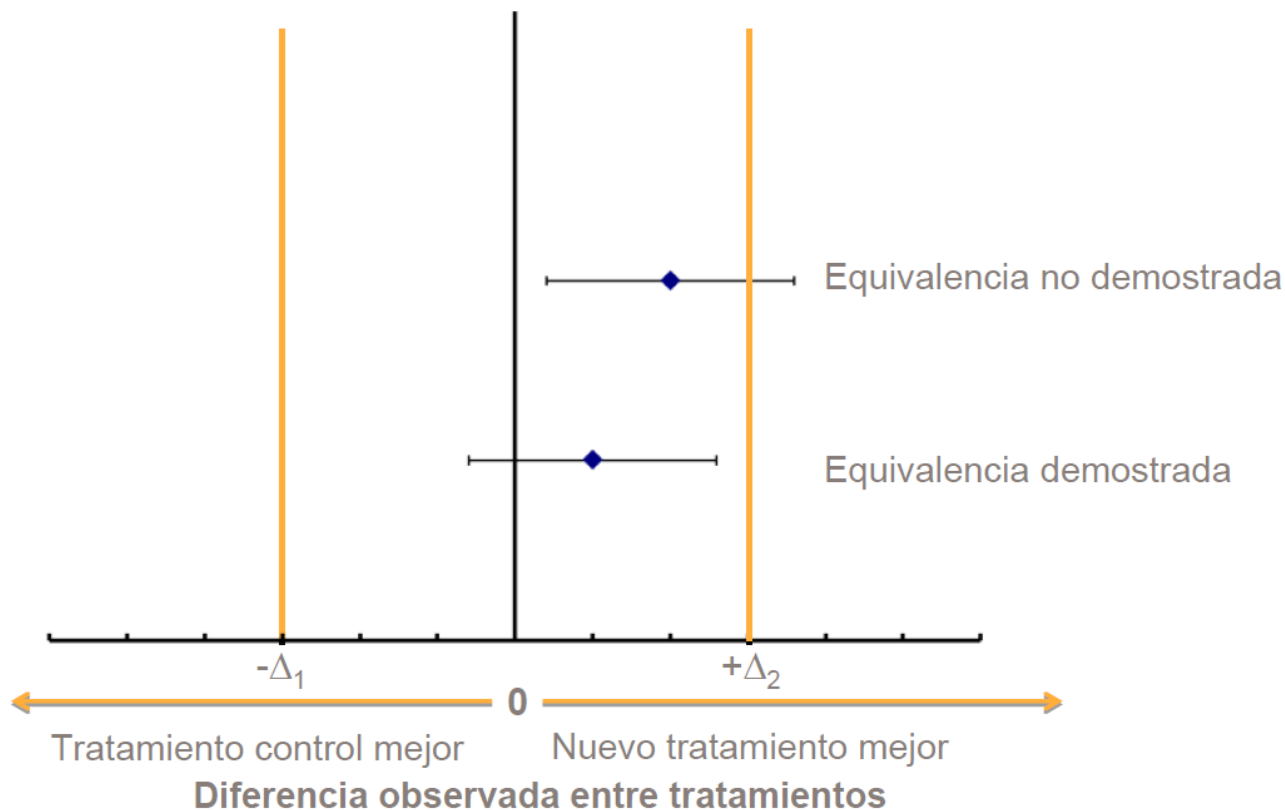
### DELTA:

Se debe establecer al inicio del estudio

Máxima diferencia entre los tratamientos que consideramos clínicamente irrelevante.

Es un intervalo definido entre 2 límites  $\Delta_1$  y  $\Delta_2$

### EQUIVALENCIA:



# 5- Clasificación de los EECC según objetivo

## *No-inferioridad*

**3. No-Inferioridad:** Si no es mejor, ¿es el nuevo tratamiento tan bueno como el empleado en el grupo control?

**ICH-E9: A trial with the primary objective of showing that the response to the investigational product is not clinically inferior to a comparative agent (active or placebo control)**

- **Objetivo:** Demostrar que el producto en investigación no es peor que el comparador por una cantidad pre-especificada y pequeña ( $\Delta$ ).
- **ERROR:** A veces se define de forma incorrecta “demostrar que el producto de investigación no es inferior al comparador”.
- Siempre debemos escoger la mejor pauta de control existente

# 5- Clasificación de los EECC según objetivo

*No-inferioridad*

## DELTA:

$\Delta$ = Las diferencias entre el fármaco experimental y el control activo no son superiores a una magnitud prefijada.

El margen de no inferioridad se determinará por una combinación de consideraciones clínicas y estadísticas.

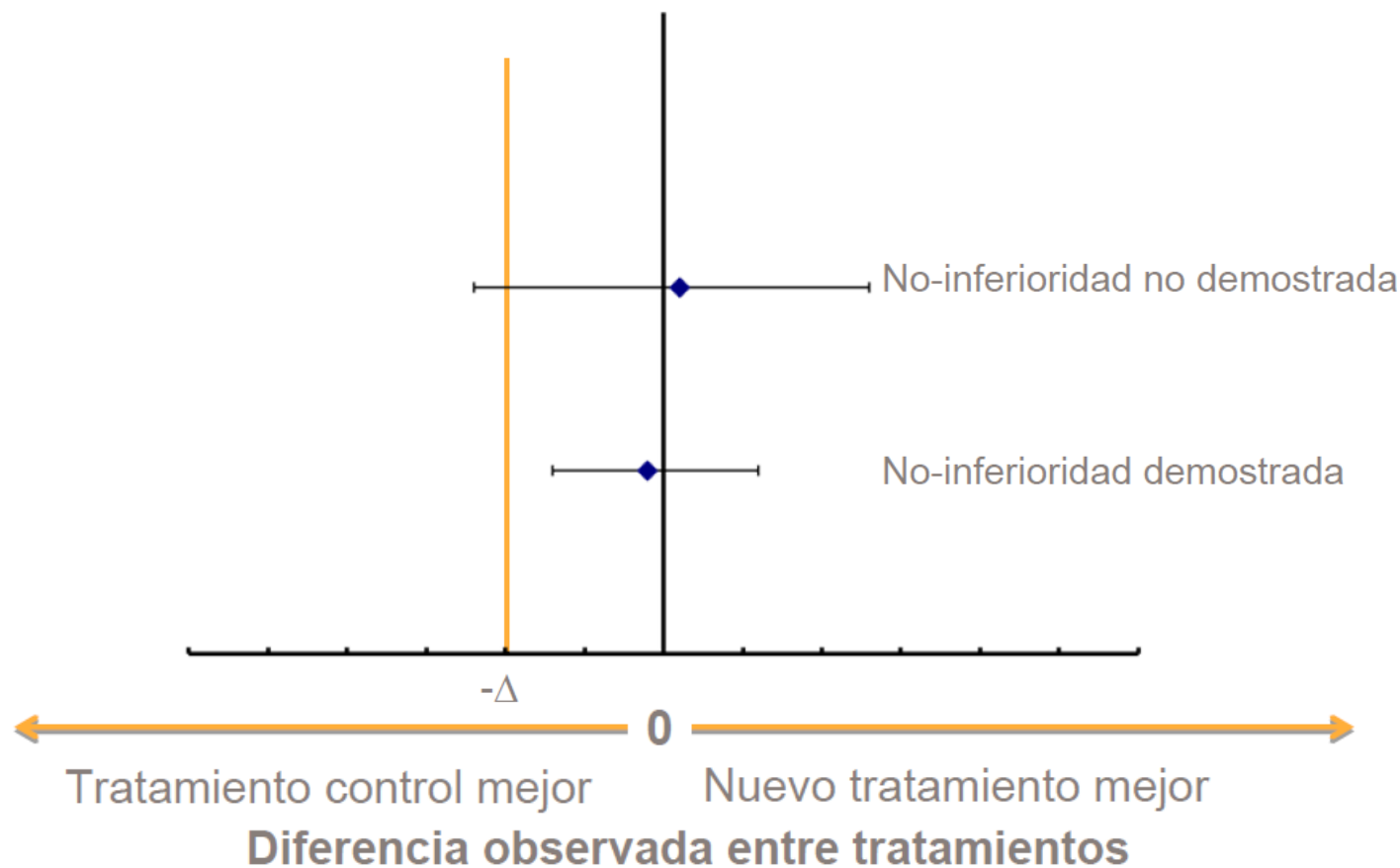
La conclusión de no-inferioridad con un  $\Delta=5\%$  nos dirá que , en el peor de los casos, el nuevo tratamiento tendrá una tasa de curación de un 5% por debajo del grupo control

### RAZONAMIENTO CLÍNICO:

- Diferencias del 5% son clínicamente irrelevantes
- El precio a pagar por obtener otros beneficios (ej: mejor perfil de seguridad)

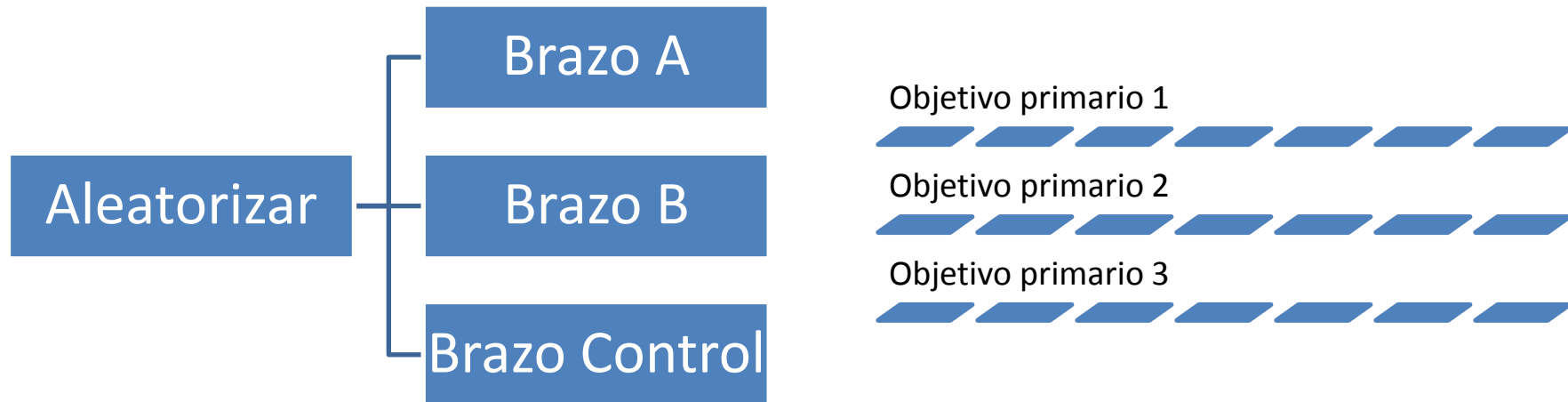
# 5- Clasificación de los EECC según objetivo

*No-inferioridad*



# 5- Clasificación de los EECC según objetivo

MULTI-ARM/MULTIOBJECTIVE



Incrementa la probabilidad de éxito erróneamente

Corregir por el método de Bonferroni

Nuestra  $H_0$  está formada por  $m$  hipótesis, una por cada comparación que queramos realizar.

- Rechazaremos la  $H_i$  si  $p_i \leq \frac{\alpha}{m}$  (Bonferroni)

**Se debe intentar siempre que sea posible:**

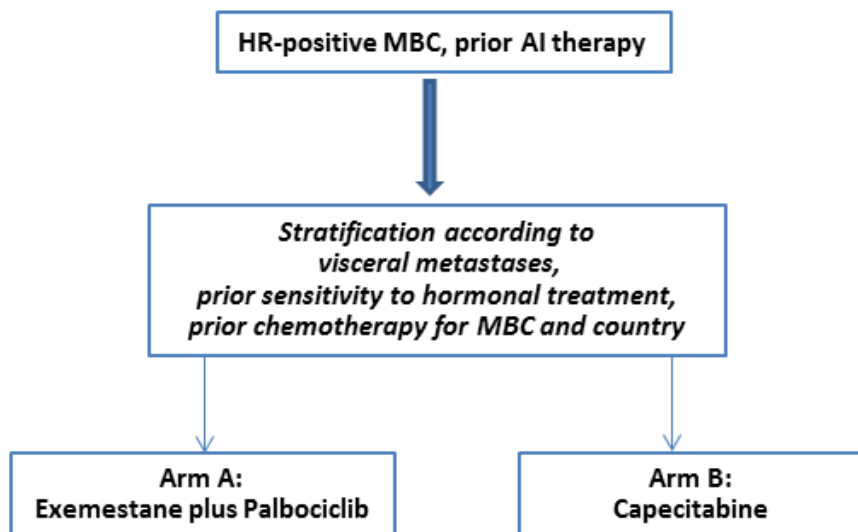
- Identificar una variable principal relevante
- Seleccionar solo un contraste de tratamientos principal



# 5- Clasificación de los EECC según objetivo

## MULTIOBJECTIVE

**Phase III study of Palbociclib (PD-0332991) in combination with Endocrine therapy (exemestane or fulvestrant) versus chemotherapy (capecitabine) in Hormonal Receptor (HR) positive/HER2 negative Metastatic Breast Cancer (MBC) patients with Resistance to non-steroidal Aromatase inhibitors**



1- To demonstrate that palbociclib in combination with fulvestrant is superior to capecitabine in prolonging Progression-Free Survival (**PFS**) in postmenopausal women with ER positive/HER2 negative MBC whose disease was resistant to non-steroidal aromatase inhibitors, regardless of the ESR1 mutational status.

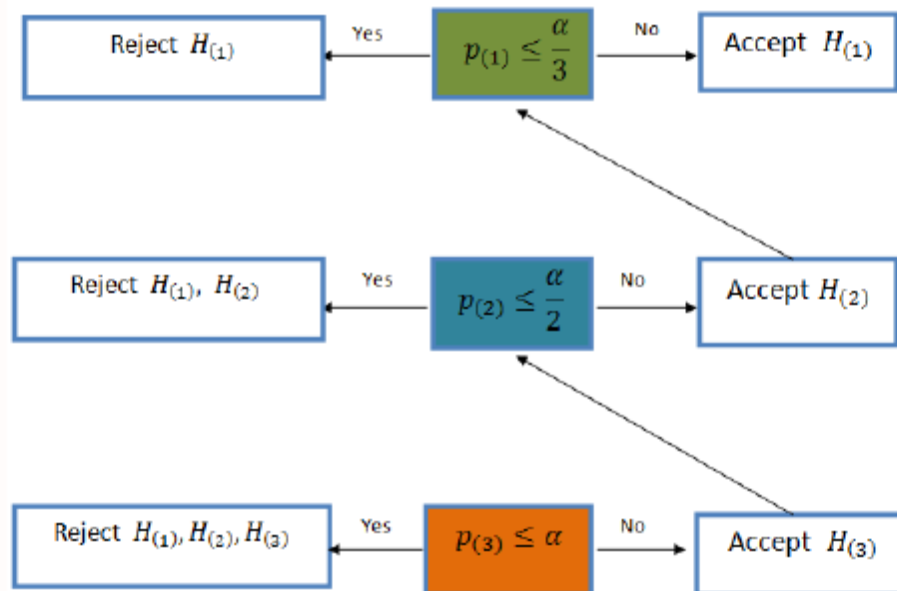
2- To demonstrate that palbociclib in combination with endocrine therapy (exemestane or fulvestrant) is superior to capecitabine in prolonging **PFS** in postmenopausal women with ER positive/HER2 negative MBC whose disease was resistant to non-steroidal aromatase inhibitors and whose tumor had estrogen receptor (*ESR1*) mutational status as wild type at study entry.

# 5- Clasificación de los EECC según objetivo

MULTIOBJECTIVE

## HORCHBERG PROCEDURE

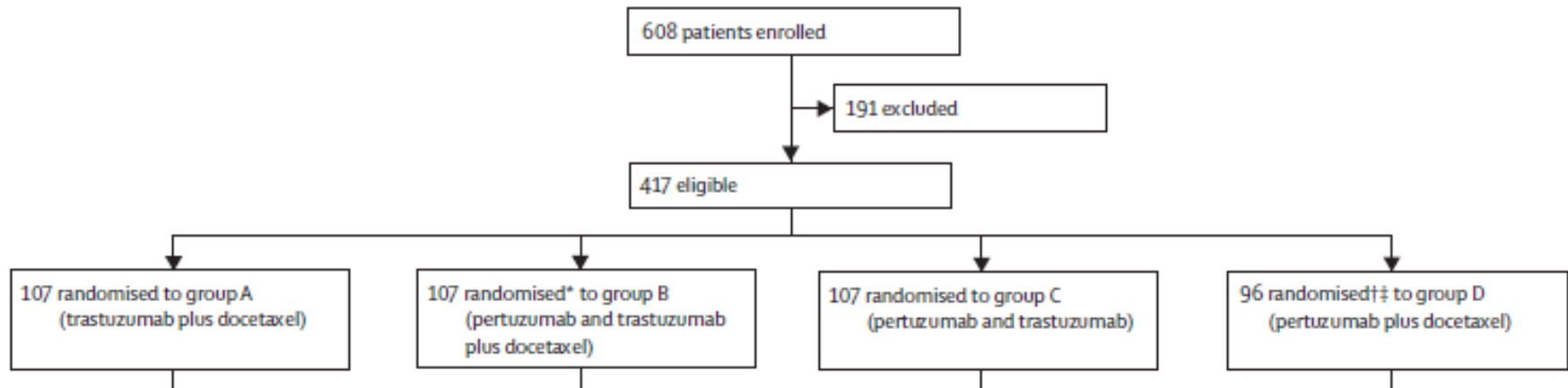
- Start with the least significant hypothesis; as soon as one rejects a null hypothesis, all further hypotheses are automatically rejected



# 5- Clasificación de los EEC según objetivo

MULTI-ARM

Efficacy and safety of neoadjuvant pertuzumab and trastuzumab in women with locally advanced, inflammatory, or early HER2-positive breast cancer (NeoSphere): a randomised multicentre, open-label, phase 2 trial



# Resultados

# 6- Resultados

## a) Supervivencia

**Análisis de supervivencia:** Conjunto de técnicas que permiten estudiar la variable “tiempo hasta que ocurre un evento” y su dependencia de otras variables explicativas. Los métodos más utilizados son el método de **Kaplan-Meier** y el método actuarial o de **tabla de vida**.

Ej: Se quiere estudiar si un tratamiento A aumenta la SLE (Supervivencia libre de enfermedad) respecto al estándar B.

En el análisis de supervivencia podemos estar interesados en puntos específicos, por ejemplo , supervivencia a 1 año o podemos dar la mediana de supervivencia.

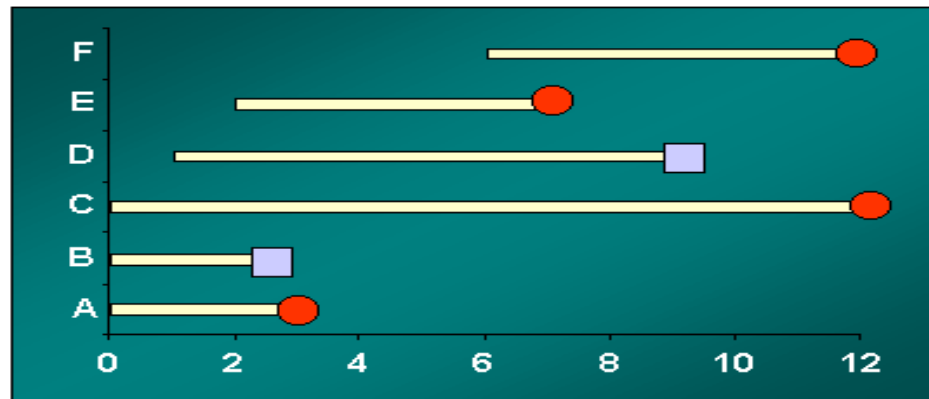
# 6- Resultados

## a) Supervivencia

**EVENTOS:** Muerte, progresión...

**DATOS CENSURADOS:** Pacientes que al final del estudio no han presentado el evento de interés

- ✓ Fin del estudio.
- ✓ Pérdida.
- ✓ Ocurrencia de un evento competitivo.

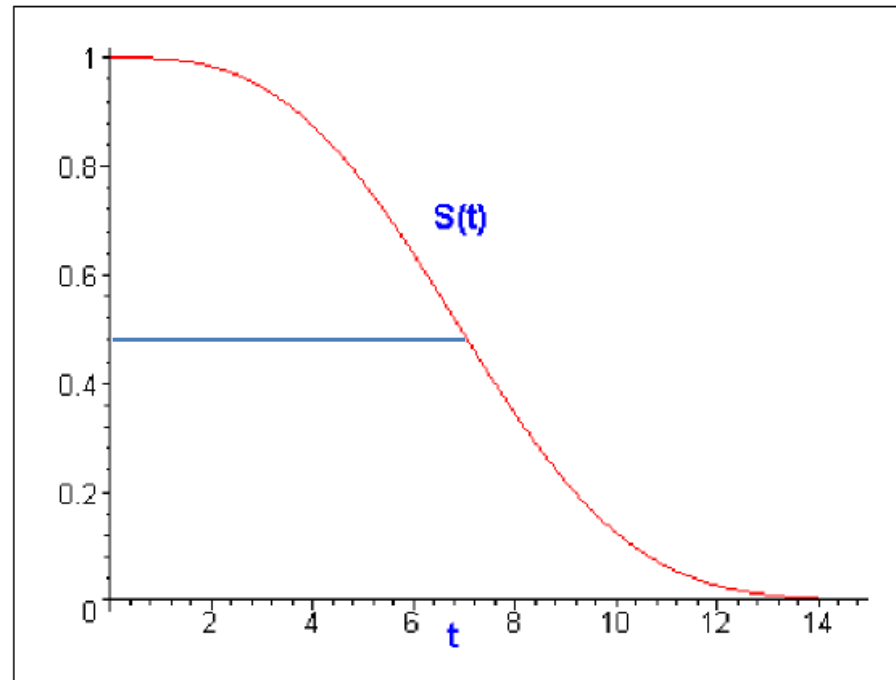


# 6- Resultados

## a) Supervivencia

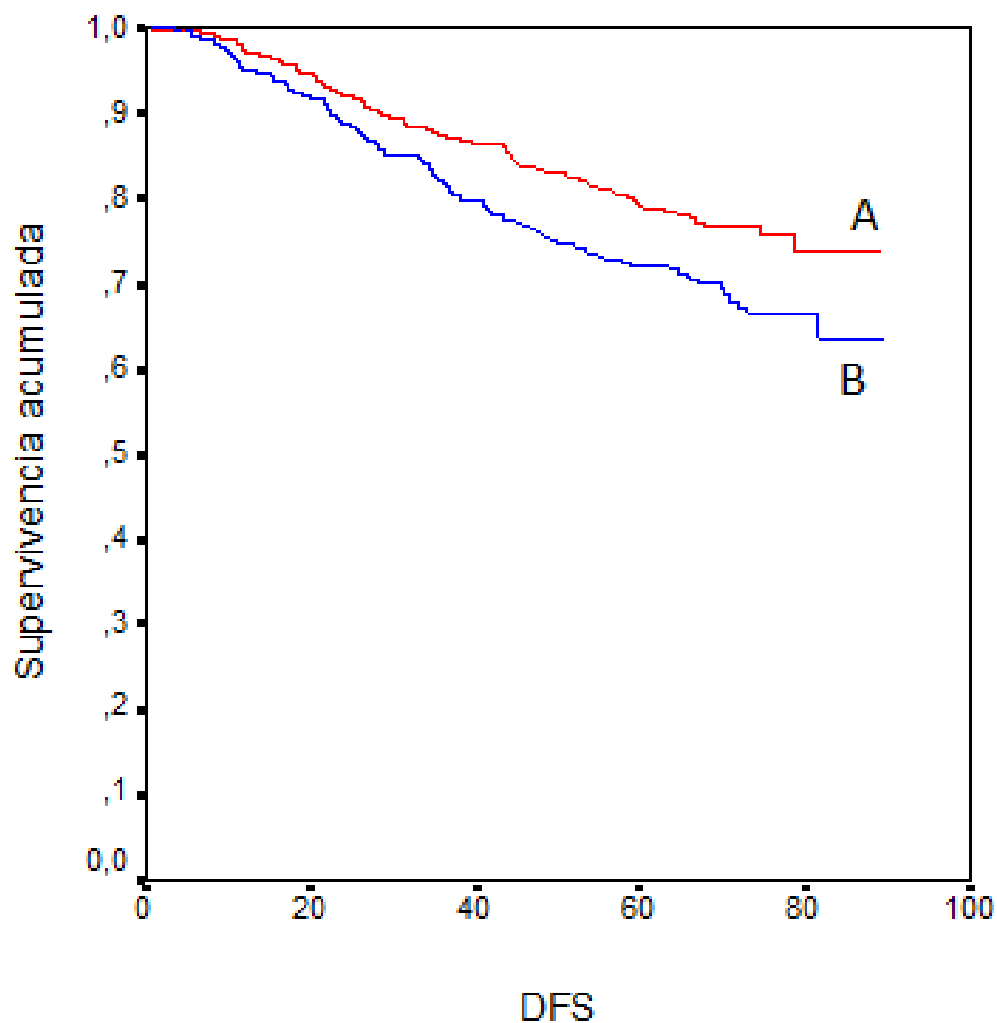
**Mediana de supervivencia** se define como el valor del tiempo con probabilidad de supervivencia igual a 0.50. Se estima con el percentil 50 de la distribución, que corresponde al primer tiempo con probabilidad acumulada de supervivencia  $S(t_i)$  igual o inferior a 0.5.

- *50% de los pacientes sobrevive al menos 7 meses.*
- *80% de los pacientes sobrevive al menos 4.8 meses.*
- *El porcentaje de pacientes que sobrevive 10 meses o más es del 12%*



# 6- Resultados

## a) Supervivencia



Las curvas de supervivencia permiten describir la evolución de un grupo de sujetos desde un estado inicial hasta un estado terminal (progresión, muerte,..) mediante la representación gráfica de las proporciones de supervivencia para diferentes intervalos de tiempo.

Las curvas de supervivencia representan la probabilidad de sobrevivir (no progresar) más allá de un tiempo  $x$  (eje horizontal). Una probabilidad acumulada de DFS a los 5 años igual a 0.72 en los pacientes B quiere decir que los pacientes de esta población tienen un 72% de probabilidades de estar “vivos” 5 años después del inicio de la enfermedad.



# 6- Resultados

## a) Supervivencia

- Análisis de Supervivencia:

Variables	N	Event	Censored	p-value (log-rank)	HR (IC95%)	p-value Cox
Rama de tratamiento						
A	406	91	315	0.0083	Ref.	0.009
B	420	128	292		1.4 (1.1-1.9)	

El haber recibido B multiplica el riesgo de recaída por 1.4 respecto a la rama A.

**SLE-60 meses** → Probabilidad de que un sujeto tarde un tiempo superior a 60 meses en recaer.

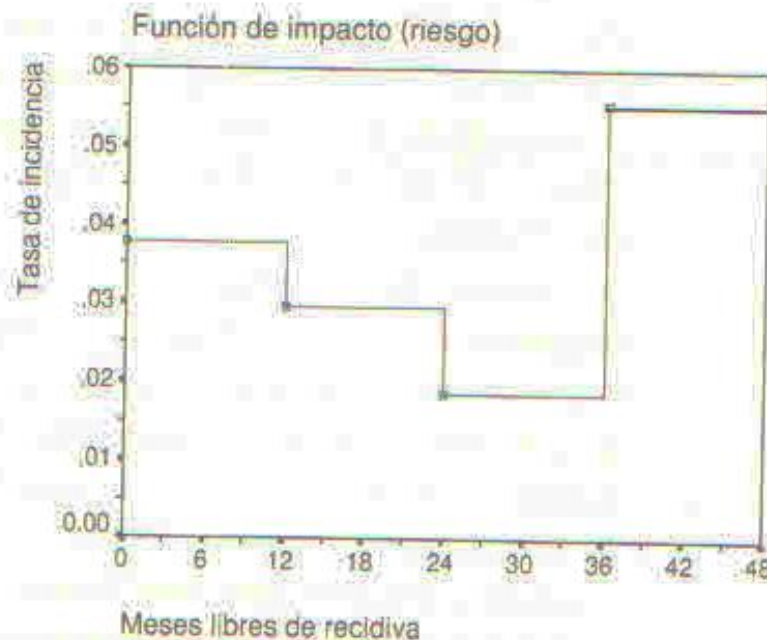
**p-log rank** → Es un test muy potente para detectar diferencias entre las funciones de supervivencia de 2 o mas brazos.

**HR (Razón de riesgos)** → Riesgo estimado de muerte (recaída) para los sujetos de la rama A respecto de la rama B (como es mayor que 1 los sujetos de la rama A tienen mas probabilidad de recaer que los de la rama B).

# 6- Resultados

## a) Supervivencia

- Hazard function o hazard rate=  $\lambda(t)$ . Esta es la probabilidad instantánea de muerte, es decir, la probabilidad de muerte en el instante  $t$  dado que ha vivido hasta el instante justo antes de  $t$ .



La tasa de eventos es alta durante el primer año, disminuye en los 2 siguientes años y presenta el máximo riesgo en los pacientes que han estado 3 años sin evento

# 6- Resultados

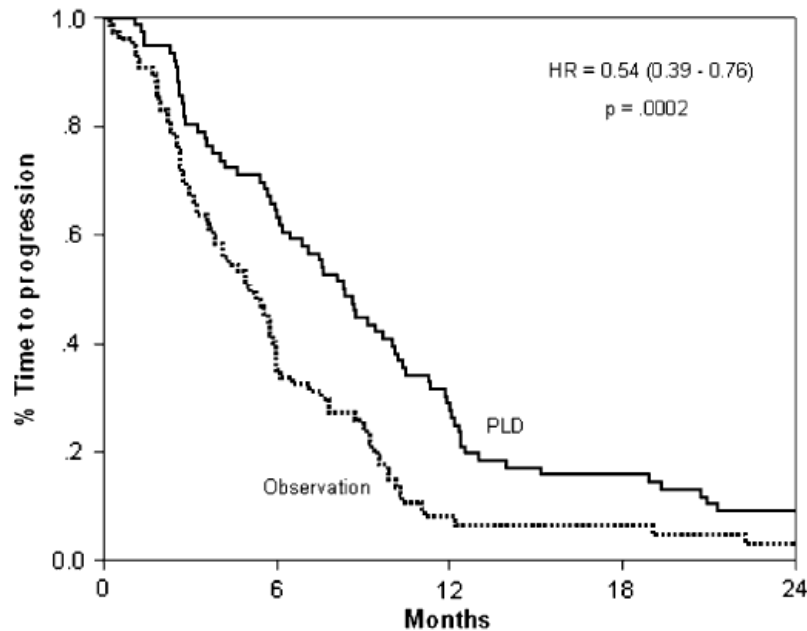
## a) Supervivencia

Breast Cancer Res Treat  
DOI 10.1007/s10549-010-0860-9

CLINICAL TRIAL

### Maintenance treatment with Pegylated liposomal doxorubicin versus observation following induction chemotherapy for metastatic breast cancer: GEICAM 2001-01 study

Emilio Alba · Manuel Ruiz-Borrego · Mireia Margelí · Álvaro Rodríguez-Lescure · Pedro Sánchez-Rovira · Amparo Ruiz · Jose Ramón Mel-Lorenzo · Manuel Ramos-Vázquez · Nuria Ribelles · Elisa Calvo · Antonio Casado · Antonia Márquez · David Vicente · José Angel García-Sáenz · Miguel Martín



No. at Risk	0	6	12	18	24
PLD	78	54	26	12	8
Observation	77	39	8	4	2

PLD significantly improved TTP by 3.3 months (8.4 vs. 5.1 months; hazard ratio [HR] = 0.54, 95% CI: 0.39 to 0.76, P = 0.0002) compared with Observation

**Número de expuestos en riesgo:** es el número de individuos que llegan vivos al intervalo menos la mitad de los que abandonan en ese intervalo.  
Estimación de los que tienen riesgo de tener evento en algún momento del periodo observado.

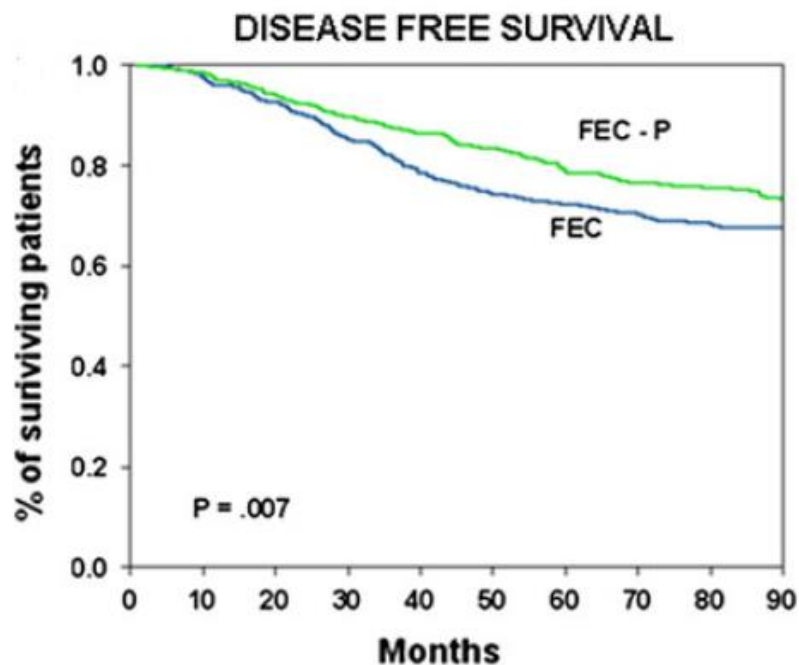
# 6- Resultados

Breast Cancer Res Treat (2010) 123:149–157  
DOI 10.1007/s10549-009-0663-z

CLINICAL TRIAL

## Molecular predictors of efficacy of adjuvant weekly paclitaxel in early breast cancer

Miguel Martín · Álvaro Rodríguez-Lescure · Amparo Ruiz · Emilio Alba · Lourdes Calvo · Manuel Ruiz-Borrego · Ana Santaballa · César A. Rodríguez · Carmen Crespo · Mar Abad · Severina Domínguez · Jesús Florián · Cristina Llorca · Miguel Méndez · María Godes · Ricardo Cubedo · Adolfo Murias · Norberto Batista · María José García · Rosalía Caballero · Enrique de Álava



No. at Risk	0	10	20	30	40	50	60	70	80	90
FEC	632	629.5	616	599	581	560	536	496.5	344	167.5
FEC - P	614	609	602	589	573.5	562	538	489.5	372	177.5

### Results

At a median follow-up of 7 years, a total of 364 DFS events have been recorded (205 in the FEC arm and 159 in the FEC-P arm (hazards ratio: 0.75; 95% CI: 0.61–0.93;  $P = 0.007$ ). The 7-year DFS was significantly superior in the FEC-P arm (75% vs. 68%;  $P = 0.007$  log-rank test). There have been 217 deaths (125 in the FEC arm and 92 in the FEC-P arm (hazards ratio: 0.74; 95% CI: 0.56–0.96;  $P = 0.026$ ). The 7-year OS was also significantly superior in the FEC-P arm (84 vs. 79%;  $P = 0.025$  log-rank test) (Fig. 1a, b).

**Probabilidad acumulada de supervivencia en un tiempo concreto.**

# 6- Resultados

## b) Análisis Uni/Multivariante

**Análisis Univariante:** Conjunto de técnicas que permiten estudiar si una variable (Y) depende de otra variable (X).

**Análisis Multivariante:** Conjunto de técnicas que permiten analizar la relación entre una variable dependiente (Y) y varias variables independientes ( $X_1, X_2, X_3, \dots$ ).

Y	Regresión
Binomial (éxito/fracaso)	Logística
Cuantitativa	Lineal
Tiempo de supervivencia	COX (riesgos proporcionales)

Las variables independientes pueden ser tanto cuantitativas como categóricas.

# 6- Resultados

## b) Análisis Uni/Multivariante

Para la construcción de un modelo de regresión multivariante se siguen los siguientes pasos:

1. Se realiza un análisis univariante para cada una de las variables independientes que estemos interesados en incluir en el análisis multivariante. Estas variables serán determinadas por el clínico. Con este análisis queremos comprobar la influencia individual de cada variable independiente sobre la variable dependiente.
2. Se seleccionan las variables con un p-valor pequeño (menor de 0,20) y que sean independientes entre sí (Ej: No podemos coger el Subtipo y el HER2).
3. Se pregunta al clínico si se tiene que meter alguna interacción.

### INTERACCIÓN

Se da **interacción** entre 2 variables independientes si, el efecto de una es diferente según los valores de la otra en la relación con la variable dependiente.

*Ejemplo: ¿la relación entre tabaco e Infarto cambia de forma significativa según la edad de los sujetos? Se dará interacción cuando la relación entre el tabaco y el infarto es diferente en los diferentes grupos de edad. Por ejemplo la relación es débil entre los menores de 50 años (RR=1,2) y muy fuerte entre los mayores de 50 años (RR=3,6).*

# 6- Resultados

## b) Análisis Uni/Multivariante

4. Se incluyen todas las variables seleccionadas en el paso 2 y 3 y se genera un modelo multivariante. Se utilizan varias técnicas (Forward/ Backward/ Stepwise) por lo que se puede obtener más de un modelo multivariante aceptable estadísticamente.
5. Se selecciona el mejor de los modelos según criterios clínicos.

IMPORTANTE: Se debe realizar un adecuado diagnóstico del modelo de regresión para estudiar si cumple adecuadamente las condiciones de aplicación.

**TODOS LOS PROCESOS DE MODELIZACIÓN COMPORTAN UNA CONTINUA INTERACCIÓN ENTRE EL ESTADÍSTICO Y EL INVESTIGADOR**

# 6- Resultados

## b) Análisis Uni/Multivariante - Regresión Logística

### Análisis univariante

Univariate Logistic Regression tpCR	p-value	OR (IC95%)
Treatment (EC-L+D as reference)	<b>0.0381</b>	2.783 (1.058 – 7.323)
RE_central (positive as reference)	<b>0.0144</b>	3.373 (1.274 – 8.931)
RP_central (positive as reference)	<b>0.0191</b>	3.375 (1.221 – 9.330)
RH_central Cate (RE+/RP+ as reference)	<b>0.0315</b>	
RE+/RP-	0.9963	2.104 (0.575 – 7.692)
RE-/RP-	0.0258	4.451 (1.460 – 13.569)
Ki67 (cuantitativa)	<b>0.0690</b>	1.019 (0.999 – 1.039)
PTEN (cuantitativa)	<b>0.0935</b>	0.890 (0.776 – 1.020)
pERK (cuantitativa)	0.2873	1.004 (0.997 – 1.011)
ERK (cuantitativa)	0.5638	0.996 (0.984 – 1.009)

### Análisis multivariante

Odds Ratio Estimates				Pr > ChiSq
Effect	Point Estimate	95% Wald Confidence Limits		
tratamiento EC-H+D vs EC-L+D	2.953	1.070	8.154	0.0366
RE_central Negativo vs Positivo	3.557	1.292	9.795	0.0141



# 6- Resultados

## b) Análisis Uni/Multivariante - Regresión de Cox

La **Regresión de Cox** permite analizar la relación entre la variable tiempo hasta que ocurre un evento (variable dependiente) y otras variables Xs, que pueden ser binarias, categóricas o continuas (variables independientes).

**Hazard Ratio (HR)** representa el riesgo de tener el evento de interés. Si la variable es categórica en cada grupo frente otro de referencia y si es continúa a medida que aumenta o disminuye en una unidad.

- Si el **HR es inferior a 1** esto indica que es un efecto protector.
- Si el **HR es superior a 1** indica que es un factor de riesgo.

# 6- Resultados

## b) Análisis Uni/Multivariante - Regresión de Cox

### Análisis univariante

**Table 2 Univariate Cox proportional hazards ratios and 95% confidence intervals in univariate analyses for association between EndoPredict, clinicopathological variables and metastasis-free survival**

Variables	HR (95% CI <sup>a</sup> )	P-value
EP	1.207 (1.134 to 1.285)	<0.0001
EPclin	1.916 (1.625 to 2.259)	<0.0001
Age	0.977 (0.960 to 0.993)	0.0065
Tumor size		0.0159 <sup>b</sup>
≤1	Reference value	
>1 to ≤2 cm	1.038 (0.457 to 2.357)	0.9290
>2 to ≤5 cm	1.919 (0.884 to 4.168)	0.0994
>5 cm	2.102 (0.762 to 5.798)	0.1511
Number of positive nodes		<0.0001 <sup>b</sup>
1 to 3	Reference value	
4 to 10	1.631 (1.085 to 2.451)	0.0187
>10	4.911 (3.022 to 7.979)	<0.0001
Grade		0.0233 <sup>b</sup>
G1	Reference value	
G2	2.263 (1.157 to 4.428)	0.0171
G3	2.883 (1.446 to 5.746)	0.0026
Unknown	1.844 (0.749 to 4.538)	0.1829
Treatment arm		0.6067 <sup>b</sup>
FEC	Reference value	
FEC-P	0.910 (0.635 to 1.304)	
ER (Allred score <sup>c</sup> )	0.950 (0.886 to 1.019)	0.1505
PR (Allred score <sup>c</sup> )	0.923 (0.874 to 0.974)	0.0033
Ki67 (quantitative)	1.017 (1.004 to 1.030)	0.0080

<sup>a</sup>EP: EndoPredict; EPclin: Combined molecular and clinical score; CI: Confidence interval; ER: Estrogen receptor; HR: Hazard ratio; MFS: Metastasis-free survival; PR: Progesterone receptor. Number of patients included in the analyses is 555. <sup>b</sup>P-value of the variables' overall effect on MFS. <sup>c</sup>Allred et al. [18].

- ✓ Se comprueba la influencia (individual) de cada variable en la variable tiempo hasta un evento.

*Clinical validation of the EndoPredict test in node-positive chemotherapy-treated. ER+/HER2- breast cancer patients: results from the GEICAM/9906 trial. Martín M. et al. (Breast Cancer Research. April 12, 2014).*

# 6- Resultados

## b) Análisis Uni/Multivariante - Regresión de Cox

### Análisis multivariante

**Table 3 Multivariate Cox proportional hazards ratios and 95% confidence intervals for the association between EndoPredict, selected clinicopathological variables and metastasis-free survival**

Variables	HR (95% CI <sup>a</sup> )	P-value
EP	1.126 (1.041 to 1.219)	0.0031
<b>Number Positive Nodes</b>		<0.0001 <sup>b</sup>
1 to 3	Reference value	
4 to 10	1.420 (0.932 to 2.166)	0.1030
>10	3.605 (2.102 to 6.185)	<0.0001
Age	0.983 (0.966 to 1.001)	0.0628
<b>Tumor size</b>		0.6631 <sup>b</sup>
≤1 cm	Reference value	
>1 to ≤2 cm	0.789 (0.343 to 1.816)	0.5774
>2 to ≤5 cm	1.042 (0.466 to 2.331)	0.9196
>5 cm	0.880 (0.301 to 2.577)	0.8159
<b>Treatment arm</b>		0.9331 <sup>b</sup>
FEC	Reference value	
FEC-P	1.016 (0.697 to 1.482)	
<b>Grade</b>		0.4650 <sup>b</sup>
G1	Reference value	
G2	1.662 (0.830 to 3.329)	0.1519
G3	1.589 (0.747 to 3.377)	0.2290
Unknown	1.198 (0.470 to 3.052)	0.7051
ER (Allred score <sup>c</sup> )	0.980 (0.903 to 1.063)	0.6207
PR (Allred score <sup>c</sup> )	0.965 (0.902 to 1.032)	0.2947
Ki67 quantitative	1.001 (0.986 to 1.016)	0.8982

<sup>a</sup>CI: Confidence interval; EP: EndoPredict; EPclin: Combined molecular and clinical score; ER: Estrogen receptor; HR: Hazard ratio; MFS: Metastasis-free survival; PR: Progesterone receptor. <sup>b</sup>P-value of the variable's overall effect on MFS. <sup>c</sup>Allred et al. [18]. Multivariate analyses included 534 patients and 116 events.

- ✓ Las variables a incluir en el modelo multivariante deben ser seleccionadas con criterio clínico y estadístico (tras el análisis univariante).
- ✓ Mediante métodos de selección de variables (stepwise, backward, forward), elegimos aquellas variables que predecirán o estimarán mejor la variable tiempo hasta un evento.

*Clinical validation of the EndoPredict test in node-positive chemotherapy-treated. ER+/HER2-breast cancer patients: results from the GEICAM/9906 trial. Martín M. et al. (Breast Cancer Research. April 12, 2014).*

# 6- Resultados

## b) Análisis Uni/Multivariante - Regresión de Cox

### Análisis univariante

Univariate Analysis					
Variables	N	Events	Censored	HR (CI 95%)	p-value Cox
<b>Treatment</b>					<b>0.0077</b>
A	566	58	508	1.779 (1.165 – 2.717)	
B	581	34	547	Ref.	
<b>Hormonal Receptor</b>					<b>&lt;0.0001</b>
Negative	295	46	249	3.361 (2.205 – 5.121)	
Positive	807	41	766	Ref.	
<b>Her2-status</b>					0.1155
Negative	1028	78	950	Ref.	
Positive	60	9	51	1.743 (0.873 – 3.480)	
<b>Ki67</b>	1062	82	980	1.015 (1.004 – 1.027)	<b>0.0065</b>
<b>Age</b>	1147	92	1055	0.979 (0.958 – 1.000)	<b>0.0537</b>
<b>Histopathologic Type</b>					0.5322
Ductal	977	81	896	Ref.	

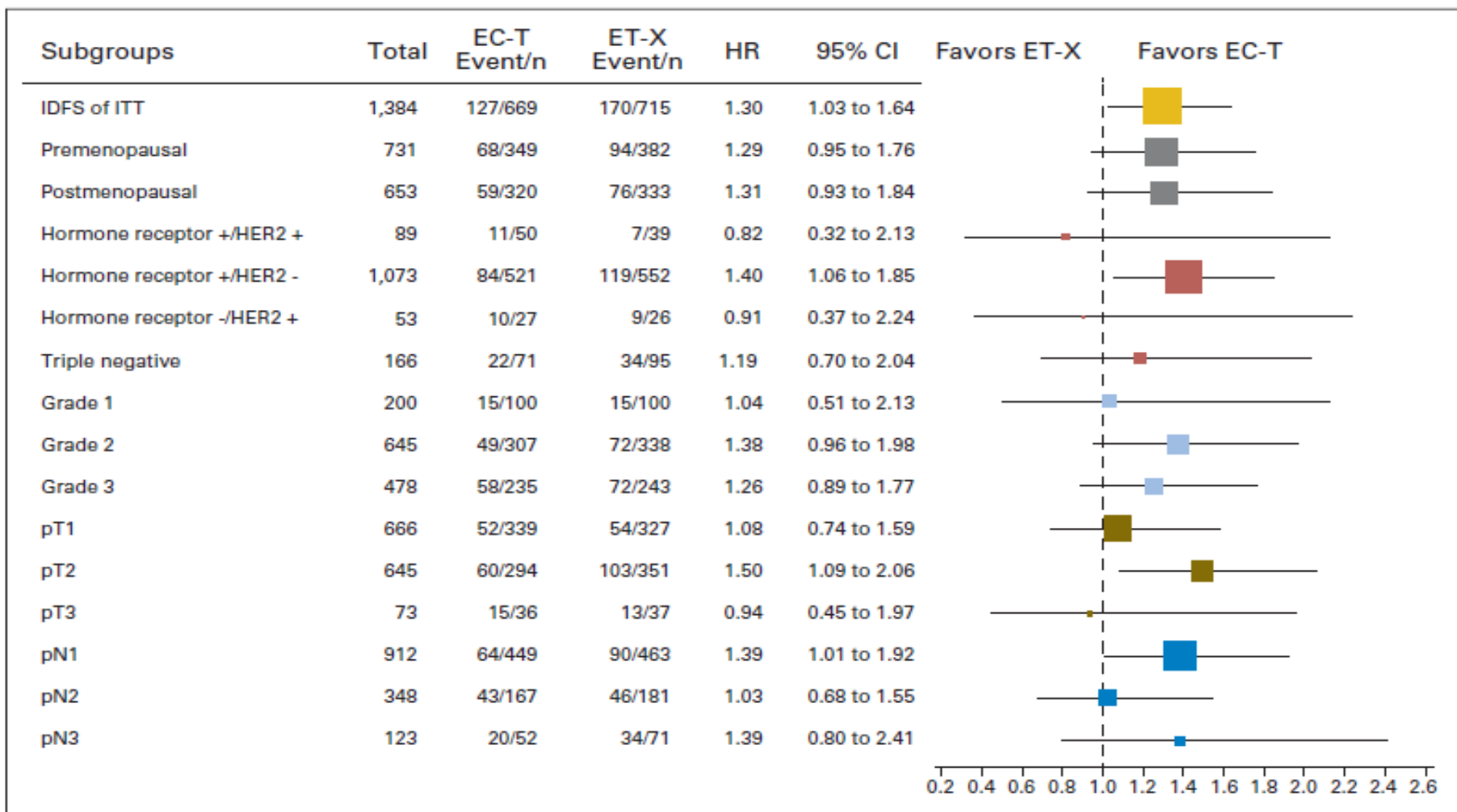
### Análisis multivariante

Multivariate Model for Disease-Free Survival			
Variables n=1009, events=83	p-value	HR	CI 95%
Treatment (A reference)			
B	0.5491	1.390	0.474 – 4.077
Subtype (St. Gallen 2013) (TN reference)	0.0159		
Luminal A	0.0006	0.043	0.007 – 0.260
Luminal B1	0.0015	0.127	0.035 – 0.453
Luminal B2	0.0681	0.304	0.085 – 1.093
HER2	0.04735	0.672	0.226 – 1.995
Subtype St. Gallen 2013*Treatment	0.0910	0.781	0.587 – 1.040

# 6- Resultados

## c) Forest Plot

### Análisis de subgrupos- FOREST PLOT



**Fig 3.** Subgroup analysis of invasive disease-free survival (IDFS). C, cyclophosphamide; E, epirubicin; HER2, human epidermal growth factor receptor 2; HR, hazard ratio; ITT, intention to treat; pN, pathologic node involvement (at surgery); pT, pathologic tumor size (at surgery); T, docetaxel; X, capecitabine.

- 1- **“Clinical Trials in Oncology”**. Chapman & Hall /CRC
- 2- “Fundamentos de Diseño y Estadística. UD4: Descripción de datos de supervivencia”. Domenech, J.M. (2004). Signo.
- 3- “Análisis de Supervivencia y modelos de regresión proporcionales de Cox”. Domenech, J.M. (2005). Signo.
- 4- “Experimental Design”. Kirk, R. E. Brooks/Cole.
- 5- “An introduction to generalized linear models”. Annette Dobson. Chapman & Hall.
- 6- “Applied Survival Analysis. Regresion modeling of time to event data”. Hosmer & Lemeshow. John Wile & Sons.

1- Apuntes de Bioestadística. Unidad de Bioestadística Clínica del Hospital Ramón y Cajal.  
[http://www.hrc.es/investigacion/bioest/M\\_doce\\_nte.html](http://www.hrc.es/investigacion/bioest/M_doce_nte.html)

2- Cálculos de tamaños muestrales  
<https://stattools.crab.org>

Programa ENE 3.0- (Gratis).

**¡GRACIAS!**

