

**Departament de Farmacologia i de Terapèutica**

**Universitat Autònoma de Barcelona**

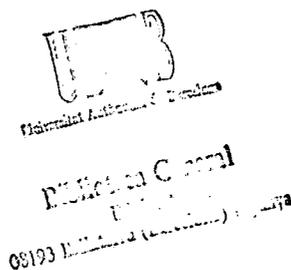
**Aspectos metodológicos y estadísticos en el  
diseño secuencial de ensayos clínicos.**

**Memoria presentada por  
Ferran Torres Benítez  
Grado de Doctor en Medicina**

**Barcelona, diciembre 1997**

DEPARTAMENT DE FARMACOLOGIA I DE TERAPÈUTICA  
UNIVERSITAT AUTÒNOMA DE BARCELONA

**Aspectos metodológicos y estadísticos en el diseño secuencial  
de ensayos clínicos.**



**Memoria presentada por**

**Ferran Torres Benítez**

**Para optar al grado de**

**Doctor en Medicina**

**Director:**

**Prof. Víctor Moreno Aguado**

**Tutor:**

**Prof. Josep Torrent Farnell**

**Barcelona, diciembre de 1997**

VICTOR MORENO AGUADO, Professor Titular del Departament de Pediatria, Ginecologia, Obstetrícia i Medicina Preventiva, com a Director i JOSEP TORRENT FARNELL, Professor Titular del Departament de Farmacologia i de Terapèutica, com a Tutor, ambdós de la Universitat Autònoma de Barcelona.

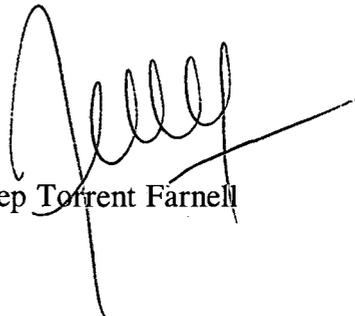
CERTIFIQUEN:

Que la Tesi Doctoral "**ASPECTES METODOLOGICS I ESTADÍSTICS EN EL DISENY SEQÜENCIAL D'ASSAIGS CLINICS**" presentada per **FERRAN TORRES BENITEZ** ha estat realitzada sota la seva direcció i tutoria i reuneix els requisits necessaris per la seva tramitació, presentació i posterior defensa davant el Tribunal corresponent.

I per que així consti i produeixi els efectes oportuns, signen aquest certificat a Barcelona, a 16 de desembre de 1997.



Victor Moreno Aguado



Josep Torrent Farnell

INTRODUCCIÓN .....	8
Infarto agudo de miocardio.....	8
Relevancia sanitaria del IAM. ....	8
Estrategias terapéuticas disponibles.....	9
Tratamiento antiagregante del IAM.....	10
Naturaleza del proceso de obstrucción coronaria .....	10
Fármacos antiagregantes.....	12
Aspirina.....	12
Tienopiridinas: Ticlopidina y clopidogrel .....	13
Otros antiagregantes .....	15
Triflusal.....	15
Justificación de la realización del estudio Triflusal en el Infarto de Miocardio (TIM) .....	17
Diseño metodológico .....	18
El ensayo clínico. Un poco de historia. ....	18
Etapas del ensayo clínico.....	19
Aspectos estadísticos .....	19
Diseño clásico.....	20
Análisis estadístico: evaluación del efecto comparado de los tratamientos .....	20
Tamaño de muestra.....	23
Mágnitud del efecto .....	23
Varianza .....	23
Errores estadísticos .....	24
Diferentes métodos para el cálculo de tamaño muestral.....	27
Diseños de equivalencia terapéutica.....	28
Razones para justificar estudios de equivalencia terapéutica .....	29
Metodología.....	29
Aproximación clásica por poder .....	30
Aproximación por intervalos de confianza y test doblemente unilateral de Schuirmann .....	31

## INDICE

---

Selección de la amplitud de los IC y del nivel de significación .....	34
Otras aproximaciones .....	35
Metodología secuencial .....	36
Justificación de los métodos secuenciales .....	36
Monitorización de la información.....	36
Análisis intermedios .....	37
Monitorización ingenua .....	37
Terminación prematura de ensayos clínicos.....	39
Ventajas y limitaciones de los diseños secuenciales .....	40
Aplicabilidad del diseño secuencial.....	41
Tipos de diseño secuencial .....	43
Historia de los métodos secuenciales.....	43
Uso y reconocimiento de los métodos secuenciales .....	44
Bases del diseño secuencial .....	46
Parametrización y medida del efecto .....	46
Para datos binarios.....	46
Para otros datos.....	47
Diseño estadístico clásico basado en Z y V.....	48
Ejemplo para datos binarios.....	49
Análisis convencional .....	49
Análisis mediante aproximación secuencial .....	50
Desarrollo estadístico*.....	50
Característica operante.....	53
Desarrollo estadístico*.....	53
Revisión de los diferentes métodos secuenciales .....	55
Reestimación del tamaño muestral .....	55
Métodos secuenciales por grupos .....	56
Desarrollo estadístico*.....	58
Funciones de gasto de alfa .....	59
Base estadística* .....	60
Diseños con funciones de gasto de $\alpha$ mediante EaSt .....	61

## INDICE

---

Métodos basados en límites continuos (basados en la función de verosimilitud).....	66
Otros contextos .....	68
Diseño .....	69
Modelos .....	71
Plan de inspecciones .....	75
Corrección por monitorización discreta.....	75
Análisis al finalizar el estudio.....	77
Nivel de significación .....	77
Estimación e intervalos de confianza.....	77
Estimador no sesgado de $\theta$ .....	78
Análisis estratificado y Análisis multivariante .....	79
“Underruning” y “Overruning” .....	80
Estimación del efecto absoluto de los tratamientos .....	81
Análisis de variables secundarias .....	82
Intervalos de confianza repetidos .....	84
Restricción estocástica.....	86
Base metodológica* .....	86
Métodos Bayesianos .....	88
Equivalencia terapéutica secuencial .....	90
OBJETIVOS .....	93
Objetivos principales .....	93
Objetivos secundarios .....	93
MATERIAL Y MÉTODO.....	94
Planificación general del diseño del estudio TIM .....	<b>¡Error! Marcador no definido.</b>
PLANIFICACIÓN DEL DISEÑO .....	94
Diferencia relevante .....	94
Características estadísticas.....	95
Estudio de alternativas al diseño clásico.....	96
Diseños contemplados .....	96
Requerimientos de tamaño muestral .....	98

## INDICE

---

Estimadores de las previsiones teóricas.....	98
Comprobación del cumplimiento de las especificaciones teóricas mediante simulaciones de estudios.....	101
Tamaño muestral.....	104
Ahorro en el número de pacientes .....	105
Valoración de las especificaciones estadísticas.....	105
Estudio TIM.....	107
Diseño .....	107
Estructuras y comités. Flujo de los datos y controles de calidad.....	107
Monitorización secuencial. Criterios de finalización del estudio .....	108
Análisis estadístico .....	109
BÚSQUEDA DEL TAMAÑO MUESTRAL ÓPTIMO .....	111
DISEÑOS TEÓRICOS .....	111
DISEÑO CLÁSICO.....	111
DISEÑOS SECUENCIALES ÓPTIMOS .....	111
FUNCIONES DE GASTO DE ALFA.....	112
MÉTODOS BASADOS EN LÍMITES CONTINUOS .....	115
OTROS DISEÑOS .....	116
RESTRICCIÓN ESTOCÁSTICA .....	116
INTERVALOS DE CONFIANZA REPETIDOS .....	116
MÉTODOS SECUENCIALES POR GRUPOS .....	116
EQUIVALENCIA TERAPÉUTICA .....	117
EQUIVALENCIA TERAPÉUTICA MEDIANTE METODOLOGÍA SECUENCIAL.....	118
MÉTODOS BAYESIANOS.....	118
ESTIMACIONES TEÓRICAS DE TAMAÑO MUESTRAL .....	119
GLOBALES.....	119
N mínima –Mín(t)t- y N máxima –Máx(t)t- .....	119
P50(t)t, P90(t)t y E(t)t.....	120
EN FUNCIÓN DE THETA.....	121
DESCRIPCIÓN DE LAS SIMULACIONES.....	130

## INDICE

---

GLOBALES.....	130
N mínima y N máxima.....	130
P50(t), P90(t) y E(t) .....	131
EN FUNCIÓN DE THETA .....	133
N mínima y N máxima.....	134
P50(t), P90(t) y E(t) .....	135
COMPROBACIÓN MEDIANTE SIMULACIONES.....	139
AHORRO DE LOS MÉTODOS SECUENCIALES.....	145
AHORRO: REDUCCIONES EN EL NÚMERO DE PACIENTES .....	145
Por thetas.....	147
AHORRO: PORCENTAJE DE ESTUDIOS QUE NO REQUIEREN EL TAMAÑO MUESTRAL CONVENCIONAL.....	155
GLOBAL .....	155
POR THETAS .....	156
CARACTERÍSTICA OPERANTE .....	158
ESTIMACIONES TEÓRICAS.....	158
ESTIMACIONES SIMULADAS.....	160
DISEÑO, MONITORIZACIÓN Y ANÁLISIS ESTADÍSTICO DEL ESTUDIO TIM164	
DISEÑO .....	164
MONITORIZACIÓN .....	166
PRE-REVISIÓN OFICIAL .....	166
Otras alternativas secuenciales .....	168
POST- REVISIÓN.....	172
Otras alternativas secuenciales .....	174
ANÁLISIS .....	175
Análisis principal .....	175
Análisis variables secundarias .....	177
Análisis pre-revisión.....	181
Planificación del diseño .....	183
Especificaciones del diseño .....	183
Selección de la variable principal .....	183

## INDICE

---

Diferencias clínicamente relevantes .....	184
Diferencias esperadas .....	185
Protección frente al error de tipo I .....	185
Protección frente al error de tipo II.....	186
Selección del modelo óptimo basándose en las especificaciones teóricas .....	188
Aplicabilidad de los métodos.....	188
Comprobación mediante simulaciones de los modelos seleccionados.....	191
Identificación de un error de el diseño.....	191
<b>DISEÑO, MONITORIZACIÓN Y ANÁLISIS ESTADÍSTICO DEL ESTUDIO</b>	
TIM .....	193
DISEÑO.....	193
MONITORIZACIÓN .....	193
PRE-REVISIÓN OFICIAL .....	193
Otras alternativas secuenciales .....	195
POST- REVISIÓN.....	196
Otras alternativas secuenciales .....	198
ANÁLISIS .....	199
Análisis principal .....	199
Análisis variables secundarias .....	201
Análisis pre-revisión .....	203
CONCLUSIONES .....	205
Respecto a los objetivos principales .....	205
Búsqueda de un diseño metodológico-estadístico teórico óptimo.....	205
Comprobación de los supuestos teóricos previamente a la realización del estudio mediante simulaciones.....	206
Realización del diseño, monitorización del estudio y análisis estadístico de los resultados .....	206
Respecto a los objetivos secundarios.....	208
ANEXOS .....	209
DATOS TEÓRICOS PARA DISEÑOS SECUENCIALES .....	209
DATOS SIMULADOS PARA DISEÑOS SECUENCIALES.....	215

## INDICE

---

AHORRO.....	221
CARACTERÍSTICA OPERANTE .....	227
PROGRAMA UTILIZADO PARA LAS SIMULACIONES DE ESTUDIOS EN FORTRAN 90.....	233
PROGRAMAS EN SAS UTILIZADOS PARA LEER Y ANALIZAR LOS DATOS SIMULADOS.....	237
LECTURA DE DATOS SIMULADOS .....	237
ANÁLISIS DE LOS DATOS SIMULADOS.....	239
Bibliografía.....	205

## INTRODUCCIÓN

### *Infarto agudo de miocardio*

El infarto agudo de miocardio (IAM) consiste en la necrosis isquémica del músculo miocárdico debida a la interrupción aguda del flujo arterial coronario. La obstrucción coronaria es a su vez la consecuencia final de un proceso complejo y dinámico donde interactúan como factores etiopatogénicos la aterosclerosis, la lesión del endotelio arterial y la agregación plaquetaria, entre otros. Más allá del compromiso hemodinámico que pueda condicionar la pérdida de funcionalismo de la pared miocárdica necrosada, diversas complicaciones locales del infarto y diversas manifestaciones sistémicas que reflejan la extensión de las alteraciones arteriales subyacentes configuran una de las enfermedades más prevalentes y de mayor trascendencia sanitaria.

### *Relevancia sanitaria del IAM.*

El infarto de miocardio continúa siendo la principal causa de muerte en la mayoría de los países desarrollados, a pesar de las mejoras introducidas en las pasadas décadas en la prevención primaria y secundaria de la enfermedad. En España se ha estimado en torno a 27.000 el número de IAM que acontecen anualmente, con una tasa de mortalidad por IAM en la población entre 25 y 74 años en torno a 60/100.000<sup>1</sup>.

Prácticamente la mitad de las víctimas del infarto de miocardio fallecen antes de llegar al hospital. De los miles de pacientes que ingresan en un hospital con el diagnóstico de IAM, un 7 a 15% fallecen en el transcurso del ingreso hospitalario, y otro 7 a 15% fallece en el año siguiente<sup>2</sup>.

A pesar de la disponibilidad y el uso de medidas terapéuticas efectivas para limitar el tamaño de la necrosis miocárdica, los pacientes que experimentan un IAM están sometidos a un mayor riesgo de sufrir un segundo evento cardíaco en el periodo posterior al infarto (reinfarto, insuficiencia cardíaca o muerte súbita). De hecho, las

causas de muerte más frecuentes tras un infarto agudo de miocardio son la insuficiencia cardíaca, las arritmias, la rotura cardíaca o el reinfarto <sup>3</sup>. El accidente cerebral vascular acontece también con mayor frecuencia en pacientes convalecientes de un IAM.

### *Estrategias terapéuticas disponibles*

El tratamiento del infarto agudo de miocardio se centró inicialmente en el tratamiento o prevención de sus complicaciones inmediatas. Así, antes de la generalización del empleo de la trombolisis, los regímenes terapéuticos comprendían básicamente el uso de antianginosos y de antiarrítmicos. El desarrollo de técnicas diagnósticas para el estudio de la hemodinámica cardíaca condujo al desarrollo de los métodos terapéuticos semi-invasivos, que se acompañaron de una marcada mejoría en el pronóstico de la enfermedad. Posteriormente se han desarrollado múltiples estrategias farmacológicas destinadas tanto a limitar el tamaño de la lesión isquémica durante la fase aguda del infarto como a minimizar las consecuencias a corto y largo plazo de la lesión miocárdica y del trastorno trombogénico subyacente, las cuales han permitido mejorar globalmente el pronóstico de los pacientes con infarto agudo de miocardio <sup>4</sup>.

Actualmente, el tratamiento del IAM se centra en tres aspectos:

El tratamiento agudo; destinado por una parte a limitar el área de necrosis y a reestablecer precozmente el flujo coronario (trombolisis, angioplastia y cirugía derivativa), y por otra parte a evitar o limitar las complicaciones inherentes a la isquemia aguda, desde el dolor de origen miocárdico hasta la insuficiencia cardíaca, la hipotensión o los trastornos del ritmo.

La terapéutica adyuvante, que pretende evitar las complicaciones precoces y tardías del IAM. A partir de extensos estudios se ha establecido claramente el papel beneficioso de los bloqueantes beta-adrenérgicos como reductores del riesgo cardiovascular en el post-infarto. Las ventajas del uso de los inhibidores del enzima conversor de la angiotensina han sido puestos recientemente en evidencia

como consecuencia del creciente entendimiento de los procesos implicados en los mecanismos de los riesgos secundarios a un infarto agudo de miocardio. La importancia de los agentes hipolipemiantes, en particular de los inhibidores de la HMGCoA debe tenerse también en consideración, tras la demostración de sus beneficios reflejados en un aumento de la supervivencia y en una reducción de la incidencia de reinfarto y de los requerimientos de cirugía coronaria derivativa tras un primer infarto de miocardio. Por contra, la anticoagulación con dicumarínicos se reserva para casos seleccionados. Otros fármacos, como los antagonistas de los canales del calcio, y especialmente las dihidropiridinas de vida media corta, no se consideran como tratamientos de primera línea en estos pacientes. Asimismo los antiarrítmicos en general, a excepción de la amiodarona se consideran relativamente contraindicados <sup>5</sup>.

La terapéutica específica del proceso trombótico causante del infarto; es decir, el tratamiento destinado a prevenir a corto y a largo plazo la reobstrucción arterial (tratamiento farmacológico antiagregante).

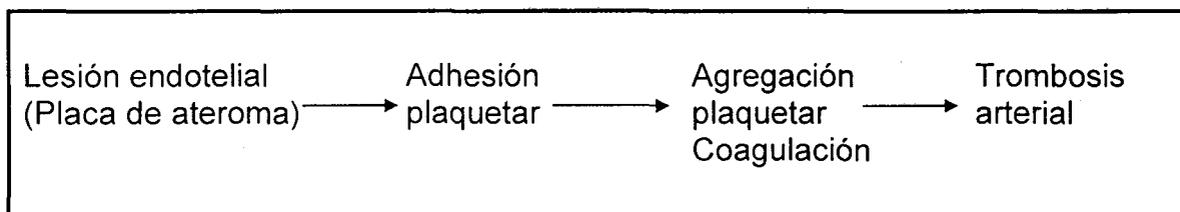
### ***Tratamiento antiagregante del IAM***

#### *Naturaleza del proceso de obstrucción coronaria*

La trombosis coronaria que subyace en un infarto de miocardio es un proceso complejo que se inicia con la lesión de la pared vascular arterial. La lesión de la pared desencadena el proceso hemostático, que se inicia típicamente con la adhesión de las plaquetas al endotelio dañado o a las estructuras subendoteliales expuestas (proceso de adhesión). Simultáneamente las proteínas de la fase fluida del plasma reaccionan con el endotelio e inician la activación por contacto de la coagulación. Los tejidos expuestos o los macrófagos que se encuentran en la matriz extracelular del vaso exponen factor tisular o tromboplastina a la sangre disparándose de esa

## INTRODUCCIÓN

---



forma la fase extrínseca de la coagulación. Las plaquetas, que han formado una monocapa protectora que recubre la lesión, liberan diversos mediadores que, a su vez, activan procesos destinados a la agregación plaquetar, a la deposición de fibrina y a la formación final de un trombo hemostático (procesos de agregación, contracción y secreción plaquetaria, y activación de la coagulación). Finalmente la oclusión completa de la luz arterial compromete el flujo coronario y condiciona la isquemia y posterior muerte celular del territorio miocárdico dependiente del vaso obstruido.

En condiciones de flujo arterial se producen predominantemente trombos ricos en plaquetas, mientras que en la trombosis venosa se producen predominantemente trombos que son ricos en fibrina y en células sanguíneas. Por ello los fármacos utilizados en el tratamiento y la prevención de la enfermedad trombótica arterial o venosa difieren en función de la composición predominante del trombo. Así el tratamiento y la prevención de las trombosis venosas se centra en disminuir la actividad de la coagulación, mientras que el tratamiento y la prevención de las trombosis arteriales está dirigido principalmente a modificar la adhesión y agregación plaquetar <sup>6-8</sup>.

La intervención terapéutica sobre las plaquetas tiene en cuenta los diversos estímulos y procesos que producen la activación plaquetar. Estos diversos mecanismos de activación (Tromboxano A<sub>2</sub>, colágeno endotelial, trombina, ADP, serotonina), principalmente a través de un mecanismo mediador común que consiste en una elevación del calcio intracelular, derivan en tres efectos principales: la secreción de los gránulos intracelulares de la plaqueta, la exposición de receptores de superficie para proteínas plasmáticas (fibrinógeno y factor de Von Willebrand) y alteraciones

de la estructura lipídica de la membrana plaquetaria, que inducen la aceleración de la coagulación plasmática.

La activación plaquetaria y sus efectos están modulados por un cierto número de sustancias reguladoras de las cuales la más importante es el AMP cíclico. Éste estimula una bomba ATPasa calcio/magnesio que disminuye la concentración citoplasmática de calcio en la plaqueta, bloqueando de esta manera los efectos de los agonistas plaquetarios. La prostaciclina endotelial estimula la enzima adenilciclasa encargada de la producción de AMPc; la enzima fosfodiesterasa cataliza su degradación. En concentraciones suficientes el AMPc inhibe no solo la agregación plaquetaria, la secreción y el cambio de forma de la plaqueta, sino también su adhesión al sustrato. Otros compuestos cuyo papel no está completamente esclarecido en la activación plaquetaria incluyen una ADPasa de la superficie de la célula endotelial y la trombomodulina, un potente inhibidor de la trombina que se encuentra en la superficie luminal del endotelio <sup>9</sup>.

### *Fármacos antiagregantes*

#### **Aspirina**

El ácido acetilsalicílico o aspirina es probablemente el fármaco más conocido y uno de los más utilizados mundialmente. Destacadas inicialmente sus propiedades analgésicas, antipiréticas y antiinflamatorias, en las últimas 3 décadas el interés que este fármaco centenario ha suscitado se ha ido focalizando en sus propiedades antiagregantes.

La aspirina induce un déficit funcional persistente en las plaquetas, clínicamente traducido en una prolongación del tiempo de sangría. Dicho déficit se debe principalmente a que la aspirina produce en las plaquetas una inhibición irreversible de la enzima ciclooxigenasa mediante acetilación del enzima, e impide así la síntesis del tromboxano A<sub>2</sub>, un potente inductor de la agregación plaquetar.

No obstante, existen otros mecanismos protrombóticos dependientes de las plaquetas que no son completamente inhibidos por el ácido acetilsalicílico, como por ejemplo la secreción de los gránulos alfa en respuesta al ADP y otros agonistas, la unión del fibrinógeno o la potenciación de la adhesión y de la agregación plaquetar inducida por adrenalina.

La aspirina se considera el tratamiento antiagregante de referencia en la prevención secundaria del infarto de miocardio y en determinadas formas de enfermedad vascular cerebral. En la prevención secundaria del IAM se han realizado diversos ensayos clínicos que conjuntamente han evaluado más de 20.000 pacientes con antecedentes de IAM. El análisis conjunto de todos los estudios <sup>10</sup> muestra una prevención en los pacientes tratados con aspirina respecto de los tratados con placebo de 13 muertes por 1000 pacientes tratados en 1 año (reducción relativa del 15%), 18 reinfartos no mortales por 1000 pacientes tratados en 1 año (reducción relativa del 31%), y 6 accidentes vasculares cerebrales no mortales por 1000 pacientes tratados en 1 año (reducción relativa del 39%). Hasta la fecha ningún otro antiagregante ha demostrado una clara superioridad en eficacia a la aspirina en la prevención secundaria del IAM.

Por lo que refiere a su seguridad, la aspirina presenta un perfil bien conocido de efectos indeseables, entre los que destacan los efectos adversos sobre el tracto gastrointestinal y aquellos derivados de la prolongación del tiempo de sangría e inhibición del funcionalismo plaquetario. Así, el sangrado gastrointestinal y la hemorragia cerebral son los dos efectos adversos graves que se contraponen a los beneficios de la aspirina <sup>11</sup>. Mientras la toxicidad gastro-intestinal parece estar relacionada con las dosis elevadas, el aumento del riesgo de sangrado acontece incluso a dosis muy bajas <sup>6,7</sup>.

### **Tienopiridinas: Ticlopidina y clopidogrel**

Las tienopiridinas actúan interfiriendo de forma selectiva y específica con la activación de las plaquetas mediada por ADP, al inhibir la secreción de los gránulos alfa que contienen proteínas adhesivas como el fibrinógeno, factores mitogénicos, y otros. De este modo impiden la exposición del lugar de unión del fibrinógeno del complejo del receptor GPIIb/IIIa, si bien no interfieren directamente con el mismo. Causan una inhibición de la actividad plaquetaria semejante a la producida por aspirina, si bien no interfieren con el metabolismo del ácido araquidonico. Asimismo inhiben de forma selectiva la expresión, ocupación o función de un subtipo de receptor de ADP de la membrana plaquetaria que media la inhibición de la adenilatociclasa, de modo que el resultado es una mayor presencia de AMPc intraplaquetario. Su mecanismo de acción, no obstante, no está plenamente aclarado.

Las tienopiridinas son inactivas “in vitro”, precisan de 3 a 5 días de tratamiento para iniciar su acción “in vivo” y no alcanzan su efecto máximo sobre el tiempo de sangría hasta varios días después. El efecto farmacológico antiagregante de la ticlopidina persiste hasta 10 días después de finalizar su administración y va desapareciendo de forma paralela al tiempo de supervivencia de las plaquetas. Éstas características condicionan que no se trate de productos adecuados para la consecución de un efecto antiagregante inmediato.

La ticlopidina ha mostrado en diversos ensayos clínicos su eficacia en la prevención secundaria del accidente vascular cerebral, del infarto de miocardio y de la vasculopatía arterial periférica. En ningún caso su eficacia ha sido menor que la de aspirina. No obstante produce neutropenia con una frecuencia en torno al 2,5% de los pacientes tratados, que si bien suele tener un carácter reversible, en un tercio de los casos puede llegar a ser severa. Este hecho condiciona que su uso esté limitado a pacientes en los que la aspirina se encuentra contraindicada, así como obliga a la realización de determinaciones hematológicas seriadas durante los primeros meses del tratamiento <sup>11</sup>.

## INTRODUCCIÓN

---

El clopidogrel es un fármaco estrechamente relacionado con la ticlopidina. La principal ventaja del clopidogrel frente a la ticlopidina radica en que no aporta un incremento sustancial del riesgo de padecer neutropenia. En un gran ensayo clínico que ha incluido cerca de 20.000 pacientes se ha observado una eficacia discretamente mayor del clopidogrel respecto a la aspirina en la prevención de acontecimientos isquémicos, en particular de accidentes vasculares cerebrales, si bien está en discusión que dicha ventaja en términos absolutos posea relevancia clínica. (CAPRIE Steering Committee)

### **Otros antiagregantes**

El dipiridamol es un derivado pirimidínico con acción antiagregante y vasodilatadora que actúa sobre la concentración plaquetaria de AMPc mediante inhibición de la fosfodiesterasa plaquetaria y mediante inhibición de la reutilización del ADP. Se especula que posee una acción sinérgica con la aspirina, si bien éste particular no se ha evidenciado en ensayos clínicos.

La sulfipirazona se utilizaba inicialmente como úricosúrico. Sus propiedades antiagregantes fueron estudiadas posteriormente, y existen ensayos clínicos que evaluaron su eficacia en prevención secundaria de eventos cardiovasculares.

(Inhibidores específicos del receptor GP IIb/IIIa)

### **Triflusal**

El triflusal es un antiagregante plaquetario relacionado estructuralmente con la aspirina. Experimentalmente se han evidenciado determinadas diferencias en la farmacocinética y mecanismo de acción de ambos fármacos que indican que podrían no ser equivalentes en cuanto a efectividad clínica y perfil de seguridad y tolerabilidad.<sup>12,13</sup>

## INTRODUCCIÓN

---

Al igual que la aspirina, el triflusal es un inhibidor irreversible del enzima ciclooxigenasa, y en consecuencia bloquea la formación de tromboxanos y el proceso de la agregación plaquetar <sup>14</sup>. En términos relativos el triflusal es menos potente que la aspirina como inhibidor de la actividad de la ciclooxigenasa, ya que para producir un grado determinado de inhibición de ésta enzima son precisas dosis 10 a 20 veces superiores a aquellas de aspirina que evidencian un efecto semejante. El principal metabolito intermediario del triflusal, el ácido 3-hidroxi-4 trifluoro- metilbenzoico (HTB), cuya semivida de eliminación es prolongada, posee a su vez actividad inhibidora de la ciclooxigenasa.

Por otra parte, a diferencia de la aspirina, el triflusal y su metabolito activo se comportan “in vitro” como inhibidores de la fosfodiesterasa a concentraciones similares a las terapéuticas, por lo que incrementan la concentración de AMP cíclico en las plaquetas y en el endotelio vascular e inhiben así la movilización del calcio intracelular que media la activación plaquetar <sup>13</sup>. Ésta propiedad se traduce en un mecanismo de efecto adicional a la inhibición de la ciclooxigenasa, y confiere al triflusal un perfil farmacodinámico diferencial de la aspirina. Asimismo se ha observado otro posible mecanismo de acción diferencial mediado por óxido nítrico <sup>15</sup>.

El tratamiento con aspirina origina normalmente una prolongación del tiempo de sangría que se traduce clínicamente en un incremento del riesgo de hemorragia <sup>16</sup>. En los estudios clínicos realizados con triflusal no se ha observado una prolongación del tiempo de sangría, lo que podría suponer un menor riesgo hemorrágico en los pacientes tratados con triflusal <sup>17</sup>.

## INTRODUCCIÓN

---

### *Justificación de la realización del estudio Triflusal en el Infarto de Miocardio (TIM)*

En el año 1992 se disponía de datos sobre diversos ensayos clínicos realizados con triflusal que sugerían que éste fármaco era eficaz en la prevención secundaria de acontecimientos trombóticos cardiovasculares. (Tabla)

---

Angioplastia coronaria transluminal percutánea

---

Angina inestable vs placebo, aparición de infarto no fatal 4,2 vs 12,3%

---

By-pass reestenosis 2,6 triflusal + dip, 10% aas + dip 12% pcb

---

trombosis venosa profunda en cirugía de cadera No diferencias con aas ni pcb

---

Enfermedad oclusiva arterial periférica 2 (mejoría en la distancia caminada y en la distancia caminada sin dolor de 64% vs 22.5% con placebo)(suma de síntomas reducidos en un 61% y un 27% pcb)

---

Accidente vascular cerebral (piloto: menor incidencia de eventos vasculares en el grupo triflusal)

---

Considerando:

El perfil farmacodinámico del producto que sustentaba la plausibilidad de un perfil farmacológico y de seguridad diferencial con la aspirina

Los datos preliminares de estudios realizados en enfermedad coronaria (angina inestable, angioplastia y by-pass aorto coronario), que sugerían diferencias en eficacia favorables a triflusal en la prevención de eventos cardiovasculares trombóticos.

La ausencia de información sobre la eficacia del triflusal en la prevención secundaria del IAM.

La escasez de datos sobre seguridad del triflusal procedentes de ensayos clínicos comparativos.

Se consideró pertinente realizar un estudio comparativo entre triflusal y aspirina en la prevención secundaria del infarto de miocardio.

### *Diseño metodológico*

#### *El ensayo clínico. Un poco de historia.*

Para evaluar la eficacia de un medicamento, actualmente hay un consenso general en la comunidad científica en reconocer al ensayo clínico como la herramienta estándar. Aunque haya distintas acepciones y casos particulares distintos, el ensayo clínico es un estudio experimental, en el que se compara un agente terapéutico con un tratamiento de referencia concurrente, aleatorizado y enmascarado con el fin de minimizar o evitar en la medida de lo posible cualquier tipo de sesgo.

Una buena revisión metodológica y práctica del ensayo clínico puede encontrarse en Bakke, Carné, et al. 1994<sup>18</sup>; otras lecturas útiles recomendadas son Pococok<sup>19</sup> y Spilker<sup>20</sup>.

En la evolución de la metodología científica, la realización del primer “experimento terapéutico” con la existencia de controles del cual se tiene información, data de 1747 con el estudio de James Lind para contrastar distintas dietas para el tratamiento del escorbuto.

Aunque desde entonces se pueden destacar distintos hitos históricos en el desarrollo metodológico de lo que hoy en día se considera el ensayo clínico, se acepta que no es hasta después de la Segunda Guerra Mundial cuando se desarrolla el primer ensayo clínico controlado y aleatorizado. En efecto, fue en 1946 cuando Sir Bradford Hill diseñó y llevó a cabo un ensayo clínico para evaluar la estreptomina en el tratamiento de la tuberculosis. Hoy en día, se puede decir que básicamente la

metodología es la misma y las diferencias se deben a cambios por requerimientos burocráticos y por la implementación de las normas de buena práctica clínica.

### *Etapas del ensayo clínico*

La realización del ensayo clínico es un proceso complejo en el cual se pueden diferenciar 5 etapas: la definición de los objetivos, el diseño, la realización o desarrollo de la parte experimental, el análisis estadístico, y finalmente la interpretación de los resultados.

### *Aspectos estadísticos*

La necesaria e inestimable aportación la estadística al desarrollo global de un ensayo clínico comprende la participación en la definición de los objetivos, selección de la variable principal, predeterminación del tamaño muestral, monitorización estadística y análisis definitivo. Asimismo puede y debe jugar un papel muy importante en la interpretación de los resultados de un ensayo en concreto o de estimaciones conjuntas de diversos estudios, nos estamos refiriendo a los metaanálisis.

### *Diseño clásico*

#### *Análisis estadístico: evaluación del efecto comparado de los tratamientos*

Aunque cronológicamente antes del análisis se deben establecer las hipótesis de trabajo y realizar la predeterminación del tamaño muestral, para facilitar la comprensión de esta tesis se abordarán primero los aspectos del análisis.

Dadas las características de la variable principal del estudio TIM, sólo se expondrá la metodología estadística específica para modelos con variables binarias, aunque la mayoría de métodos que se estudiarán pueden ser generalizables a otros tipos de variables.

Para evaluar el efecto de los tratamientos es aconsejable proceder sistemáticamente, realizando primero análisis simples y posteriormente los más complicados.

Si trabajamos una variable binaria (éxito fracaso), a partir de los datos del estudio, se debe construir una tabla de doble entrada, de 2x2, que resuma los resultados como la expresada en el siguiente cuadro:

	Tratamiento experimental	Tratamiento control	
Exito	a	b	a+b
Fracaso	c	d	c+d
	a+c	b+d	a+b+c+d

En un ensayo clínico, los marginales inferiores (a+c y b+d) y la suma de estos (a+b+c+d) quedan predeterminados por el tamaño muestral del diseño. Los marginales de la derecha (a+b y c+d) expresarán el efecto global en los pacientes sometidos a ambos tratamientos.

## INTRODUCCIÓN

---

Se puede calcular la incidencia acumulada de éxitos, o fracasos, en cada uno de los grupos de estudio al finalizar el seguimiento:

$$I_e = a / (a+c)$$

$$I_c = b / (b+d)$$

El efecto del tratamiento experimental, medido como una comparación respecto al tratamiento control puede expresarse en una escala absoluta o relativa.

La diferencia de incidencias en escala absoluta se calcularía a partir de la resta entre las dos incidencias:

$$I_e - I_c$$

Como medida relativa del efecto se puede emplear el riesgo relativo (RR) o el odds ratio (OR). El riesgo relativo se calcula como el cociente de la incidencia de éxitos en el grupo experimental respecto al grupo control.

$$RR = I_e / I_c$$

$$RR = a / (a+b) / b / (c+d)$$

Esta medida de riesgo se acostumbra a reservar para estudios epidemiológicos de cohortes, y como medida de asociación se puede calcular el odds ratio -OR- o razón de odds.

El odds expresa el cociente de la probabilidad de que ocurra un suceso respecto a la probabilidad de que no ocurra:

$$\text{odds} = p / (1-p)$$

## INTRODUCCIÓN

---

Por ejemplo, si la probabilidad de que al tirar un dado se obtenga un 6 es de  $1/6$  (0.167), el odds de obtener un 6 es de  $0.167/(1-0.167) = 0.2$  que se puede expresar en la forma 1:5 (de cada seis veces, 1 saldrá un 6 y en cinco no saldrá). Mientras que la probabilidad de un suceso puede estar comprendida entre 0 y 1, el odds de un suceso puede estar entre 0 e infinito. El odds de un suceso con probabilidad 0.5 es 1. Para evitar la asimetría de esta escala se utiliza comúnmente el logaritmo del odds, con ello se proyecta el espacio de probabilidad  $[0,1]$  a otro simétrico respecto al 0 con margen de menos infinito a más infinito. Esta transformación matemática es muy útil en estadística pues permite convertir la distribución de probabilidad de una proporción (generalmente binomial) en una distribución de características que se aproximan a la normal.

El OR en el caso del ensayo clínico, y aplicándolo a la tabla anterior, se expresaría como el cociente de las odds de la relación de éxitos respecto a fracasos para cada grupo de tratamiento.

$$\text{Odds en grupo experimental} = a / c$$

$$\text{Odds en grupo control} = b / d$$

$$\text{OR} = (a/c) / (b/d) = (a*d) / (c*b)$$

El odds ratio indica cuántas veces más riesgo de tener éxito posee la población expuesta al tratamiento experimental respecto a la expuesta al control, entendiendo la odds como una manera de medir el riesgo. De hecho el OR es un estimador del riesgo relativo, pues el valor de ambos parámetros se aproxima en el caso de enfermedades raras. Si invertimos el OR, éste indicará la estimación de riesgo de fracaso en lugar de éxito.

$$\text{OR}_{\text{fracaso}} = 1 / \text{OR}_{\text{éxito}} = (c*b) / (a*d)$$

Una explicación más detallada con otros estimadores de efecto puede encontrarse en Armitage y Berry <sup>21</sup> y Sentís  $\{\}$ .

### **Tamaño de muestra**

En un ensayo clínico, al igual que en cualquier estudio, son tres los parámetros de los que depende el tamaño de muestra necesario para detectar diferencias entre el efecto de dos o más grupos de tratamiento<sup>22</sup>:

La magnitud del efecto que se desee detectar.

La varianza de la medida del efecto, que en caso de datos binarios depende de las proporciones de éxito (o fracaso).

El nivel de protección frente a los errores de primera y de segunda especie ( $\alpha$  y  $\beta$ ).

### ***Mágnitud del efecto***

Debe definirse la escala en que interesa la magnitud del efecto. En general se emplea la diferencia entre las tasas de eventos en cada grupo o la razón entre las tasas de eventos.

Definida la escala, debe cuantificarse la magnitud de la diferencia que se pretenda detectar a la que denominaremos  $\theta_R$ . Ello se debe basar en conocimientos previos de la enfermedad y de los tratamientos.  $\theta_R$  debe elegirse de manera realista entre el mínimo riesgo detectable y el máximo riesgo que sería esperable encontrar.

En general, lo aconsejado es que  $\theta_R$  sea el mínimo valor de sea interesante detectar. Si sólo interesa detectar grandes diferencias no serán necesarios muchos pacientes, pero por otro lado no se dispondrá de suficiente poder estadístico para detectar valores pequeños de  $\theta_R$ .

### ***Varianza***

## INTRODUCCIÓN

---

La varianza de la diferencia o razón de proporciones depende de las propias proporciones. Cuanto más se acerca al 50% mayor es la variabilidad y mayor el tamaño muestral requerido.

$$\sigma^2 = p_0*(1-p_0) + p_1*(1-p_1)$$

### *Errores estadísticos*

Por último, el tamaño de muestra depende del nivel de protección que se desee tener frente a los errores de primera y segunda especie. Se recordará que en una prueba de hipótesis pueden cometerse dos tipos de error. En primer lugar, puede creerse que existe un mayor beneficio para un grupo, pero éste en realidad puede ser debido al azar. A este error se le denomina de tipo I o de primera especie y a la probabilidad de cometerlo  $\alpha$ . La protección frente a este error se basa en fijar un nivel de significación ( $\alpha$ ) bajo en el momento de tomar la decisión de rechazar la hipótesis nula.

A continuación se presenta una tabla que facilita la comprensión de las distintas posibilidades y sus consecuencias.

		Conclusión de la prueba	
		H <sub>0</sub>	H <sub>A</sub>
Realidad	H <sub>0</sub>	OK 1- $\alpha$	Riesgo para el paciente $\alpha$ (0.05)
	H <sub>A</sub>	Riesgo para el promotor $\beta$ (0.2)	OK 1- $\beta$

## INTRODUCCIÓN

---

En función de cómo se plantee la hipótesis nula ( $H_0$ ) y la alternativa ( $H_1$ ), el riesgo de primera especie se puede considerar a una o dos colas. En una prueba a una cola,  $H_1$  se plantea como: “el tratamiento experimental es mejor que el control” y  $H_0$  como “el tratamiento experimental no es mejor”

$H_0$ : Efecto  $\leq 0$

$H_1$ : Efecto  $> 0$

En una prueba a dos colas,  $H_1$  se plantea con dos posibilidades; “El tratamiento experimental es mejor o peor que el control” y  $H_0$  “el tratamiento experimental no es distinto que el control”

$H_0$ : Efecto = 0

$H_1$ : Efecto  $> 0$    ó   Efecto  $< 0$

En los ensayos clínicos raramente es aceptado por las entidades reguladoras estudios con hipótesis a una cola. Solamente son tolerados en algún caso concreto, como cuando el grupo de control es el placebo. En todo caso queda a criterio de los investigadores, y posteriormente a la aceptación de los Comités Éticos de Investigación Clínica e instituciones reguladoras. Al requerirse solamente protección para una de las colas, las pruebas unilaterales requieren menor tamaño muestral.

En la fase de diseño del estudio también es fundamental definir qué poder estadístico se pretende obtener. El poder estadístico es la probabilidad de detectar que los tratamientos en estudio son distintos cuando éstos realmente lo son. El error de aceptar la  $H_0$  cuando ésta es falsa se denomina de tipo II o de segunda especie. La probabilidad de cometer este error se denomina beta y es la complementaria del poder estadístico, habitualmente expresado en porcentaje:

$$\text{poder} = (1 - \beta) * 100$$

Así como clásicamente se acepta que alfa se debe fijar en un 5%, no existe un consenso respecto a qué niveles de  $\beta$  son tolerables. Se conoce que la mayoría de los

## INTRODUCCIÓN

---

estudios negativos lo son por causa de un bajo poder estadístico <sup>23</sup>, y muchos autores recomiendan diseñar los estudios con un poder mínimo del 80%.

El nivel de protección frente a los errores de primera y segunda especie se traduce en la fórmula para la estimación del tamaño de la muestra en un factor calculado en función de  $\alpha$  y  $\beta$ .

$$f(\alpha, \beta) = (U_\alpha + U_\beta)^2$$

Este factor puede consultarse en la siguiente tabla:

$\alpha$ (1 cola)	0,050	0,025	0,010	0,005	0,001
$\alpha$ (2 colas)	0,100	0,050	0,020	0,010	0,005
$\beta$ (1cola)					
0,200	6,183	7,849	10,036	11,679	15,460
0,100	8,564	10,507	13,017	14,879	19,113
0,050	10,822	12,995	15,770	17,814	22,421
0,025	12,995	15,366	18,372	20,573	25,505
0,010	15,770	18,372	21,647	24,031	29,339
0,005	17,814	20,573	24,031	26,540	32,104
0,001	22,421	25,505	29,339	32,104	38,198

Desde el punto de vista regulador se debe vigilar más el error de tipo I, puesto que implica el riesgo de permitir posteriormente la exposición de la población a un tratamiento ineficaz. Para el promotor parece poco eficiente admitir diseños con un poder bajo por las consiguientes repercusiones económicas. Aún con el poder habitual del 80%, de cada 5 estudios que se realicen con un fármaco eficaz, 1 no será capaz de rechazar la  $H_0$  al nivel de significación establecido.

Si el diseño es a dos colas existe un tercer error posible, llamado de tipo III o de tercera especie. Consiste en rechazar la hipótesis nula en una dirección de las

alternativas cuando lo cierto es la alternativa contraria; es decir, se acepta superioridad de un tratamiento cuando lo cierto es que el superior es el otro tratamiento. La probabilidad de este tipo de error es muy pequeña y, en general, suele ignorarse. Si el diseño es a una cola este error no puede contemplarse.

### *Diferentes métodos para el cálculo de tamaño muestral*

Cuando se trabaja con variables aleatorias binarias se puede calcular el tamaño de la muestra necesario para detectar una diferencia a partir de la prueba exacta de Fisher<sup>24, 25, 26</sup>. Esta prueba se basa en la distribución de probabilidad del número de éxitos en el tratamiento experimental -casilla (a)- condicional a que todos los totales marginales de las casillas de 2x2 sean fijos. Esta distribución condicional se debe calcular por métodos numéricos y supone un costo de cálculo cuando el tamaño de muestra es muy grande. Por este motivo se suelen utilizar transformaciones que modifiquen los datos de las tablas de manera que sigan otra distribución de probabilidad más sencilla como la normal.

Existen numerosos métodos alternativos para calcular el número de sujetos necesarios cuando se utilizan variables binarias<sup>27, 28, 29, 30, 31, 32, 33, 34</sup>. Una revisión de varios métodos y programa puede encontrarse en Dupont & Plummer 1990<sup>35</sup>. Para una revisión más detallada con fórmulas y tablas ver Machin & Campbell<sup>36</sup>.

### *Diseños de equivalencia terapéutica*

El objetivo habitual de un ensayo clínico suele ser la demostración de superioridad de un tratamiento experimental frente de un control. No obstante, el continuo desarrollo de nuevos fármacos ha requerido nuevas estrategias y diseños.

El concepto de ensayo clínico controlado con placebo se desarrolló en los años 60, cuando había la necesidad de demostrar la eficacia sin ambigüedades mediante la aceptación de la superioridad del nuevo tratamiento frente a placebo. En el análisis de estos ensayos clínicos se utilizaba la metodología estadística basada en la prueba de hipótesis clásica <sup>37</sup>.

Con la existencia cada vez más frecuente de fármacos de eficacia probada, la posibilidad de realizar estudios controlados frente a placebo es cada vez más difícil por problemas éticos <sup>37</sup>. De todas maneras, la metodología de los ensayos clínicos controlados frente a placebo puede ser íntegramente transferida a ensayos en los que se busca la superioridad frente a un tratamiento establecido.

Cuanto más eficaces se van mostrando los nuevos fármacos, más difícil se hace la demostración de superioridad. Así, haría falta cada vez más, definir diferencias relevantes a detectar de menor magnitud con lo que el tamaño muestral puede llegar a ser enormemente elevado y racionalmente absurdo. Dado que cada vez el número de fármacos comercializados va en aumento, y que el número de posibles indicaciones en las que el placebo puede ser estudiado disminuye, hay pocos estudios en los que sea posible justificar el uso del placebo.

Por otro lado, a partir de la metodología utilizada en los estudios de bioequivalencia, se han ido consolidando nuevas estrategias para la demostración de equivalencia terapéutica entre dos tratamientos activos. De esta manera se han desarrollado un tipo de estudios denominados por algunos autores como estudios de equivalencia con control activo (active control equivalence studies ACES) <sup>38</sup>.

### *Razones para justificar estudios de equivalencia terapéutica*

¿Por que puede existir la necesidad de llevar a cabo nuevos estudios con fármacos que no ofrecerán más eficacia que la aportada por los tratamientos ya establecidos? Los ensayos clínicos hacen una estimación del efecto global sobre un grupo de pacientes seleccionado. La respuesta individual puede variar, incluso en gran medida, entre un fármaco y otro, y por lo tanto unos sujetos determinados pueden beneficiarse más que otros. Además, la eficacia es una, pero no la única de las propiedades de un medicamento. Fármacos de segunda o tercera generación de un grupo determinado, a menudo han sido seleccionados por diferentes propiedades farmacocinéticas, menor potencial para interacciones, menor coste económico, mejor tolerabilidad, etc.

### *Metodología*

En primer lugar debe remarcarse la importancia que tiene el hecho de ser imposible de demostrar la equivalencia absoluta, solamente se puede probar que la diferencia real no va más allá de unos límites preestablecidos<sup>39</sup>.

En los últimos años se han publicado bastantes artículos de opinión o de revisión de ensayos clínicos de equivalencia terapéutica tanto en revistas específicas de estadística como en revistas de ámbito médico<sup>37,38,40-47</sup>. En éstos se revisa la metodología y hechos importantes a tener en cuenta para el diseño de estos estudios.

El número de diseños existentes es bastante numeroso por lo que dados los objetivos de la tesis nos limitaremos a revisar los más aceptados. Todos ellos tienen en común la definición a priori de un rango de tolerabilidad o de aceptación de equivalencia. Aunque en algunos casos pueda aplicarse la regla del 20% usada en bioequivalencia, la definición de ese rango debe individualizarse para cada caso concreto y debe ser lo suficientemente estrecho para discriminar entre la respuesta del fármaco de eficacia establecida (control) y la respuesta de la evolución espontánea del proceso estudiado, el cual puede estimarse basándose en la respuesta al placebo.

### Aproximación clásica por poder

El planteamiento clásico para abordar la equivalencia terapéutica se basa en la primera aproximación por potencia. Bajo este planteamiento, la condición de equivalencia se identifica con la  $H_0$  clásica de igualdad, mientras que la  $H_1$  representa la no equivalencia. Puesto que  $H_0$  puede no rechazarse debido a su escasa potencia, el planteamiento anterior sería completamente inadmisibles en ausencia de un cálculo muestral y una hipótesis clara de trabajo que garantice una potencia adecuada (como mínimo un 80%). Una extensión de este método consiste en realizar el cálculo de la potencia a posteriori y comprobar que ésta sea suficiente.

$$H_0: \theta_T - \theta_R = 0 \quad \Rightarrow \text{Equivalencia}$$

$$H_A: \theta_T - \theta_R \neq 0 \quad \Rightarrow \text{Inequivalencia}$$

Los problemas de este método son consecuencia de que el valor crítico de rechazo de la hipótesis nula se calcula en base a la diferencia de medias (o proporciones) estandarizada (dividida por su error estándar). De esta manera el rechazo de la hipótesis nula, es decir, conclusión de no equivalencia, puede producirse tanto por una gran diferencia de medias como por un error estándar pequeño. Así, pequeñas diferencias entre fármacos, sin relevancia alguna, podrían llevar a concluir no equivalencia por exceso de precisión y lo mismo puede suceder en sentido contrario. Grandes diferencias podrían satisfacer el criterio de equivalencia por falta de precisión, sobretodo si no se realiza el cálculo de poder a posteriori<sup>48</sup>.

Como consecuencia de estos planteamientos, el riesgo para los pacientes se sitúa a un nivel alto, habitualmente del 20%, y se reduce el asumido por el promotor como se describe en la siguiente tabla:

		Conclusión de la prueba	
		$H_0$	$H_A$
Realidad	$H_0$	OK $1-\alpha$	Riesgo para el promotor $\alpha (0.05)$
	$H_A$	Riesgo para el paciente $\beta (0.2)$	OK $1-\beta$

### Aproximación por intervalos de confianza y test doblemente unilateral de Schuirmann

Actualmente para la valoración de eficacia se defiende cada vez más la aproximación mediante intervalos de confianza, incluso en los diseños clásicos <sup>49-53</sup>. Esta aproximación ha ganado uso especialmente en estudios de bioequivalencia y se ha mostrado como una buena herramienta para la equivalencia terapéutica.

Básicamente consiste en que una vez definido a priori el rango de aceptación de equivalencia,  $(\Delta_1, \Delta_2)$ , el intervalo formado por los límites inferior y superior de la diferencia calculado a posteriori (LI, LS) esté completamente dentro del rango de equivalencia preestablecido.

Otra aproximación por IC descrita por Westlake en 1976 <sup>54</sup> emplea intervalos simétricos alrededor del 0, en lugar de centrados en la diferencia. Aunque en su momento este método fue bastante utilizado en bioequivalencia hoy en día se prefieren los IC clásicos.

Schuirmann en 1987 <sup>48</sup>, propone un método basado en dos pruebas de hipótesis unilaterales simultáneas:

$$H_{01}: \theta_T - \theta_R \leq \Delta_1$$

$$H_{A1}: \theta_T - \theta_R > \Delta_1$$

y

$$H_{02}: \theta_T - \theta_R \geq \Delta_2$$

$$H_{A2}: \theta_T - \theta_R < \Delta_2$$

lo que equivale a

$$H_0: \theta_T - \theta_R \leq \Delta_1 \quad \text{o} \quad \theta_T - \theta_R \geq \Delta_2 \quad \Rightarrow \text{Inequivalencia}$$

$$H_A: \Delta_1 < \theta_T - \theta_R < \Delta_2 \quad \Rightarrow \text{Equivalencia}$$

Si se rechazan las dos hipótesis nulas al nivel de significación  $\alpha$ , puede concluirse equivalencia con un valor de p global igual al mayor de las dos pruebas realizadas. El incremento del riesgo  $\alpha$  global no sucede aquí puesto que las dos hipótesis nulas no pueden ser simultáneamente ciertas. Este método es equivalente a la aproximación mediante intervalos de confianza clásicos al  $100*(1-2\alpha)\%$ .

### **Capturar y pegar figura de zona de decisión de equivalencia comparada\$.**

Hauk y Anderson en diversos trabajos conjuntos han estudiado otras aproximaciones paramétricas similares al test doblemente unilateral<sup>55-57</sup> (respectivamente en 1983, 1984 y 1986). En el trabajo de Schuirmann se critica estas aproximaciones {1291} puesto que la región de decisión de aceptación de equivalencia es independiente de que el error estándar sea muy grande (ver figura anexa).

gráficas:

Región de aceptación en función de se y dif. comparación de IC, schuirmann y Hauck & Anderson <sup>55</sup>

Para el cálculo del tamaño muestral, aunque las fórmulas originales son muy complejas <sup>58</sup>, se han publicado tablas y curvas de potencia para el modelo aditivo <sup>59</sup>), para el modelo multiplicativo <sup>60,61</sup>) y fórmulas aproximadas para el modelo aditivo <sup>62</sup>, y para el modelo multiplicativo <sup>63</sup>.

Ejemplo de 6 estudios distintos comparando las decisiones de equivalencia. En todos ellos la desviación típica es idéntica y vale 10. El intervalo de aceptación de equivalencia de va entre -10 y +10.

Ej.	N por grupo	Diferencias observadas	ES(dif)	Aproximación por potencia			IC – test doblemente unilateral			
				P	1-β	Equivalencia	P	IC 90%		Equivalencia
1	100	3	1.41	0.035	> 0.99	NO	<0.001	0.66	5.34	SI
2	20	3	3.16	0.348	0.87	SI	0.017	-2.33	8.33	SI
3	14	3	3.78	0.437	0.72	NO	0.038	-3.45	9.45	SI
4	100	6	1.41	<0.001	> 0.99	NO	0.003	3.66	8.34	SI
5	20	6	3.16	0.065	0.87	SI	0.108	0.67	11.33	NO
6	14	6	3.78	0.124	0.72	NO	0.149	-0.45	12.45	NO

A diferencia de la aproximación por potencia, los IC y los diseños doblemente unilaterales ofrecen una protección superior para el paciente a costa de aumentar el riesgo para el promotor; habitualmente los riesgos se sitúan en un 5% ( $\alpha$ ) y un 20 ( $\beta$ ) respectivamente.

		Conclusión de la prueba	
		H <sub>0</sub>	H <sub>A</sub>
Realidad	H <sub>0</sub>	OK 1- $\alpha$	Riesgo para el paciente $\alpha$ (0.05)
	H <sub>A</sub>	Riesgo para el promotor $\beta$ ( $\geq 0.2$ )	OK 1- $\beta$

**Selección de la amplitud de los IC y del nivel de significación**

Parece bastante claro que en estos estudios, como en otros, el riesgo que se debe controlar de forma primordial sea la protección para futuros pacientes. Aunque ha habido durante algún tiempo cierta discusión acerca del nivel de significación global y de la amplitud de los intervalos de confianza, si el riesgo  $\alpha$  global se establece en el 5%, los intervalos se deben construir al  $100*(1-2*\alpha) \%^{64}$ .

Efectivamente, dado que las dos hipótesis nulas del test de Shuirmann son sucesos incompatibles, así como también es cierto que la verdadera estimación del efecto no puede ser superior e inferior al intervalo de confianza calculado, si calculasemos los intervalos de confianza al 95%, estamos operando con un riesgo  $\alpha$  global del 2.5%.

Así, aunque la decisión pueda variar en cada caso, para niveles de protección globales de riesgo  $\alpha$  del 5% los intervalos se deben calcular al 90% y en caso de calcular los intervalos al 95% debemos tener en cuenta que el nivel de significación queda fijado en el 2.5%.

### **Otras aproximaciones**

Modelos semejantes han sido descritos para intervalos de confianza no paramétricos <sup>65,66</sup>, para pruebas de hipótesis con variables binarias basándose en el intercambio de las dos hipótesis <sup>67,68</sup> con sus correspondientes fórmulas para predeterminación del cálculo muestral <sup>69-71</sup>, y para datos binarios apareados <sup>72</sup>). Kieser <sup>73</sup> propone un método intermedio en el que se puede hacer un test de equivalencia, y si ésta se alcanza, se puede probar la superioridad sin gasto de alfa. También se han descrito varias alternativas mediante un enfoque bayesiano aunque su uso en equivalencia terapéutica no está hoy por hoy extendido <sup>74-78</sup>.

Finalmente, proveniente del ámbito de la bioequivalencia, se está trabajando en el sentido de dar importancia al a equivalencia no global, sino ponderando el efecto en cada sujeto, como por ejemplo con test de razones individuales de equivalencia <sup>79</sup>. Estas tendencias requieren datos apareados y de momento no se están aplicando a ensayos de eficacia.

### *Metodología secuencial*

#### *Justificación de los métodos secuenciales*

En algunos campos de investigación, tales como en agricultura o experimentación animal, se dispone la información toda al mismo tiempo. En cambio, en otros casos como ensayos clínicos o estudios epidemiológicos, la información se va recuperando paulatinamente en el tiempo.

Parece interesante la posibilidad de realizar análisis intermedios para aprovechar este hecho con el fin de inspeccionar los datos antes de haber concluido todos los pacientes y detener el estudio si se llega a una conclusión clara.

Ante este hecho podemos tomar las siguientes actitudes:

- (1) Ignorarlo con lo que no se detiene el estudio hasta el final programado; es lo que sucede en un diseño clásico.
- (2) Monitorizar efectos adversos: Decisión de la continuación del estudio independiente de la eficacia
- (3) Aprovechar la información a medida que se acumula: diseño secuencial.

#### *Monitorización de la información*

Es recomendable que los ensayos clínicos, sobretodo los de mayor duración, exista un Comité de Monitorización de Datos para poder resolver problemas detectados durante su desarrollo. En algunos casos pueden efectuar funciones de control de calidad de los datos, pero sus objetivos deben ser de más alto nivel e intervenir en las decisiones que afecten en la continuidad de un estudio <sup>80</sup>.

Se recomienda que esté formado por pocas personas, 3 ó 4, y que esté integrado por al menos un clínico especialista y un estadístico. Ninguno de los integrantes debe estar relacionado con el reclutamiento ni constar como investigador <sup>81</sup>.

### **Análisis intermedios**

Estos análisis, sobre todo en ensayos clínicos, desempeñan una función importante. En primer lugar permiten evaluar si se cumple el protocolo respecto a los factores como la tasa de reclutamiento, el adecuado cumplimiento de los criterios de inclusión y exclusión o el balance entre los grupos respecto a los factores pronósticos. En segundo lugar permiten un estudio de los efectos adversos de los tratamientos de manera que se pueda interrumpir el ensayo clínico si alguno de los mismos es claramente desfavorable debido a la producción de efectos no deseados. En tercer lugar, los análisis intermedios son un buen sistema para que los investigadores mantengan el interés por el estudio, especialmente en estudios multicéntricos o de larga duración.

### **Monitorización ingenua**

La monitorización ingenua es el hecho de inspeccionar los datos de un estudio con diseño fijo. En un análisis intermedio no se deben realizar análisis para comparar la eficacia de los dos tratamientos, a no ser que ello se haya previsto en el protocolo y se haya diseñado adecuadamente mediante una técnica secuencial. Si se realizan estos análisis y se ignoran los resultados no ocurre nada, pero es difícil mantener un estudio hasta el final si ya se sabe que existen grandes diferencias a favor de uno de los tratamientos. Si se interrumpe un estudio porque uno de los tratamientos es mejor, pero el protocolo no preveía este evento, las conclusiones estarán sesgadas. En primer lugar, el nivel de significación del estudio calculado mediante una prueba de hipótesis clásica será inválido <sup>82</sup>. El valor de  $\alpha$  que se obtenga será menor que el real, es decir, la probabilidad de equivocarse al decir que existen diferencias será mayor a la calculada mediante la prueba clásica. En segundo lugar, la estimación de

la magnitud del efecto comparado entre tratamientos y su intervalo de confianza también serán inválidos. La diferencia estará sobrevalorada y el intervalo de confianza será más estrecho de lo debido, por estos motivos el análisis frecuentista habitual es inválido.

Un problema añadido de los aspectos estadísticos, es que la evaluación de futuras inclusiones puede estar influido por los resultados previos tanto en el mismo estudio como en estudios similares concurrentes. Este hecho puede ser de vital importancia en estudios no enmascarados y puede suponer un fuerte sesgo <sup>83</sup>.

En resumen, el análisis estadístico de un estudio antes de finalizar el reclutamiento total de sujetos da lugar a conclusiones erróneas si no se emplea una metodología adecuada. El error de tipo I se incrementa de manera importante al aumentar el número de inspecciones. De hecho aún siendo cierta la hipótesis de igualdad, si se realizaran infinitas inspecciones, se podría, llegar a, rechazar  $H_0$  con toda seguridad<sup>82</sup>. En la tabla siguiente se expresa la probabilidad de error para rechazar la hipótesis de igualdad siendo esta cierta, al aumentar el número de inspecciones:

Nº de inspecciones	P(rechazar H0   H0)
1	0.05
2	0.08
5	0.14
10	0.19
50	0.32
100	0.37
1000	0.53
Infinito	1

### **Terminación prematura de ensayos clínicos**

El criterio estadístico no es el único que puede hacer terminar un estudio antes de lo previsto. A continuación se presenta un lista sintetizando motivos de diversa índole:

Basados en los datos acumulados del ensayo:

El tratamiento experimental es claramente peor que el control

Hay pocas posibilidades de demostrar que el tratamiento experimental sea mejor  
(no hay evidencia ni probabilidad razonable de demostrar beneficio)

El tratamiento experimental ya ha demostrado que es claramente mejor

Efectos secundarios inesperados e inaceptables

Basado en el desarrollo global del estudio:

Imposibilidad de reclutar suficientes pacientes a ritmo necesario

Falta de cumplimiento en un elevado número de pacientes

Insuficiente financiación

Basado en información externa al estudio:

Datos de otros estudios que establecen inequívocamente beneficio o perjuicio del tratamiento

Datos de la práctica clínica habitual que sugieren efectos secundarios no sospechados e inaceptables

Desarrollos que cambian los estándares del tratamiento

Retirada del tratamiento del mercado

Por el otro lado, puede suceder que la ventaja del tratamiento experimental sea moderada, o bien, la varianza mayor de la esperada, con lo que el poder no se mantiene. En ambos casos parece razonable la prolongación del estudio.

Una revisión del tema con algunos ejemplos reales de estudios que acabaron prematuramente de forma correcta e incorrecta puede encontrarse en Eur. Heart J. 1994; 15: 721-738 <sup>81</sup>.

### *Ventajas y limitaciones de los diseños secuenciales*

Las principales ventajas que se obtiene al escoger estos métodos son:

Ética: La asignación aleatoria de pacientes a grupos de tratamientos es ética siempre que existan dudas acerca de qué tratamiento es mejor. En el momento en que se disipan estas dudas deja de ser ético seguir randomizando y exponiendo sujetos a un tratamiento inferior o menos seguro.

Economía: Al permitir concluir antes en general, se requiere menos recursos de tiempo, dinero y esfuerzo.

Poder estadístico: Algunos modelos secuenciales permiten asegurar el poder estadístico predefinido en el diseño independientemente de la variabilidad.

Incentivación: Asimismo, el hecho de conocer resultados puede incentivar a los investigadores (y a los promotores) en la tarea de reclutamiento y seguimiento del estudio.

Por otra parte, los diseños secuenciales tienen las siguientes limitaciones:

Dependencia del modelo: un modelo y el parámetro theta ( $\theta$ ) deben ser elegido a priori.

Aproximación en el análisis: el análisis permite las inspecciones intermedias, pero utiliza una forma más aproximada de distribución para el test estadístico que el análisis convencional. No obstante, a efectos prácticos esto sólo puede ser un problema en ensayos de tamaño muestral pequeño.

Están diseñados para una sola respuesta y dos grupos de comparación: diseños con más de una respuesta o más de dos tratamientos precisan métodos más complicado, basados en aproximaciones.

### *Aplicabilidad del diseño secuencial*

Dadas las características de los métodos secuenciales, pueden considerarse de utilidad cuando se presentan las siguientes características:

Gran tamaño de muestra necesario

El objetivo del estudio es único y está bien definido

Se prevé una larga duración del estudio

Se conoce el resultado de un paciente pronto, mucho antes de que finalice la fase de inclusión

Por otro lado se debe descartar su uso cuando:

- El número de pacientes necesario es reducido

- El estudio tiene objetivos múltiples

- El resultado de los pacientes se conoce tarde, una vez finalizado el período de inclusión

- La logística del estudio no permite realizar análisis intermedios

Factores que no limitan el empleo de los métodos secuenciales

- Diseño paralelo o cruzado

- Tipo de variable de medida: binaria, cuantitativa, tiempo de supervivencia, ordinal

- Necesidad de ajustar el efecto de covariables o factores mediante estratificación o modelos multivariantes

- Razones de asignación de los pacientes a los grupos distintas de 1:1

- Más de 2 grupos de tratamientos

### *Tipos de diseño secuencial*

Los métodos secuenciales pueden clasificarse en los siguientes grupos:

Tests de significación repetidos o métodos secuenciales por grupos.

Aproximación por funciones de gasto de alfa.

Intervalos de confianza repetidos.

Restricción estocástica.

Métodos Bayesianos

Límites continuos (métodos basados en la función de verosimilitud)

Reestimación del tamaño muestral.

### *Historia de los métodos secuenciales*

El nacimiento de los métodos secuenciales se remonta a mediados de nuestro siglo. En 1947 A. Wald <sup>84</sup> publicó el primer libro sobre el tema. El poco éxito de tan interesante idea probablemente se debió a la complejidad en el cálculo de las pruebas estadísticas, a la rigidez de los diseños que inicialmente se investigaron y a la falta de una motivación clara para seguir investigando sobre el tema; las primeras aplicaciones se realizaron en el campo del control de calidad industrial.

A partir de la década de los 60, con la aparición de los ordenadores y el desarrollo de la industria farmacéutica, actividad que precisa ensayar numerosos fármacos en búsqueda de alternativas eficaces, se reactivó en gran medida la investigación sobre

las técnicas de análisis secuencial tras la publicación por parte de P. Armitage del libro "Sequential Medical Trials"<sup>85</sup>.

El mismo autor reeditó su obra en 1975 con una puesta al día de las investigaciones realizadas hasta entonces. Este libro fue el punto de partida de los métodos actuales. En él se describe con detalle la teoría de las pruebas de significación repetidas, base de los modernos métodos de análisis secuencial por grupos.

De todos modos, los métodos descritos por P. Armitage eran poco flexibles; aunque podían utilizarse diferentes tipos de variables, los métodos se basaban en diseños apareados y era preciso la monitorización continua de los datos. Estas características limitaban mucho la posibilidad de aplicar las técnicas secuenciales a ensayos clínicos.

### *Uso y reconocimiento de los métodos secuenciales*

En la actualidad existen numerosos grupos de estadísticos teóricos trabajando en el tema; se están buscando métodos más flexibles y poco a poco se van generando las técnicas estadísticas habituales en el futuro.

Desde los primeros trabajos publicados hasta los años 80 el uso ha sido ocasional: Kilpatrick & Oldham 1954<sup>86</sup> y Freiereich et al. 1966<sup>87</sup>. A partir de entonces el uso ha sido creciente, especialmente en ensayos clínicos en oncología, patología cardiovascular, y en SIDA, sobretodo con los métodos de O'Brien & Fleming comúnmente con diseños de límites continuos y funciones de gasto de alfa. El número de ensayos clínicos publicados hasta hoy en día que usan métodos secuenciales es muy grande y tenemos numerosos ejemplos<sup>88 89 90 91 92 93 94 95 96 97 87 98 99 100 91 86 101 102 103 104 105 106 90 107 108 109 110 111</sup>

## INTRODUCCIÓN

---

Aunque por sus características y limitaciones su ámbito de aplicación se fundamenta sobretodo para estudios de eficacia que se encuadran en fase III, puede utilizarse para otros usos, como por ejemplo bioequivalencia, monitorización paralela de seguridad, y fases más tempranas del ensayo clínico<sup>112</sup>. Se han desarrollado aproximaciones para ensayos clínicos no controlados de fase II mediante una aproximación del modelo triangular de Whitehead con software incluido <sup>113</sup>.

La entidad reguladora norteamericana Food and Drug Administration (FDA) rechaza habitualmente los ensayos clínicos con inspecciones no previstas sino están muy bien justificadas. Los métodos que con más frecuencia evalúan son estudios que utilizan el método de O'Brien y Fleming y ha aprobado protocolos basados en métodos de límites continuos con el programa PEST <sup>114</sup>. También se han aprobado estudios por parte de las administraciones europeas, incluidos estudios que han usado métodos secuenciales de límites continuos <sup>98</sup>.

*Bases del diseño secuencial*

**Parametrización y medida del efecto**

En los modelos secuenciales, la medida de las diferencias entre tratamientos se denomina theta ( $\theta$ ). Para diseños con variables binarias,  $\theta$  equivale al logaritmo neperiano del odds ratio. En base a los valores observados se calculan dos estadígrafos: Z y V. Z es un estimador de las diferencias acumuladas entre los tratamientos y V mide la cantidad de información. En el apartado “Desarrollo estadístico \*” se da una explicación detallada y explican las bases de este apartado.

Theta =  $\theta$  ( $\theta_R$  = diferencias relevante especificada en el diseño)

E = grupo experimental

C = grupo control

*Para datos binarios*

E: éxito       $Pe = Se/Ne$

C: éxito       $Pc = Sc/Nc$

$$\theta = \log ( Pe*(1-Pc) / Pc*(1-Pe) )$$

$$\theta = \log (Pe/(1-Pe)) - \log (Pc/(1-Pc))$$

	<i>E</i>	<i>C</i>	<i>Total</i>
<i>Éxito</i>	Se	Sc	S
<i>Fracaso</i>	Fe	Fc	F
	Ne	Nc	N

## INTRODUCCIÓN

---

$$Z = (N_c * S_e - N_e * S_c) / N$$

$$V = N_e * N_c * S * F / N^3$$

$$Z^2 / V = \chi_1^2 = \sum (O-E)^2 / E$$

$$\Rightarrow Z/V \sim N(0,1)$$

Para  $N_e$  y  $N_c$  grandes

$$V \approx N_e * N_c / N * p * (1-p)$$

donde  $p$  es el promedio de éxitos en el ensayo ( $S/N$ ).

Si  $N_e = N_c$ ,

$$V \approx \frac{1}{4} N * p * (1-p)$$

### *Para otros datos*

También están descritas las fórmulas de  $Z$ ,  $V$  y  $\theta$  para datos normales, ordinales, supervivencia y con distribución de poisson <sup>112</sup>.

Respuesta	$\theta$	Test equivalente	Asunción
Binaria	log-odds-ratio	Pearson's $\chi^2$	—
Poisson	Log-event-rate-ratio	$\chi^2$	—
Supervivencia	Log-hazards-ratio	logrank test	Riesgos proporcionales
Ordinal	log-odds-ratio	Mann-Whitney	odds proporcionales

## INTRODUCCIÓN

---

---

Normal	$(\mu_E - \mu_E) / \sigma$	$\sim t\text{-test}$	igual $\sigma^2$
--------	----------------------------	----------------------	------------------

---

### Diseño estadístico clásico basado en Z y V

Para detectar una diferencia de referencia  $\theta_R$  se puede utilizar Z como prueba estadística. Si Z es mayor a un cierto valor de k se rechaza la hipótesis nula con un riesgo de primera especie alfa y un poder estadístico 1-beta. Los requerimientos de la prueba son:

$$P(Z \geq k; 0) = \alpha$$

$$P(Z \geq k; \theta_R) = 1 - \beta$$

donde los valores entre paréntesis que siguen al punto y coma representan el verdadero valor de  $\theta$ . Como  $Z \sim N(\theta V, V)$ , si

$$\theta = 0 \quad Z \sim N(0, V)$$

y si

$$\theta = \theta_R \quad Z \sim N(\theta_R V, V)$$

de donde se deduce que

$$\Phi[k / \sqrt{V}] = 1 - \alpha \quad \Rightarrow \quad k / \sqrt{V} = U_\alpha$$

y

$$\Phi[(\theta_R V - k) / \sqrt{V}] = 1 - \beta \quad \Rightarrow \quad (\theta_R V - k) / \sqrt{V} = U_\beta$$

En estas expresiones  $\Phi()$  representa la función de probabilidad de la distribución normal y  $u_\alpha$  o  $u_\beta$  son los valores de la distribución normal que dejan un área (una probabilidad)  $\alpha$  o  $\beta$  respectivamente por encima. A partir de este sistema de ecuaciones puede obtenerse fácilmente la expresión

$$(u_\alpha + u_\beta)^2 = \theta_R^2 V$$

## INTRODUCCIÓN

---

El valor de  $V$ , la información que  $Z$  contiene sobre  $\theta$ , depende del tamaño de muestra y puede utilizarse la expresión anterior para calcular el número de sujetos necesarios.

$$V = [(u_\alpha + u_\beta) / \theta_R]^2$$

Para cada forma de  $\theta$  que depende del tipo de variable a estudiar y la parametrización del efecto derivan unos estadísticos  $Z$  y  $V$  específicos.

### *Ejemplo para datos binarios*

Se describe un ejemplo de un estudio hipotético analizado de forma convencional y mediante aproximación secuencial con valores de  $Z$  y  $V$ . Dado que únicamente se ha realizado una única inspección, los resultados son superponibles:

Valores observados en el estudio (O)

	Experimental	Control	
Éxito	272	269	541
Fracaso	101	126	227
	373	395	768

### *Análisis convencional*

Valores esperados (E)

	Experimental	Control
Éxito	262,7513	278,2487
Fracaso	110,2487	116,7513

(O-E)

## INTRODUCCIÓN

---

	Experimental	Control
Éxito	9,2487	-9,2487
Fracaso	-9,2487	9,2487

$$(O-E)^2 / E$$

	Experimental	Control
Éxito	0,3255	0,3074
Fracaso	0,7759	0,7327
		2,1415

$$\begin{aligned} \chi^2 &= \Sigma(O-E)^2 / E &&= 2,1415 \\ OR &= (272*126)/(101*269) = 1,261438 \\ \theta &= \ln(OR) &&= 0,232252 \\ EE_{OR} &= (1/a + 1/b + 1/c + 1/d)^{1/2} = 1,172155 \\ EE_{\theta} &= \ln(EE_{OR}) &&= 0,158844 \end{aligned}$$

### *Análisis mediante aproximación secuencial*

$$\begin{aligned} Z &= 9,248698 \\ V &= 39,943433 \\ \theta &= Z/V &&= 0,231545 \\ OR &= \exp(\theta) &&= 1,261438 \\ EE_{(OR)} &= \sqrt{V} &&= 0,158226 \\ Z^2/V &(\sim \chi_1^2) &&= 2,1415 \end{aligned}$$

### **Desarrollo estadístico\***

Si la medida del efecto de los tratamientos sólo dependiera de un parámetro, la derivación de Z y V sería sencilla. Si se dispone de N observaciones  $x_1, x_2, \dots, x_n$ , de

## INTRODUCCIÓN

---

una variable aleatoria  $X$  que tiene función densidad de distribución  $f(x;\theta)$ , el logaritmo de la función de verosimilitud de  $\theta$  es

$$\ell(\theta; x_1, x_2 \dots x_t) = \sum_{i=1}^t \ln f(x_i; \theta)$$

Si  $\theta$  es pequeño, una serie de Taylor puede aproximar la función  $\ell(\theta)$

$$\ell(\theta) = \ell(0) + \theta \ell_{\theta}(0) + \frac{1}{2} \theta^2 \ell_{\theta\theta}(0) + o^3$$

$$\ell(\theta) = \text{const} + \theta Z - \frac{1}{2} \theta^2 V + o^3$$

$$Z = \ell_{\theta}(0) \quad \text{y} \quad V = - \ell_{\theta\theta}(0)$$

donde  $\ell_{\theta}(0)$  es el valor de la primera derivada del logaritmo de la función de verosimilitud suponiendo diferencias nulas,  $\ell_{\theta\theta}(0)$  es el mismo valor de la segunda derivada y  $o^3$  simboliza el término de error de la aproximación de la serie.

El estadístico  $Z$  cuantifica la evidencia de las diferencias acumuladas entre los grupos de comparación a lo largo del estudio y  $V$  mide la información que  $Z$  posee sobre  $\theta$ . En términos estadísticos,  $Z$  es el estimador eficiente (efficient estimator o efficient score) y  $V$  es la información de Fisher <sup>115</sup>.  $V$  aumenta en función de  $N$  (el número de pacientes) y  $Z$  varía en función de las diferencias. Si los tratamientos son similares  $Z$  tendrá un valor cercano a 0, si son diferentes, el valor se alejará de 0 de la siguiente manera: en caso de que el tratamiento experimental es mejor con signo positivo y en caso contrario con signo negativo. Cuando la estimación de las diferencias ( $\theta$ ) es pequeña y el número de pacientes ( $N$ ) es grande,  $Z$  tiene una distribución aproximadamente normal con media  $\theta V$  y varianza  $V$ .

$$Z \sim N(\theta V, V)$$

Como  $Z$  es una suma,

$$\sum_{i=1}^n \left[ \frac{d}{d\theta} \ln \{f(x_i; \theta)\} \right],$$

## INTRODUCCIÓN

---

es asintóticamente normal. De hecho para  $n$  grande y  $\theta$  pequeña,  $Z \sim N(\theta V, V)$

Cuando la función de verosimilitud depende de otros parámetros ( $\varnothing$ , que puede ser un vector), además del que nos interesa, la derivación de  $Z$  y  $V$  se complica un poco, pues los parámetros extra deben estimarse también a partir de la muestra de observaciones.

Si se sustituyen éstos parámetros por sus estimadores de máxima verosimilitud para un valor de  $\theta$  ( $\hat{\varnothing}(\theta)$ ), entonces el logaritmo de la función de verosimilitud del parámetro de interés depende sólo de  $\theta$  y puede utilizarse la serie de Taylor para su aproximación. En este caso,

$$Z = l_{\theta}(0, \varnothing_0)$$

$$V = -l^{\theta\theta}(0)$$

donde  $l^{\theta\theta} = l_{\theta\theta} - l_{\varnothing\theta} l^{-1}_{\varnothing\theta} l_{\theta\varnothing}$  y  $\varnothing_0 = \hat{\varnothing}(0)$

$l_{\theta}$  es la primera derivada del logaritmo de la función de verosimilitud respecto a  $\theta$  y  $l_{\theta\theta}$ ,  $l_{\varnothing\theta}$ ,  $l_{\varnothing\varnothing}$ ,  $l_{\theta\varnothing}$  son las segundas derivadas respecto a los parámetros  $\theta\theta$ ,  $\varnothing\theta$ ,  $\varnothing\varnothing$  y  $\theta\varnothing$  respectivamente.

### **Característica operante**

La característica operante (CO) es la probabilidad de aceptar la  $H_0$  en función del valor de una prueba de hipótesis basada en un parámetro  $\theta$  que cuantifica las diferencias entre grupos.

La comparación de curvas CO obtenidas mediante simulación con las teóricas permiten comprobar si la prueba estadística cumple las especificaciones prefijadas. Existen dos puntos de la curva que se han fijado en el diseño; cuando  $\theta$  vale 0, no hay diferencias, la probabilidad de aceptar  $H_0$  debe ser  $1 - \alpha$ . Cuando las diferencias tienen la magnitud de las especificadas en el diseño,  $\theta = \theta_R$ , la probabilidad de aceptar  $H_0$  debe ser  $\beta$ .

$$\theta = 0 \Rightarrow \text{CO} = 1 - \alpha$$

$$\theta = \theta_R \Rightarrow \text{CO} = \beta$$

### **Desarrollo estadístico\***

La función de la curva CO teórica para una prueba clásica se puede deducir a partir de una aproximación normal de la distribución de probabilidad del estadígrafo de la prueba  $Z$ . Para una hipótesis cualquiera

$$P(Z \geq k; \theta) = 1 - \text{CO}$$

y en concreto

$$P(Z \geq k; 0) = \alpha$$

donde los valores en el paréntesis detrás del punto y coma representan el verdadero valor de  $\theta$ . Como  $Z \sim N(\theta V, V)$ ,

$$\Phi \left[ \frac{k}{\sqrt{V}} \right] = 1 - \alpha \Rightarrow \frac{k}{\sqrt{V}} = U \alpha$$

y

$$\phi [ (\theta V - k) / \sqrt{V} ] = 1 - CO$$

A partir de estas ecuaciones puede obtenerse fácilmente la expresión

$$CO = 1 - \phi ( \theta \sqrt{V} - u_{\alpha} )$$

Para la medida de  $\theta$  de variable binomiales,  $\theta = \ln(OR)$ ;

$$V = \frac{1}{4} N p (1 - p)$$

donde

$$p = (p_0 + p_1)/2$$

$$p_1 = p_0 * e^{\theta} / \{1 + p_0 * (e^{\theta} - 1)\}$$

$$\Rightarrow p = [ p_0 * (1 + e^{\theta} / (1 + p_0 * (e^{\theta} - 1))) ] / 2$$

Se puede constatar que el valor de CO depende del tamaño de la muestra (N) calculado para unas especificaciones ( $\alpha$ ,  $\beta$  y  $\theta_R$ ), de la probabilidad de éxito en el grupo control ( $p_0$ ), y de la diferencia real ( $\theta$ ).

### *Revisión de los diferentes métodos secuenciales*

#### **Reestimación del tamaño muestral**

Es un compromiso entre el diseño clásico y el diseño secuencial. Aunque algunos métodos necesitan romper el ciego de la asignación <sup>116</sup>, existen alternativas tanto para datos binarios <sup>117</sup>) como para datos normales <sup>118</sup> que mantienen el enmascaramiento. Una revisión de estos métodos con ejemplos puede encontrarse en el libro *Biopharmaceutical sequential statistical applications*<sup>119</sup>.

Para una aproximación pragmática se diseña el estudio mediante la técnica habitual frecuentista. Ya que se necesita la tasa de partida o desviación estándar, se calcula un tamaño muestral aproximado. Se recluta aproximadamente el 50% del tamaño prefijado y se calculan los parámetros necesarios sin romper el ciego. Seguidamente se recalcula el tamaño muestral con los datos obtenidos en la primera inspección.

7 3  
Biblioteca General  
Edifici A  
08193 Bellaterra (Barcelona) Espanya

### **Métodos secuenciales por grupos**

Esta estrategia, también llamada pruebas de significación repetidas, fue introducida por Armitage <sup>120</sup> y revisada por Pocock <sup>121</sup>.

Consiste en definir *a priori* un número máximo determinado,  $K$ , de análisis intermedios. Éstos se realizarán a medida que se cuente con la información referente a un grupo de pacientes de tamaño preestablecido, según las especificaciones del ensayo.

Se emplea en cada análisis intermedio un nivel de significación  $\alpha'$  menor que el habitual, de 0,05. Si se encuentran diferencias significativas a ese nivel, el estudio finaliza. En caso contrario, se incluye otro grupo de pacientes de igual tamaño y se repite la prueba estadística con todos los pacientes acumulados.

Si al finalizar el análisis con  $K$  grupos no se encuentran diferencias, el estudio finaliza con resultado negativo y se acepta la hipótesis negativa con riesgo  $\beta$ . El valor de  $\alpha'$  se calcula de manera que, en función del número de análisis intermedios a realizar ( $K$ ) se mantenga, en conjunto, el nivel de significación inicial prefijado  $\alpha$ , habitualmente el 0.05. Si  $\alpha'$  es menor que  $\alpha$ , el número total de pacientes necesario para finalizar el ensayo puede ser mayor que el de un estudio con tamaño fijo si la hipótesis nula es cierta.

Existe la posibilidad de emplear valores de  $\alpha'$  variables en cada análisis intermedio, de manera que en los primeros, cuando hay pocos pacientes, sea difícil rechazar la hipótesis nula, mientras que en análisis sucesivos ello sea más fácil. Este diseño propuesto por O'Brien y Fleming (O'Brien PC & Fleming TR, 1979) <sup>122</sup>, hace difícil la finalización prematura del estudio a no ser que las diferencias sean enormes, lo cual pone en entredicho el interés de este diseño secuencial. Por otro lado, el método de Pocock facilita mucho la finalización en los primeros análisis intermedios y el pequeño tamaño de muestra empleado, aunque permite concluir que los tratamientos

## INTRODUCCIÓN

---

son diferentes, proporciona estimaciones de las diferencias con unos intervalos de confianza excesivamente amplios.

Pocock en 1982 <sup>123</sup>, sugiere un diseño intermedio entre el suyo inicial y el de O'Brien & Fleming. Algunos autores, como Machin <sup>124</sup>, están de acuerdo con Peto <sup>125</sup> por su aproximación pragmática que refleja lo que sucede en la práctica habitual. Ésta consiste en fijar el nivel de significación para todas las inspecciones al nivel de 0.001 excepto la última.

A continuación se presentan en una tabla para algunos de los métodos por grupos, los valores de  $\alpha'$  para cada inspección, teniendo en cuenta que el número máximo de éstas es 5 y que  $\alpha$ , el nivel de significación global, se establece en 0.05.

Inspección	Pocock 1977	O'Brien & Fleming 1979	Peto 1976	Pocock 1982
1	0.0158	0.00001	0.001	0.010
2	0.0158	0.0013	0.001	0.017
3	0.0158	0.0084	0.001	0.017
4	0.0158	0.0225	0.001	0.017
5	0.0158	0.041	0.049	0.021

Otras variantes menos utilizados pueden encontrarse en DeMets & Ware <sup>126</sup> y Emerson & Fleming <sup>127</sup>.

Aunque los métodos secuenciales por grupos son sencillos de emplear, imponen una rigidez en su diseño más estricta que el diseño clásico con el tamaño fijo, pues los análisis deben realizarse exactamente al obtener los resultados de un grupo de pacientes de tamaño predeterminado. Algunos autores han estudiado este problema <sup>128</sup> y en general no se modifican los resultados por pequeñas desviaciones del plan,

como realizar análisis en función del calendario en lugar de hacerlos según el número de pacientes, siempre que la tasa de inclusión sea más o menos constante.

Inicialmente el diseño de estos métodos se realizó para variables con distribución normal. Pocock adaptó los métodos RPST de Armitage<sup>120 82</sup> y demostró mediante simulaciones que los resultados podían aplicarse a otros tipos de variables (p.e. binomial y exponencial). O'Brien & Fleming 1979 también aplicaron el método para variables dicotómicas.

Actualmente estos métodos pueden utilizarse mediante tablas y fórmulas publicados por los autores antes citados, con algún programa diseñados para tal fin<sup>129</sup> o mediante la aproximación de los métodos de gasto de alfa<sup>130</sup>, que se revisaran más adelante.

Como principales problemas, estos métodos presentan gran dificultad para calcular el p-valor real, y la estimación del efecto al terminar el estudio.

Revisiones de los distintos métodos secuenciales por grupos pueden encontrarse en Lewis<sup>131</sup> y Pocock & Geller<sup>132</sup>.

### ***Desarrollo estadístico\****

Para el análisis convencional se procede al cálculo de Z y V para  $\theta$

Se rechaza  $H_0: \theta=0$  si  $|Z| \geq j\sqrt{V}$

( si  $j = 1.96 \Rightarrow \alpha=0.05$  ya que  $Z \sim N(\theta V, V)$ )

Si este paso se realiza más de una vez  $P(\text{rechazar } H_0 \mid H_0)$  se incrementa con el número de inspecciones (ver monitorización ingenua)

Para  $i=1, \dots, n$  inspecciones

se rechaza  $H_0: \theta=0$  si  $|Z_i| \geq k_i\sqrt{V_i}$

Se definen:

- El nivel de significación global

$$\alpha = P(|Z_i| \geq k_i \sqrt{V_i}, \text{ para alguna } i=1, \dots, n | H_0)$$

- El nivel de significación nominal para la  $i$  inspección

$$\eta_i = P(|Z_i| \geq k_i \sqrt{V_i} | H_0)$$

### Funciones de gasto de alfa

Puesto que el principal problema en los análisis múltiples es que se incrementa el nivel de significación  $\alpha$  por encima de lo deseado, se han propuesto métodos para compensar el “gasto” de  $\alpha$  a medida que se realizan análisis, de manera que se mantenga el valor global deseado <sup>133</sup>.

Los métodos secuenciales por grupos requieren (1) que se especifique a priori el número de inspecciones, y (2) los análisis se realizarán con grupos de información idénticos. Si la segunda condición no se cumple, estos métodos son inaplicables <sup>134</sup>. Los métodos descritos por Lan & DeMets en 1983 <sup>133</sup> y Kim & DeMets <sup>135</sup> relajan estos requisitos.

En el fondo es un procedimiento frecuentista flexible, en el sentido de que permite realizar los análisis que se desee y cuando se desee. En la práctica, consisten en definir una función matemática que describe cómo se modifica el nivel de significación que debe emplearse en los análisis en función de la cantidad de información acumulada y del número de análisis previos. Las inspecciones se pueden realizar tanto en función del calendario como de la información recibida ((Lan & DeMets <sup>136</sup>)

Al igual que en los diseños secuenciales por grupos, si se llega a un máximo de tamaño de muestra preestablecido y no se ha rechazado la hipótesis nula, el estudio finaliza con resultado negativo; de hecho los métodos secuenciales por grupos pueden considerarse un caso concreto de estos métodos. Existen múltiples propuestas de funciones de gasto posibles, pero en el momento de diseñar el estudio TIM (1992), no se habían desarrollado programas de ordenador que facilitaran los cálculos, por lo que éstos eran de difícil aplicación.

### *Base estadística\**

Bajo la asunción de monitorización continua, se define una función de gasto de  $\alpha$

$$\alpha^*(t), t \in (0,1)$$

por

$$\alpha^*(t) = P(V \leq t \cdot V_{\max} \mid \theta = 0) = \text{nivel de significación si } V \leq t \cdot V_{\max}$$

con

$$\alpha^*(0) = 0, \alpha^*(1) = \alpha$$

Lan & DeMets sugirieron el uso de las siguientes funciones para aproximar los métodos secuenciales por grupos:

Función equivalente al diseño de O'Brien & Fleming

$$\alpha^*(t) = 2 * \{ 1 - \phi(1.96 / \sqrt{t}) \}$$

Función equivalente al diseño de Pocock

$$\alpha^*(t) = 0.05 * \log \{ 1 + (e-1) * t \}$$

Kim & DeMets 1987<sup>135</sup>, Jennison & Turnbull 1989<sup>137</sup>

$$\alpha^*(t) = 0.05 * t^\rho, (\rho > 0)$$

si  $\rho > 1$  da valores bajos de la función para valores de  $t$  pequeños  $\Rightarrow$  límites anchos con baja probabilidad de finalización temprana

si  $\rho > 1$  da valores altos de la función para valores de  $t$  pequeños  $\Rightarrow$  límites anchos con mayores probabilidades de finalización temprana

Cualquier función creciente de 0 a  $\alpha$  con  $t$  creciente entre 0 y 1 puede ser usada como función de gasto de alfa.

Hasta el momento, las funciones han aportado soluciones en caso de límites de finalización para monitorización continua. Para monitorización discreta, Slud y Wei, 1982<sup>138</sup>; Lan y DeMets, 1983<sup>133</sup>) obtienen valores discretos para los límites. Antes de iniciar el ensayo se fijan  $\alpha$ ,  $1-\beta$ ,  $\theta_R$  y  $L$ . A cada inspección  $i$ , se calculan  $Z_i$  y  $V_i$ . Se escogen valores simétricos de los límites  $u_i$  y  $l_i$  para que la probabilidad de terminar previamente o en la inspección actual, bajo  $\theta=0$ , sea  $\alpha(V_i/L)$ .

$$P(Z_j \notin (l_j, u_j) \text{ para algún } j=1, \dots, i; 0) = \alpha(V_i/L), i=1, 2, \dots$$

Aunque esta aproximación tiene en cuenta la monitorización discreta y es muy flexible, el cálculo de dichos límites es más complicado puesto que utilizan integración numérica repetida. Jennison (1993)<sup>139</sup> ofrece fórmulas para el cálculo exacto de los límites.

### ***Diseños con funciones de gasto de $\alpha$ mediante EaSt***

Actualmente, existe un programa EaSt<sup>130</sup> que permite el diseño y monitorización de estudios con dicho diseño, pero siempre que se predetermine el número de análisis intermedios, y hasta un máximo de 10. Aunque es aplicable para variables binarias, normales y supervivencia, no permite el análisis final, (determinación del valor de  $p$ , estimación), y carece de la flexibilidad que pueden tener, en general, estos métodos basados en las funciones de gasto de  $\alpha$ .

## INTRODUCCIÓN

---

Emerson & Fleming (1989) <sup>127</sup> sugirieron métodos secuenciales por grupos que permitían la finalización temprana con el rechazo temprano de  $H_1$ , pero limitado a estudios con valores iguales de los errores de tipo I y II. Wang & Tsiatis (1987) <sup>140</sup> aproximaron los métodos secuenciales por grupos mediante límites de gasto para el rechazo de  $H_0$ .

Pampallona y Tsiatis 1994 <sup>141</sup> generalizaron el método de Wang y Tsiatis para rechazo temprano de  $H_0$  o  $H_1$  pudiéndose aplicar a tests de una o dos colas.

Se definen los diseños fijando los límites de continuación mediante sus funciones de gasto de alfa.

En el programa PEST diseñado para los métodos basados en límites continuos, se presentan en los gráficos los valores de  $V$  y  $Z$  en los ejes de coordenadas y abcisas respectivamente. En el EaSt en el eje de coordenadas se representa el valor de la distribución normal  $-U = Z/\sqrt{V}$ , no confundir  $U$  de distribución normal en esta tesis con el estadístico  $Z$  secuencial- versus fracciones respecto a la información total ( $V / V_{\text{máx}}$ ).

Los posibles diseños contemplan la combinación de las características (a) número de colas, y (b) objetivo de rechazar tempranamente  $H_0$ , o  $H_0$  y  $H_1$ . En cualquier de estas combinaciones, se deberá seleccionar un valor del parámetro de forma  $\Delta$  entre 0 y 0.5, y el número de inspecciones previstas.

En la tabla siguiente se muestran las posibles combinaciones de los dos primeros factores:

## INTRODUCCIÓN

---

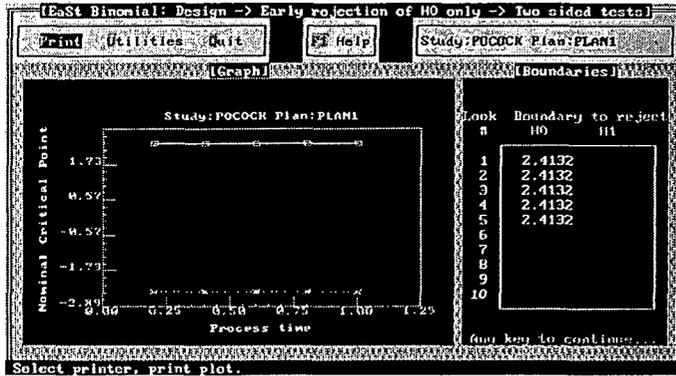
Modelo	Nº colas	Rechazo temprano de	Se finaliza si
A	1	H0	$Z \geq c * V^\Delta$
B	2	H0	$ Z  \geq c * V^\Delta$
C	1	H0 o H1	$Z \geq c * V^\Delta$ ° $Z \leq -c_2 * V^\Delta + \theta_R * V$
D	2	H0 o H1	$ Z  \geq c * V^\Delta$ ° $ Z  \leq -c_2 * V^\Delta + \theta_R * V$

En el caso de diseños a dos colas para rechazo temprano de la hipótesis nula, si el valor de  $\Delta$  es 0 estamos ante la equivalencia del modelo de O'Brien, y si el valor de  $\Delta$  es 0.5 el de Pocock (1977). Se muestran en los dos gráficos siguientes, el diseño de un ejemplo de estudio emulando los diseños de O'Brien & Fleming y de Pocock para cinco inspecciones. En el gráfico se representan los límites de aceptación de la hipótesis nula, en los ejes de coordenadas en valores de U (U de distribución normal, que no Z secuencial), para cada inspección, en el eje de abcisas.

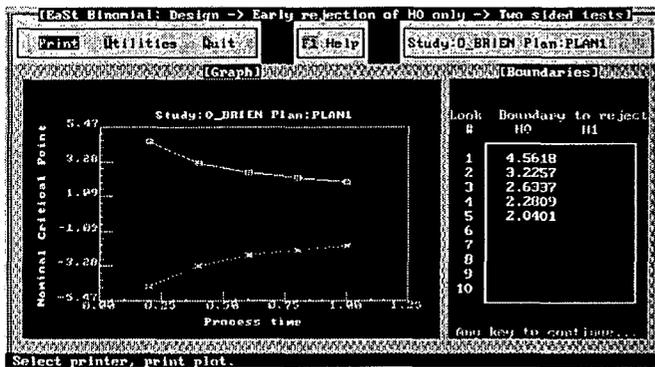
# INTRODUCCIÓN

Pocock:

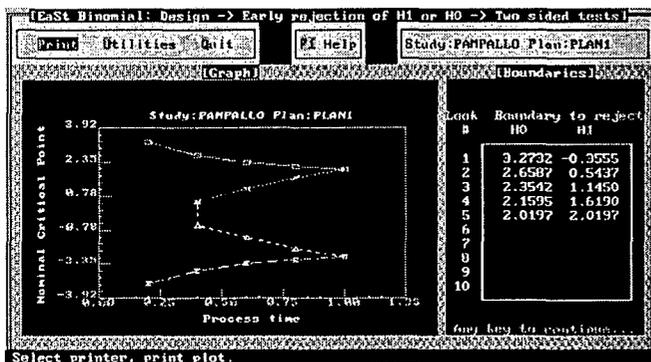
Rechazo temprano de  $H_0$ ,  $\Delta=0.5$



O'Brien & Fleming: Rechazo temprano de  $H_0$ ,  $\Delta=0$



Pampallona & Tsiatis 1994  $\Delta=0.2$



## INTRODUCCIÓN

---

Diferencias fundamentales entre los dos programas reconocidos para diseño secuencial:

	EaSt	PEST
1	Asume $Z \sim N(\theta V, V)$	$Z \sim N(\theta V, V)$
2	Variables: Normales Binarias Supervivencia	Normales Binarias Supervivencia Ordinales General ( $\theta$ )
3	Monitorización: discreta, función de gastos de $\alpha$	discreta, árbol de Navidad
4	Permite: Diseño Monitorización	Diseño Monitorización Análisis Simulaciones
5	Cálculo de Z y V: No	Sí
6	Duración de supervivencia basada: distribución exponencial	aproximación discreta

### **Métodos basados en límites continuos (basados en la función de verosimilitud)**

Son originales de Wald <sup>84</sup> y Barnard <sup>142</sup> en la década de los 40. Fueron desarrollados por Anscombe <sup>143</sup> en 1963, y han sido estudiados con detalle y perfeccionados por Whitehead y Jones <sup>144</sup>, permitiendo plantear una amplia variedad de diseños secuenciales.

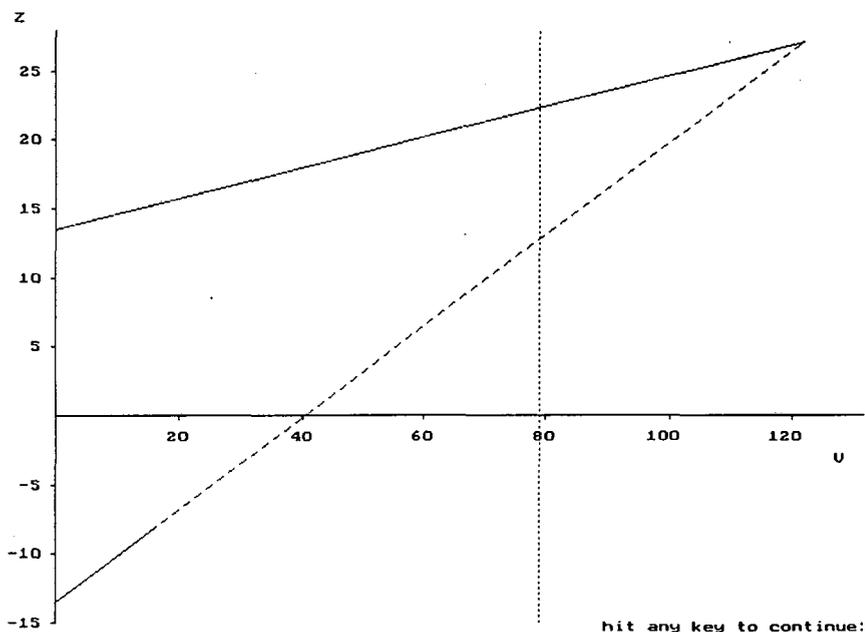
Consisten en definir dos fórmulas estadísticas a partir de los datos que se van acumulando. Una, llamada Z, valora la diferencia acumulada entre los dos tratamientos. Otra, llamada V, calcula la información acumulada y es proporcional al tamaño de la muestra. Cada vez que se desea realizar un análisis intermedio se calculan los valores de Z y V y se dibujan en una gráfica en la que se ha definido una región de continuación.

Así, mientras las coordenadas de (V,Z) se encuentren en el interior de esa región el estudio debe continuar. En el momento en que se cruce cualquiera de los límites de esa región, el estudio finaliza. En los diseños más sencillos, a una cola, si se cruza el límite superior, se debe a que existen diferencias a favor del nuevo tratamiento. Si se cruza el límite inferior, el estudio finaliza con resultado negativo. Aún en diseños a una cola, si se cruza el límite inferior muy cerca del origen, se puede concluir diferencias, de gran magnitud, a favor del tratamiento de control.

## INTRODUCCIÓN

---

Ejemplo de modelo secuencial de límites continuos con región de triangular para hipótesis unilateral



En diseños a dos colas hay dos zonas simétricas a partir del eje de abscisas. Si el corte se realiza por uno de los límites superior o inferior se concluirá superioridad del tratamiento experimental o control respectivamente. En caso contrario se aceptará la hipótesis nula.

Las ventajas de estos diseños respecto a los secuenciales por grupos radican en su flexibilidad. Pueden realizarse tantos análisis intermedios como se desee, de manera regular o irregular. Incluso se puede calcular el valor de Z y V tras la finalización de la participación de cada paciente, pero es importante que la frecuencia de las inspecciones no dependa de que el camino se esté aproximando a un límite.

Además, mientras que el diseño secuencial por grupos debe emplear la totalidad de la muestra para llegar a esa conclusión, este tipo de diseño permite finalizar pronto un estudio con resultados negativos. Se pueden emplear regiones de continuación con formas variadas. La selección de una u otra dependerá de las características y objetivos del estudio.

El cálculo de la posición de los límites se basa en métodos matemáticos laboriosos, pero existe un programa de ordenador <sup>114</sup> que realiza estos cálculos y que proporciona datos teóricos sobre el tamaño de muestra esperado en función del diseño y las características del ensayo. En promedio, los datos teóricos apuntan que estos métodos permiten un ahorro de alrededor del 40% de los pacientes necesarios en un diseño fijo. Existe siempre, sin embargo, una pequeña proporción (menor del 10%) de estudios que precisan más pacientes que los necesarios en un diseño clásico. Una buena revisión de estos métodos bajo una aproximación práctica complementada con el desarrollo estadístico-matemático se puede encontrar en el libro de Whitehead "The design and analysis of sequential clinical trials" <sup>145</sup>.

### *Otros contextos*

Aunque la base de estos métodos es inicialmente aplicable a diseños paralelos, con comparación de dos grupos y de igual razón de asignación (1:1), se pueden aplicar a otras situaciones:

Razón de asignación aleatoria distinta a 1:1.

Comparación de más de dos grupos, por ejemplo estudios de dosis respuesta.

Grupos no paralelos

### *Diseño*

Para diseñar un estudio usando la metodología secuencial de límites continuos, hace falta, al igual que en un diseño clásico, la especificación de:

$\alpha$  = nivel de significación

$1-\beta$  = poder

$\theta_R$  = la diferencia considerada relevante a buscar, siendo  $\theta_R > 0$

Dadas las hipótesis:

H0:  $\theta_R = 0$ ; los tratamientos son equivalentes

H1:  $\theta_R \neq 0$ ; los tratamientos son distintos

si  $\theta = \theta_R$ , entonces H0 debe ser rechazada en favor de H1 al nivel de significación  $\alpha$ , con probabilidad  $(1-\beta)$ .

Si  $\theta > 0$ , el tratamiento experimental -E- es mejor, si  $\theta < 0$  E es peor. El tratamiento control se expresara como C.

Se definen dos funciones dependientes de  $\theta$ :

$C^+(\theta) = P(\text{rechazar } H_0 \text{ a favor de E; } \theta)$

$C^-(\theta) = P(\text{rechazar } H_0 \text{ a favor de C; } \theta)$

$C^+(0) + C^-(0) = \alpha$

Siguiendo la nomenclatura utilizada por Whitehead {2316}, los modelos con *requerimiento de poder II* quedan contemplados en los diseños simétricos:

$C^+(\theta_R) = 1-\beta$  y  $C^(-\theta_R) = 1-\beta$

Estos diseños son útiles para conocer si

E>C,

E=C, y

E<C.

## INTRODUCCIÓN

---

En cambio, los modelos con *requerimiento de poder I* quedan referidos a diseño asimétricos:

$$C^+(\theta_R) = 1-\beta \quad \text{y} \quad C^(-\theta_R) \ll 1-\beta$$

En éstos no se controla  $C^(-\theta_R)$ . Esta opción es útil para conocer si:

(E: tratamiento experimental, C: tratamiento control)

$$E > C,$$

$$E \text{ no } > C$$

pero no para

$$E = C \text{ y}$$

$$E < C$$

De esta manera tan pronto como esté claro que E no va a ser mejor que C el estudio se puede finalizar y abandonar el fármaco E.

Los límites de la región de continuación se construyen de acuerdo al modelo seleccionado y de manera que cumplan con los requerimientos del poder establecido.

### *Modelos*

Hay múltiples formas posibles de la región de continuación. Históricamente el primer modelo fue el test razón de probabilidades secuencial (RPST) descrito por Wald (1947) [Wald & Wolfowitz (1948) <sup>146</sup>] (en inglés sequential probability ratio test –SPRT-). Este diseño con límites paralelos que optimiza el tamaño muestral si  $\theta_R \geq \theta \leq 0$ . Al ser abierto, no queda limitado el tamaño muestral máximo.

Posteriormente se describió un modelo cerrado con Armitage (1957) <sup>147</sup> y Anderson (1969) quienes propusieron el modelo RPST truncado, como forma de cerrar la región de continuación fijando una N máxima.

El modelo triangular fue inicialmente descrito por Anderson (1960) <sup>148</sup> y desarrollado por Whitehead (1983) <sup>149</sup>. Minimiza  $E(V^*)$  bajo  $\theta = \theta_R$

Otros modelos como el Open Top Design, no finalizan si  $E > C$  y pueden ser útiles para estimar  $\theta$  con precisión, y en función de que la pendiente de la recta inferior sea positiva o negativa la utilidad es distinta. En el primer caso finaliza bajo  $\theta = -\theta_R$  con lo que es útil para seguridad, mientras que en el segundo finaliza pronto bajo  $\theta = 0$  con lo que supone un ahorro económico.

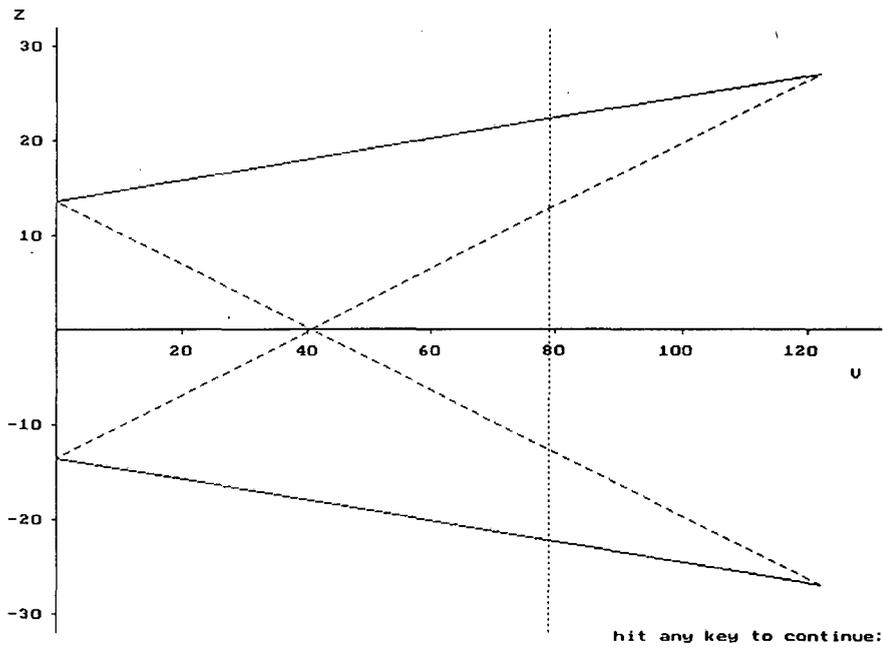
<b>Modelo</b>	<b>Número de colas</b>
Triangular	1 ó 2 colas
Razón de probabilidad Secuencial (RPST)	1 ó 2 colas
Razón de probabilidad Secuencial truncado (RPST <sub>x</sub> ; x= valor de N máxima)	1 ó 2 colas
Restricción estocástica (RP)	Sólo dos colas
Diseño superior abierto (DSA)	Sólo una cola

# INTRODUCCIÓN

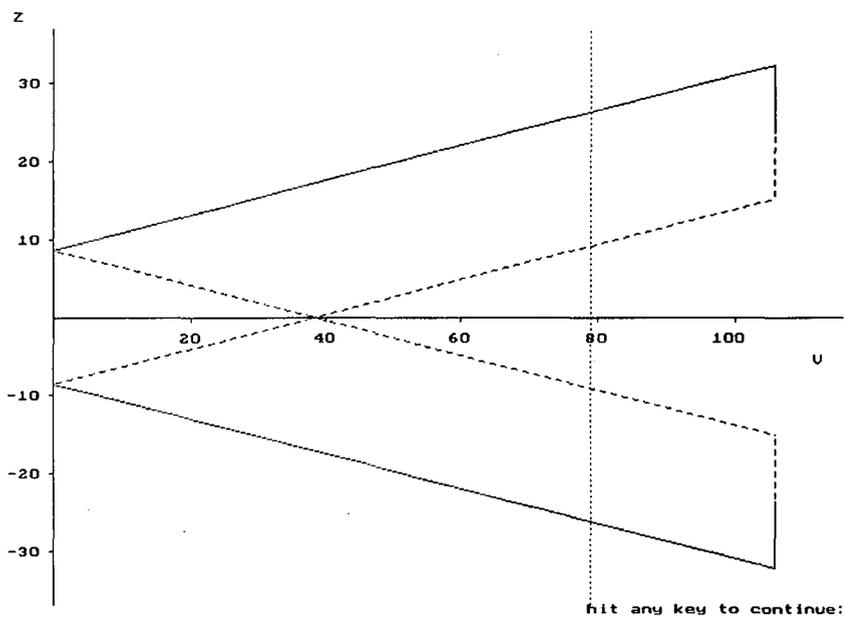
---

Ejemplos de modelo secuencial de límites continuos con distintos tipos de región de continuación para test de hipótesis bilaterales.

Modelo triangular



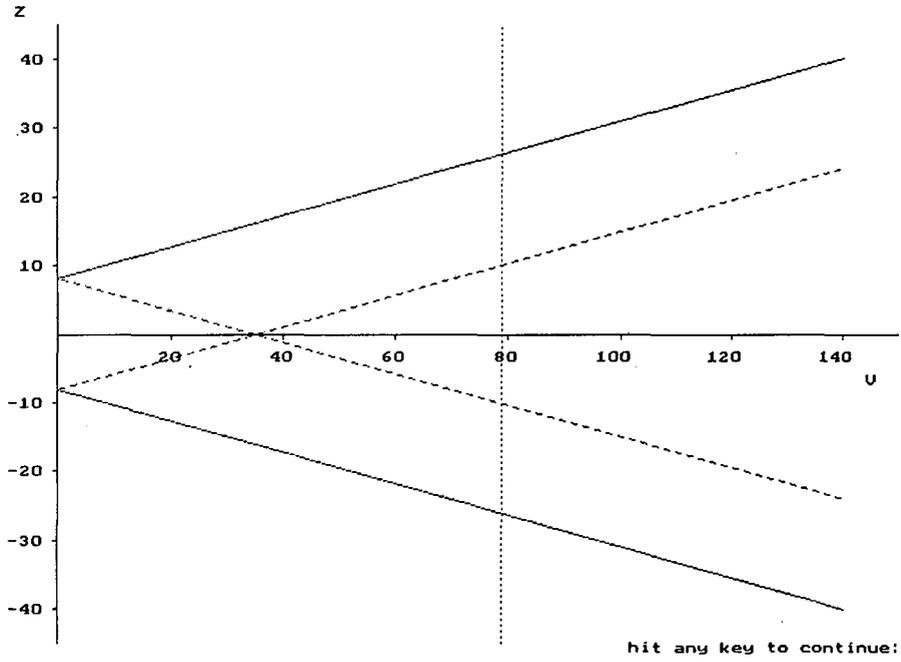
Modelo RPST truncado



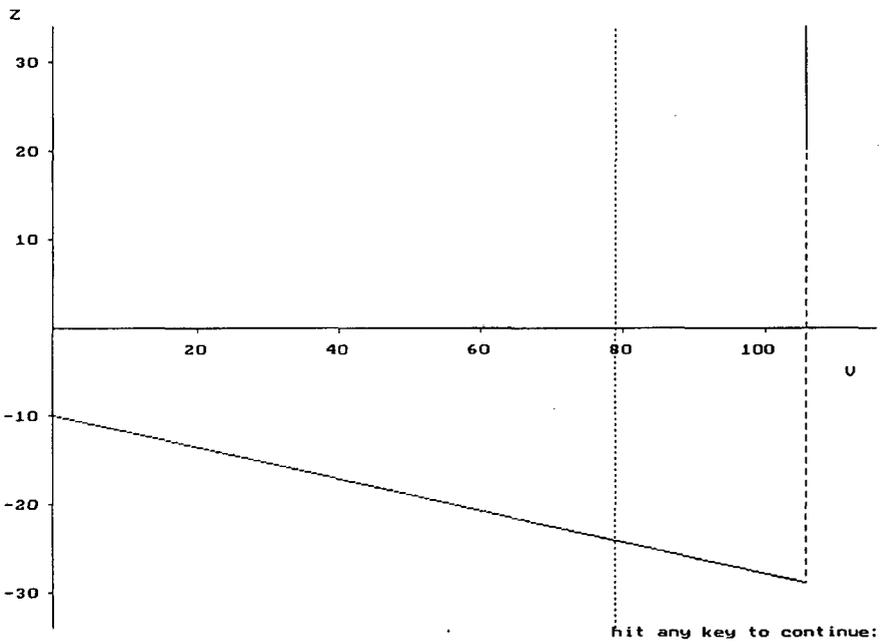
# INTRODUCCIÓN

---

## Modelo RPST



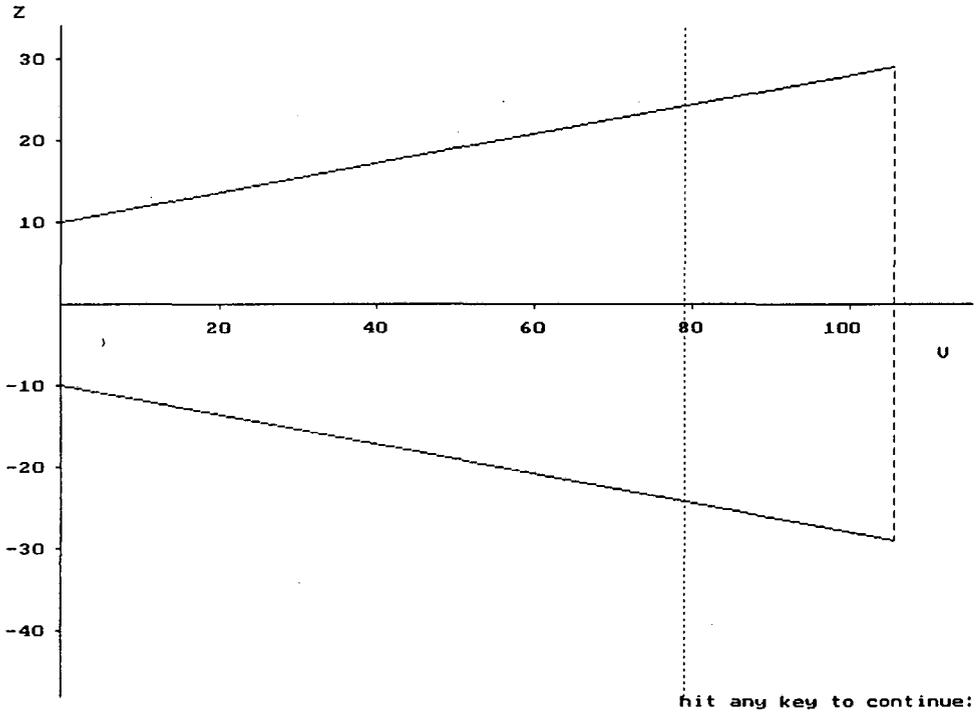
## Modelo de diseño superior abierto (Open Top Design)



# INTRODUCCIÓN

---

Modelo procedimiento restringido "restricted procedure"



### Monitorización

En cada inspección  $i$  se calcula el valor de las coordenadas dadas por  $(V_i, Z_i)$ , y si sus valores quedan dentro de la región delimitada por los límites de continuación se sigue con el estudio. En caso contrario se procederá al análisis definitivo.

### *Plan de inspecciones*

La formulación matemática solamente tiene un requisito: que  $V_i$  sea independiente de  $Z_1, \dots, Z_i, i=1, 2, \dots$

Lan & DeMets (1989) <sup>134</sup> exploran los efectos de incrementar las inspecciones cuando se está cercano a los límites de finalización, en violación del requisito, y constataron que los efectos son despreciables. Estas conclusiones fueron confrontadas por Proschan et al. (1992) <sup>150</sup> §.

Consideraciones prácticas:

Deben ser fijados con antelación y podrían ser:

- por número de respuestas,
- por tiempo,
- por número de muertes,
- por fechas de reuniones.

Los planes pueden ser flexibles, por ejemplo si se reduce la tasa de reclutamiento, reducir la frecuencia <sup>112</sup>.

### *Corrección por monitorización discreta*

La teoría fue inicialmente diseñada para monitorización continua y posteriormente adaptada a la monitorización discreta mediante una aproximación. La “inspección continua” significaría que  $Z$  y  $V$  se van calculando y representando continuamente. No obstante, dicho hecho es una abstracción matemática, puesto que incluso en el caso de que el incremento de información esté suministrada por un caso, la monitorización claramente no es continua.

## INTRODUCCIÓN

---

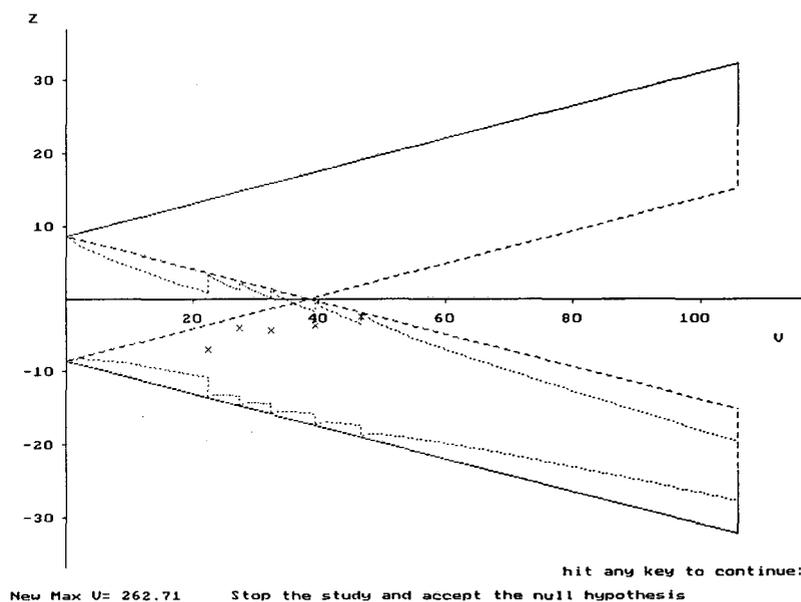
Whitehead, muestra como  $Z$  se comporta como un movimiento Browniano con tendencia de  $\theta$ . Algunas clarificaciones y ejemplos específicos pueden verse en Selke & Siegmund (Biometrika 1983)<sup>151</sup>, Lan & Zucker<sup>152</sup> y Tsiatis et al. (Biometrika 1995)  $\{S\}$ . Otros autores realizan cálculos repetidos de integración numérica para abordar la monitorización discreta.

Se aproxima la monitorización discreta por el método llamado corrección en “árbol de Navidad”, en concordancia a la forma que van cogiendo los nuevos límites corregidos a medida que se van realizando las inspecciones. Los límites nuevos son:

$$u_i = a + b \cdot V_i - 0.583 \cdot \sqrt{(V_i - V_{i-1})}$$

$$l_i = -a - b \cdot V_i + 0.583 \cdot \sqrt{(V_i - V_{i-1})}$$

y se van construyendo a medida que se procesa la información.



### *Análisis al finalizar el estudio*

#### *Nivel de significación*

Una vez finalizado el ensayo se requiere una ordenación de todos los posible valores finales  $(V_T, Z_T)$  de los estadísticos  $(V, Z)$ . Entre los posibles métodos estocásticos tenemos los descritos por Fairbanks y Madsen (Biometrika 1982) <sup>153</sup>, Rosner y Tsiatis (Biometrika 1988) <sup>154</sup>, Chang (Biometrics 1989) <sup>155</sup> y Emerson y Fleming (Biometrika 1990) <sup>156</sup>. Las diferencias entre estos métodos son escasas y Whitehead se decanta por el uso del primero y desaconseja otros métodos no estocásticos de ordenación.

Una vez seleccionado el método se puede definir la función del p-valor,  $P(\theta)$ . El valor de dicha función es la probabilidad, si la ventaja de E sobre C es igual a  $\theta$ , de obtener por azr una diferencia igual o mayor a la observada.

El valor de significación de  $p^+$  para el test de

$$H_0: \theta=0 \text{ versus}$$

$$H_1: \theta>0$$

es justamente  $p^+=P(0)$ .

Para la alternativa a dos colas, el nivel de significación es

$$p = 2 \min (p^+, 1-p^+).$$

#### *Estimación e intervalos de confianza*

A partir de los valores terminales de Z y de V se puede obtener un estimador de máxima verosimilitud del estimador de las diferencias.

$$\hat{\theta} = Z / V$$

Sin embargo, este estimador está sesgado al finalizar una prueba secuencial ya que ésta finalizó al encontrarse un valor de  $Z$  muy diferente de 0. Si se cruzan los límites externos,  $\hat{\theta}$  tiende a sobrestimar las diferencias reales y se cruzan los límites de manera que se acepta  $H_0$  nos encontramos frente a una infraestimación de las diferencias.

Una posible solución es utilizar un estimador no sesgado respecto a la mediana  $\theta_M$ . Los valores de este estimador se pueden calcular a partir del valor terminal de  $V$ .

### *Estimador no sesgado de $\theta$*

El estimador por máxima verosimilitud calculado al final del estudio

$$\hat{\theta} = Z / V$$

está sesgado. En caso de finalización aceptando  $H_0$ ,  $\hat{\theta}$  está en realidad infraestimando las diferencias, mientras que en caso de rechazar  $H_0$ , las está sobrestimando.

Whitehead (1986 *Biometrika*)<sup>157</sup> propone en su artículo una solución para conseguir el estimador insesgado de  $\theta$ ,  $\tilde{\theta}$ . Así, se podrá calcular  $\tilde{\theta}$  restandole el efecto del sesgo, siendo éste, el estimador considerado de elección.

$$\tilde{\theta} = \hat{\theta} - b(\tilde{\theta})$$

La solución se aproxima por un proceso iterativo el cual está implementado en el programa PEST v3.0.

El error estándar de  $\hat{\theta}$  es  $1/\sqrt{V}$ . Este error estándar también es incorrecto al finalizar un estudio secuencial y no se debe utilizar para el cálculo de los intervalos de confianza. La solución consiste en realizar una búsqueda de los límites de  $\theta_L$ , y  $\theta_U$  que comprendan a  $\theta$  con un 95% de probabilidad. El estimador no sesgado respecto a la mediana  $\theta_M$  se encontrará entre los dos límites pero no necesariamente en el centro.

Un intervalo de confianza al 90% ( $\theta_L, \theta_U$ ) para  $\theta$  ajustado por la naturaleza secuencial del análisis, se puede hallar resolviendo  $P(\theta_L)=0.05$  y  $P(\theta_U)=0.95$ , siendo  $P(\theta)$  la función definida por la ordenación estocástica.

### *Análisis estratificado y Análisis multivariante*

Cuando se prevea la necesidad de ajustar el efecto de uno o varios factores, los estadísticos  $Z$  y  $V$  pueden adaptarse a esta situación. Si se definen  $s$  estratos

$$Z = \sum_{i=1}^s Z_i \qquad V = \sum_{i=1}^s V_i$$

Cuando el número de posibles factores modificadores del efecto es grande, el análisis estratificado puede presentar problemas debido a la existencia de múltiples estratos vacíos. En estas situaciones los métodos multivariantes son preferibles.

Los estadísticos  $Z$  y  $V$  pueden definirse para un análisis de regresión lineal generalizada, concretamente para las regresiones logística, múltiple o de Cox en función del tipo de variable dependiente. Para el cálculo ajustado de  $Z$  y  $V$  se debe recurrir a otro paquete estadístico distinto del PEST v3.0, como SAS, BMDP o GLIM. Los detalles técnicos se describen en Whitehead J (1983 y 1992) y en el manual del programa PEST v3.0.

### *“Underruning” y “OVERRUNING”*

Puede producirse que un estudio presente cambios respecto al plan de monitorización y análisis. En primer lugar puede suceder que se finalice el estudio por alguna causa ajena al estudio, como por ejemplo cambios en estrategia terapéutica, publicaciones recientes, o bajo reclutamiento. En segundo lugar, los datos se pueden ir acumulando después de alcanzar el criterio de finalización y puede ser deseable utilizarlos. Nos estamos refiriendo a “underruning” y “overruning” respectivamente y ambos problemas tienen solución. Estos problemas fueron estudiados por Anderson <sup>158</sup> y Choi y Clark <sup>159</sup> y detallados con aproximación pragmática por Whitehead <sup>160</sup>.

En el caso del “underruning” es muy importante que la decisión de finalizar sin haber cruzado los límites sea totalmente independiente de los valores de Z, es decir de la trayectoria de las diferencias entre tratamientos. Por este motivo, no solamente es necesario que la decisión se lleve bajo enmascaramiento de los tratamientos, sino que además el comité que tome la decisión debe ser ciego también a los valores de Z.

Para solucionarlo, se define una nueva función de p-valor cuyo cálculo está implementado en el programa PEST v3.0 <sup>114</sup> al igual que el “overruning”, para detalles ver Whitehead <sup>112</sup>.

El “overruning” tiene más difícil solución y en muchos casos es previsible e inevitable. Si se producen cambios en el manejo general del estudio, o de los pacientes, se recomienda no incluir las nuevas inclusiones. Whitehead recomienda una aproximación pragmática ignorando la penúltima inspección.

Evidentemente, en estos casos el poder al final del estudio puede presentar incrementos y descensos de distinta magnitud en función de la relevancia del “overruning” y “underruning”.

### *Estimación del efecto absoluto de los tratamientos*

Como ya se ha comentado la estimación de los efectos por máxima verosimilitud al final de un ensayo secuencial está sesgado. Esto mismo se aplica tanto a la estimación del efecto de las diferencias  $\hat{\theta}$  como a los efectos individuales. De esta manera, en un estudio en el que la variable de evaluación es binaria, la estimación de las proporciones en forma convencional no es correcta.

Dado que se puede obtener un estimador no sesgado de  $\theta$ , y que la proporción promedio,  $\bar{p}$  se puede estimar como  $S/N$  y es aproximadamente independiente de  $Z$  y  $\tilde{\theta}$ , las proporciones estimadas desde el estimador no sesgado de  $\theta$ , se calcularán a partir de:

$$\tilde{\theta} = \log \left\{ \frac{[\tilde{p}_E * (1 - \tilde{p}_C)]}{[\tilde{p}_C * (1 - \tilde{p}_E)]} \right\}$$

y de:

$$\bar{p}' = a_E * \tilde{p}_E + a_C * \tilde{p}_C$$

siendo  $a_E = N_E / N$  y  $a_C = N_C / N$ .

A partir del estimador no sesgado de  $\theta$  y del efecto promedio de los tratamientos, se deducen las fórmulas para otros tipos de variables. Una explicación detallada con ejemplos y aplicaciones para otros tipos de variables puede consultarse en el libro de Whitehead <sup>112</sup>.

### *Análisis de variables secundarias*

Durante un ensayo clínico se pueden observar muchas respuestas en un paciente. Cuando se establece el nivel de significación en el 5%, la probabilidad que una de esas respuestas evaluadas sea significativa es mayor que 0.05.

En el contexto de un ensayo clínico de diseño clásico a N fija, se pueden realizar ajustes. Si está claramente especificado en el protocolo que hay una única variable, y ésta será la comparación definitiva, se puede realizar el análisis sin ningún tipo de ajuste. Algunas veces, un pequeño grupo de respuestas o variables secundarias se identifican en el protocolo como de particular importancia.

Después de un ensayo secuencial, se añade un nuevo problema por la existencia de variables secundarias ya que pueden estar correlacionadas con la principal. Si para la variable principal quedan definidas  $\theta_1$ ,  $Z^{(1)}$  y  $V^{(1)}$ , para la variable secundaria se identifican  $\theta_2$ ,  $Z^{(2)}$  y  $V^{(2)}$ . La covarianza entre  $Z^{(1)}$  y  $Z^{(2)}$  la identificaremos como W y se asumirá que:

$$V^{(2)} = s^2 V^{(1)} \quad \text{y} \quad W = r s V^{(1)}$$

El estimador convencional de  $\theta_2$  mediante máxima verosimilitud es

$$\hat{\theta} = Z^{(2)} / V^{(2)},$$

el sesgo de  $\hat{\theta}$

$$b_2 = E(\hat{\theta}_2) - \theta_2.$$

A partir de  $b_1$ ,

$$b_2 = r/s * b_1$$

Si los dos análisis no están correlacionados ( $r=0$ ), entonces  $b_2 = 0$  y las estimación de  $\hat{\theta}_2$  no está sesgada. Además, si hay mucha más información (V) en el parámetro  $\theta_2$

## INTRODUCCIÓN

---

que en  $\theta_1$  ( $s^2$  grande), el sesgo de la variable secundaria será pequeño y puede despreciarse. En caso contrario es necesario aplicar una corrección.

El sesgo para  $\hat{\theta}_2$  puede calcularse a partir del sesgo de  $\hat{\theta}_1$ :

$$\tilde{\theta}_2 = \hat{\theta}_2 - r/s * (\hat{\theta}_1 - \tilde{\theta}_1).$$

La varianza se estimará a partir de

$$\text{var}(\tilde{\Theta}_2) = \frac{r^2}{s^2} \text{var}(\tilde{\Theta}) + \frac{1-r^2}{s^2} E\left(\frac{1}{V_1}\right)$$

Finalmente, el intervalo  $\tilde{\theta}_2 \pm 1.96 * \sqrt{\text{var}(\tilde{\theta}_2)}$  es conservador y probablemente proporcionará una cobertura superior al 95% <sup>161</sup>.

### Intervalos de confianza repetidos

Desarrollados por Jennison & Turnbull (1984) <sup>162</sup>. Si el intervalo excluye al 0, el estudio puede finalizar con un resultado positivo. Si se alcanza el tamaño de muestra final, el intervalo de confianza calculado entonces tendrá la propiedad de contener el valor real de la diferencia entre tratamientos con una probabilidad del 95%. Este diseño, aunque es aplicable al ámbito del ensayo clínico, donde parece ser más útil es para estudios epidemiológicos.

Cuando el interés de un estudio consiste en estimar con gran precisión el efecto de un tratamiento puede emplearse la técnica propuesta por Jennison y Turnbull. Ésta, está basada en el método secuencial por grupos, según el cual cada vez que se acumula un grupo de pacientes de tamaño preestablecido se calculan el intervalo de confianza, en general del 95%, de la diferencia entre tratamientos.

Se calculan un par de valores limitando un intervalo,  $(\theta_l, \theta_s)$ , a partir de los datos, de manera que la probabilidad es  $1-\gamma$ , de que  $\theta_l$  y  $\theta_s$  contengan el verdadero valor de  $\theta$  entre ellos.

Se podría transcribir como:

$$P\{(\theta_l, \theta_s) \in \theta\} = 1-\gamma$$

Generalmente  $\gamma$  es positivo pero pequeño, siendo la opción  $1-\gamma = 0.95$  la más extendida. En un ensayo a tamaño fijo, el intervalo de confianza se calcula una sola vez. Este método, crea una secuencia de intervalos, una vez calculados en cada inspección, con la propiedad de contener  $\theta$  con probabilidad  $1-\gamma$ .

$\theta_{li}$  y  $\theta_{si}$  para cada inspección  $i$

$$P\{(\theta_{li}, \theta_{si}) \in \theta \text{ para } i=1,2,\dots\} = 1-\gamma$$

Cada intervalo se debe referir como un intervalo de confianza repetido para distinguirlo de uno real. Además, cada intervalo posterior cumple

$$P\{(\theta_{li}, \theta_{si}) \in \theta\} \geq 1 - \gamma,$$

y es más conservador que un intervalo de confianza real.

Estos intervalos pueden ser calculados con el objetivo de tener una impresión de la magnitud del efecto del tratamiento durante el curso de un ensayo. Estos métodos están descritos ampliamente por Jennison & Turnbull (1984<sup>162</sup>, 1989<sup>137</sup>, y 1990<sup>163</sup>). El más claro y extenso es el artículo de 1989.

Métodos de uso:

Como diseño secuencial: Se finaliza el ensayo si el intervalo calculado en la inspección  $i$  no contiene el 0.

Como regla de finalización flexible: Aunque propuesto en 1989 por Jennison and Turnbull es muy criticado. Se predefine una diferencia relevante, por ejemplo  $\theta = 0.2$ , y durante el transcurso del estudio se permite redefinir el margen de la diferencia relevante. Difícilmente sería aceptable este diseño para un ensayo clínico por la subjetividad a la que está sujeta dicha decisión.

Sin reglas de finalización (Robbins 1970)<sup>164</sup>: En algunos estudios epidemiológicos, no es posible cesar la exposición a un fármaco, o a factor de riesgo, y ello puede ser aprovechado para seguir analizando los resultados. Aunque no se ha probado en la práctica, parece un diseño interesante para algún tipo de estudio.

### Restricción estocástica

La restricción estocástica es un método utilizado en los procedimientos de control de calidad. Un ejemplo de un esquema de inspección podría consistir en seleccionar y probar 100 utensilios manufacturados de un gran lote. Si 5 o más han salido defectuosos, se desecha el lote. Aunque el esquema inicial prevé testar 100 utensilios, tan pronto como se encuentran 5 defectuosos se deja de seguir con la prueba.

En el campo del ensayo clínico no parece que sea aplicable para evaluar eficacia comparada, máxime cuando el resultado final no puede especificarse más que como aceptar o rechazar una hipótesis. Este método fue introducido por Lan, Simon & Halperin (1982) <sup>165</sup>. Como pros tenemos (1) que en condiciones adecuadas permite detener prematuramente un diseño con tamaño fijo y (2) podría usarse en estadios iniciales para desechar fármacos con baja probabilidad de éxito. En relación de este segundo punto existen desarrollos con aplicación práctica clara {2110} \$. Estas extensiones se han utilizado básicamente en investigación oncológica {2310} \$,

### *Base metodológica\**

El criterio original de Lan, Simon & Halperin para finalizar sería cualquiera de:

Se puede calcular en cualquier momento, la probabilidad de superioridad experimental, condicional en (1) el verdadero valor de  $\theta$  y (2) los datos reclutados hasta el momento.

$$C^+(\theta, Z, V) = P(Z_{fija} \geq u_{\alpha/2} * \sqrt{V_{fija}} \mid Z, V; \theta)$$

La idea de la restricción estocástica es finalizar tan pronto como

$$C^+(0, Z, V) \leq \gamma'$$

donde  $\gamma'$  es positiva pero pequeña.

Expresando la fórmula anterior en función d Z y V:

$$Z \geq u_{\alpha/2} * \sqrt{V_{fija}} + u_{\gamma'} * \sqrt{(V_{fija} - V)}$$

## INTRODUCCIÓN

---

Una vez que se permite la finalización prematura, la probabilidad de nula de concluir que el tratamiento experimental es superior es ahora mayor que  $\alpha/2$ . En realidad es

$$\alpha/\{2*1-\gamma\}$$

El caso contrario

$$C(\theta, Z, V) = P(Z_{fija} \leq -u_{\alpha/2} * \sqrt{V_{fija}} \mid Z, V; \theta)$$

Para finalizar:

$$C(0, Z, V) \geq \gamma'$$

$$\Rightarrow Z \leq -u_{\alpha/2} * \sqrt{V_{fija}} - u_{\gamma'} * \sqrt{(V_{fija} - V)}$$

Si  $\gamma'=1/2$  es una manera alternativa de interpretar O'Brien y Fleming, y si  $\gamma'<1/2$  son diseños con menor probabilidad de finalizar pronto.

Como problema tiene que en caso de  $\theta=0$ , la probabilidad de concluir superioridad puede ser mayor a  $\alpha/2$  y tener un valor entre  $\alpha/2$  y  $\alpha/\{2*(1-\gamma)\}$ ; si  $\alpha=0.05$  y  $1-\gamma=0.95$ , el valor real puede estar entre 0.025 y 0.0263. En la práctica se podría sustituir  $\alpha$  por  $\alpha*(1-\gamma)$  para solucionar este problema.

### Métodos Bayesianos

La aproximación bayesiana a la inferencia estadística es totalmente diferente a la postura frecuentista. Su metodología se basa en la obra de un clérigo del siglo XVIII, Thomas Bayes. Las bases e introducción se pueden encontrar en Lindley (1965)<sup>166</sup>, dos tomos) y Barnett (1982)<sup>167</sup>, y ejemplos de aplicación médica en Spiegelhalter (1982)<sup>168</sup>, y Racine y cols (1987)<sup>169</sup>.

La estadística frecuentista se basa exclusivamente en extraer conclusiones a partir de la información observada mediante los contrastes de hipótesis (con los errores de I y II especie), y los intervalos de confianza (que no son la probabilidad de contener el valor poblacional puesto que es fijo) para extraer conclusiones.

La filosofía bayesiana es distinta. Antes de que el investigador recoja sus datos puede formular, basándose en el conocimiento previo, su distribución subjetiva de  $\theta$  llamada distribución a priori. La información aportada por los nuevos datos se combina con la distribución a priori para generar una distribución a posteriori más precisa sobre  $\theta$ .

Así, la inferencia bayesiana deriva completamente a partir de la distribución a posteriori. El estimador bayesiano de  $\theta$  es la medida de la distribución a posteriori. Los valores de  $\theta_L$  y  $\theta_U$  pueden ser escogidos de acuerdo a la distribución posterior,  $P(\theta \leq \theta_L) = 0.025$  y  $P(\theta \leq \theta_U) = 0.0975$ , formándose de esta manera el intervalo de "credibilidad" para  $\theta$ . En lugar de calcular un nivel de significación se puede calcular el valor de  $P(\theta \leq 0)$ , es decir la probabilidad (basada en la experiencia renovada) que el tratamiento experimental es inferior al control. Este cálculo es interesante por tener una interpretación más intuitiva que el valor p, pero depende de una distribución a priori subjetiva.

En un diseño secuencial, el investigador es capaz de aumentar y renovar la información sobre  $\theta$  en cada inspección. No es necesario establecer a priori reglas de finalización, aunque se pueden definir de acuerdo a una mejoría mínima para concluir la recomendación del nuevo tratamiento o un empeoramiento máximo a partir de cual es mejor abandonar el tratamiento experimental. La metodología bayesiana no está afectada por el número de inspecciones con lo que resulta un método flexible y especialmente aplicable a las técnicas secuenciales. No obstante, no está tan desarrollada en el campo secuencial y no hay programas de uso aplicado.

El principal problema estriba en la distribución subjetiva en la que se basan los resultados posteriores. Evidentemente, la interpretación de ésta variará en función de los intereses: promotores, compañía competidora, otro grupo de investigadores, etc. De esta manera se prevé que no habrá acuerdo en el análisis de los resultados y de las conclusiones. Ya en algún caso se ha reanalizado los resultados de un estudio con distintas distribuciones a priori <sup>170</sup>).

La polémica persiste entre seguidores frecuentistas y bayesianos actualmente, y se puede observar en revistas de ámbito clínico discusiones acerca de la distintas interpretaciones de resultados en función del método <sup>171</sup> y <sup>172</sup>.

En general, aunque no universalmente, se pueden establecer correspondencias entre los métodos bayesianos y los métodos frecuentistas. Así, si se utiliza una distribución a priori adecuada y muy dispersa para indicar un estado de ignorancia inicial, las diferencias se minimizan y se obtienen las ventajas de las técnicas bayesianas con respecto al problema de las inspecciones <sup>173</sup>.

Una buena revisión desde el punto de vista práctico con ejemplos en salud pública y medicina puede encontrarse en <sup>174</sup>.

Aunque grupos de trabajo están desarrollando la metodología secuencial con base bayesiana <sup>175,176</sup>, estos diseños todavía no se pueden poner en práctica por problemas de falta de software para resolver los laboriosos cálculos.

### **Equivalencia terapéutica secuencial**

Los estudios de bioequivalencia suelen realizarse con un número pequeño de pacientes, pero en los estudios de eficacia en fase III, podemos encontrar en algunos casos con un número muy grande de pacientes, hasta varios miles. En este contexto, parece adecuado la implementación de la metodología secuencial con el principal objetivo de ahorrar pacientes. Además, a diferencia de los diseños clásicos, con un método secuencial se puede garantizar el poder al finalizar el estudio.

J Whitehead (1996) <sup>177</sup> presenta un artículo con las bases para implementar la equivalencia terapéutica a la estadística secuencial. En este trabajo muestra la aplicación directa de dos modelos a un ensayo clínico publicado <sup>178</sup> mediante simulación a partir de sus resultados.

En la práctica los tratamientos E y C son equivalentes si  $\theta \in (-\theta_R, \theta_R)$ .

Presenta 3 alternativas para poder concluir:

Aproximación por poder:

Si no se encuentran diferencias entre E y C, al nivel de significación  $\alpha$ , se concluye equivalencia.

Intervalos de confianza:

Si el intervalo porcentual construido de manera  $100(1-2\beta)$  incluye  $(\theta_L, \theta_U)$  se concluye equivalencia.

## INTRODUCCIÓN

---

Test doblemente unilateral:

Si las hipótesis nulas correspondientes a:

$$H_0(1): \theta = -\theta_R \quad \text{y} \quad H_0(2): \theta = \theta_R$$

son rechazadas a favor de

$$H_1(1): \theta > -\theta_R \quad \text{y} \quad H_1(2): \theta < \theta_R,$$

ambas al nivel de significación  $\beta$ , entonces se concluye equivalencia. Este es el método de equivalencia recomendado por Schuirmann (1987) <sup>48</sup>.

A nivel teórico, dado que la metodología secuencial mantiene las especificaciones de poder del diseño, los tres métodos son equivalentes. No obstante, a nivel práctico el mantenimiento del poder se garantizaría en una monitorización continua y siempre que el corte se hiciera exactamente en el límite. Además en algunos casos los estados de “underrunning” y “overrunning” llevan consigo incrementos y descensos en el poder. Como conclusión se desaconseja el método (1) y los métodos (2) y (3) son equivalentes.

La elección del valor de  $\beta$  es arbitrario, no obstante los valores más usados son 0.05, y 0.025. Para los diseños de equivalencia terapéutica se exige a veces valores más restrictivos como 0,025.

El valor de  $\alpha$  se puede aumentar un poco, por ejemplo a 0.1, para optimizar el tamaño muestral. En este caso si se cruzan los límites externos cerca del origen se concluiría superioridad de uno de los dos tratamientos, pero se corre el riesgo de que si se cruzan los límites externos más tardíamente no se pueda concluir ni equivalencia ni superioridad de ninguno de los dos tratamientos.

Whitehead se decanta en su trabajo por dos modelos con un comportamiento similar: (1) el modelo triangular bilateral, el cual optimiza  $E(t)$  bajo  $H_0$ , aunque haya modelos más favorables para  $P_{50}(t)$ , y (2) el modelo de Pampallona y Tsiatis con parámetro de

## INTRODUCCIÓN

---

forma  $\Delta=0.2$  intermedio. No obstante, la selección del modelo adecuado se debe hacer conforme al objetivo y es aconsejado realizar simulaciones previas.

En caso de querer demostrar que el tratamiento E no es peor que el C, Whitehead propone un diseño triangular reverso de un solo brazo. Se puede concluir que E no es mejor que C si:

se cruza el límite superior, o

$-\theta_R \leq (\theta_L, \theta_U)$  siendo  $(\theta_L, \theta_U)$  los límites superior e inferior de intervalo de confianza porcentual  $100*(1-2\beta)$ , o

$H_0$  se rechaza al nivel  $\beta$ ,

siendo las tres opciones son equivalentes.

## OBJETIVOS

### *Objetivos principales*

Esta tesis se encamina en la evaluación del posible beneficio terapéutico de un antiagregante en la profilaxis secundaria del post-infarto inmediato mediante un ensayo clínico controlado. Los objetivos que se plantea son los siguientes objetivos:

1. Búsqueda de un diseño metodológico-estadístico teórico óptimo
2. Comprobación de los supuestos teóricos previamente a la realización del estudio mediante simulaciones
3. Realización del diseño, monitorización del estudio y análisis estadístico de los resultados

### *Objetivos secundarios*

1. Revisión y descripción de métodos alternativos al diseño clásico
2. Determinación del cumplimiento de las especificaciones teóricas estadísticas de los modelos más favorables
3. Estudio del ahorro del tamaño muestral de los métodos secuenciales respecto a los diseños clásicos

## MATERIAL Y MÉTODO

### *PLANIFICACIÓN DEL DISEÑO*

Dado que esta tesis es básicamente metodológica, en ésta no vamos a tratar temas de razonamientos terapéuticos ni recomendaciones farmacológicas acerca de los datos del estudio. Solamente se entrará a comentar aquellos aspectos que pudieran haber influido en las decisiones del diseño, seguimiento o análisis definitivo del estudio TIM.

De acuerdo con las tendencias de los grandes ensayos clínicos publicados, como el estudio ISIS-2<sup>179</sup>, y las recomendaciones de los grandes grupos de trabajo en prevención cardiovascular, como el Antiplatelet Trialist's Collaboration<sup>180</sup>, se seleccionó como objetivo principal la reducción comparada entre la aspirina y triflusal en una variable compuesta por muerte o infarto no fatal o accidente vascular cerebral en el periodo agudo post-infarto -35 días-. El grupo control, aspirina, fue escogido basándose en las evidencias publicadas hasta aquel momento -finales de 1992-.

#### *Diferencia relevante*

De acuerdo con los datos del estudio ISIS-2, la reducción para la variable compuesta por IAM no fatal, AVC no fatal o muerte de cualquier causa para AAS vs placebo era del 30% con una proporción en placebo del 10%. En término de porcentajes absolutos equivale a describir los resultados como 10% de fracasos en el grupo de placebo y 7% en el de aspirina.

En el estudio TIM se estableció la magnitud de la diferencia clínicamente relevante en el 25%. Si el fármaco experimental o control fuera un 25% peor se concluiría

ineficacia con probabilidad de detectarlo de (1-Beta), y tendría eficacia similar a placebo.

De esta manera se asegura un tamaño muestral alto, de varios millares de pacientes, y si el estudio sale negativo la eficacia del fármaco perdedor se aleja en el peor de los casos un 5% de los valores del placebo.

*Características estadísticas*

El nivel de significación se fijo en el 5% bilateral convencional. El diseño unilateral fue descartado por que aunque optimiza el tamaño muestral, existe gran controversia en aceptarse a menos que el grupo de comparación pueda ser el placebo. Además, contrastando que el tratamiento experimental no es peor que el control, perderíamos la posibilidad de demostrar si el experimental es mejor. Los modelos secuenciales de límites continuos sí que detectan en diseños a una cola diferencias en sentido contrario, pero éstas tienen que tener una gran magnitud que no es esperable en este caso.

Dado que hay un abanico de posibilidades en la relación de proporción de eventos en ambos grupos, OR, theta, diferencias en porcentaje relativo y absoluto, y tamaño muestral, se presentan a continuación en dos tablas. En la primera se incrementan valores fijos del 5% en la diferencia en porcentaje absoluto de eventos y en la segunda se incrementa 0.25 en los valores de theta. En ambos casos se presentan valores de tamaño muestral por la formula convencional y por la aproximación secuencial.

Diferencias relativas	Diferencias absolutas	P(con)	P(exp)	OR	Theta	Theta Ratio	N (theta)	N
-40.0%	-4.0%	10,00%	6,00%	0,574	0,5543	1,759	1388	1436
-35.0%	-3.5%	10,00%	6,50%	0,626	0,4689	1,488	1886	1932
-30.0%	-3.0%	10,00%	7,00%	0,677	0,3895	1,236	2661	2705
-25.0%	-2.5%	10,00%	7,50%	0,730	0,3151	1,000	3961	4003
-20.0%	-2.0%	10,00%	8,00%	0,783	0,2451	0,778	6380	6420
-15.0%	-1.5%	10,00%	8,50%	0,836	0,1790	0,568	11666	11705
-10.0%	-1.0%	10,00%	9,00%	0,890	0,1164	0,369	26947	26984

## MATERIAL Y MÉTODO

-5,0%	-0,5%	10,00%	9,50%	0,945	0,0568	0,180	110460	110496
0,0%	0,0%	10,00%	10,00%	1,000	0,0000	0,000	----	----
5,0%	0,5%	10,00%	10,50%	1,056	-0,0544	-0,173	115486	115520
10,0%	1,0%	10,00%	11,00%	1,112	-0,1065	-0,338	29464	29496
15,0%	1,5%	10,00%	11,50%	1,169	-0,1566	-0,497	13349	13380
20,0%	2,0%	10,00%	12,00%	1,227	-0,2048	-0,650	7646	7676
25,0%	2,5%	10,00%	12,50%	1,286	-0,2513	-0,798	4979	5008
30,0%	3,0%	10,00%	13,00%	1,345	-0,2963	-0,940	3514	3542
35,0%	3,5%	10,00%	13,50%	1,405	-0,3398	-1,078	2623	2650
40,0%	4,0%	10,00%	14,00%	1,465	-0,3819	-1,212	2038	2064
45,0%	4,5%	10,00%	14,50%	1,526	-0,4229	-1,342	1633	1659

Theta (fracciones)	ThetaR Ratio	Theta	OR	P(con)	P(exp)	Diferencias absolutas	Diferencias relativas	N (theta)	N
3/2 (+ThetaR)	1.50	0.473	0.623	10.00%	6,48%	-3,52%	-35,22%	1860	1906
5/4 (+ThetaR)	1.25	0.394	0.674	10.00%	6,97%	-3,03%	-30,29%	2605	2649
(+ThetaR)	1.00	0.315	0.730	10.00%	7,50%	-2,50%	-25,00%	3961	4003
3/4 (+ThetaR)	0.75	0.236	0.790	10.00%	8,07%	-1,94%	-19,35%	6842	6882
1/2 (-ThetaR)	0.50	0.158	0.854	10.00%	8,67%	-1,33%	-13,31%	14952	14991
1/4 (+ThetaR)	0.25	0.079	0.924	10.00%	9,31%	-0,69%	-6,87%	57988	58024
0 (+ThetaR)	0.00	0.000	1,000	10.00%	10,00%	0,00%	0,00%	----	----
1/4 (-ThetaR)	-0.25	-0.079	1,082	10.00%	10,73%	0,73%	7,30%	54695	54728
1/2 (-ThetaR)	-0.50	-0.158	1,171	10.00%	11,51%	1,51%	15,10%	13177	13208
3/4 (-ThetaR)	-0.75	-0.236	1,267	10.00%	12,34%	2,34%	23,40%	5652	5681
(-ThetaR)	-1.00	-0.315	1,370	10.00%	13,21%	3,21%	32,10%	3090	3118
5/4 (-ThetaR)	-1.25	-0.394	1,482	10.00%	14,14%	4,14%	41,40%	1910	1936
3/2 (-ThetaR)	-1.50	-0.473	1,604	10.00%	15,13%	5,13%	51,30%	1278	1303

### *Estudio de alternativas al diseño clásico*

#### Diseños contemplados

## MATERIAL Y MÉTODO

---

Para estudiar la aplicabilidad de las distintas opciones existentes al estudio TIM se hizo una primera aproximación de acuerdo a las características teóricas de los posibles diseños. Los puntos que se tuvieron en cuenta fueron:

Tamaño muestral requerido. Existencia de la posibilidad de finalización previa al tamaño muestral preestablecido máximo y cuantificación de ésta.

Cumplimiento de las especificaciones estadísticas

Plausibilidad del método para este estudio

Reconocimiento y aceptación por parte de la comunidad científica

### *Requerimientos de tamaño muestral*

Estimadores de las previsiones teóricas

En el diseño clásico general las previsiones son fijas y el estudio no se finaliza hasta que se recluta el tamaño preestablecido. No obstante para algunos modelos secuenciales el tamaño muestral se comporta como una variable aleatoria y se debe medir en función de unos estimadores que a continuación se describen.

Para el estudio del tamaño muestral de los modelos secuenciales se escogió como patrón de medida principal el percentil 50 - $P_{50}(t)$ -. Este estimador central ofrece como ventaja respecto al valor promedio - $E(t)$ - que no está sesgado por los valores extremos. Indica qué valor de tamaño muestral se requiere superar en el 50% de estudios hipotéticos realizados.

Como medida de valores extremos se seleccionó el percentil 90 - $P_{90}(t)$ - que nos indica aquel valor que solamente sería superado en un 10% de hipotéticos estudios realizados.

El valor máximo teórico  $Máx(t)$  se encuentra en concordancia al valor del estadístico  $V$  cuando éste está en el punto más alejado del origen. Los puntos máximos corresponden a los puntos de truncamiento en los modelos RPST truncados y a los vértices externos de los modelos triangulares. En el modelo RPST no truncado el tamaño muestral máximo no está delimitado y aunque puede ser muy grande, la probabilidad de que la trayectoria discorra entre los límites de la región de continuación y que llegue a infinito es nula. Así, ya que la  $N$  máxima teórica está abierta, le hemos asignado siempre el mayor valor observado para  $N$  máxima.

En relación al valor mínimo teórico -  $Mín(t)$  -, éste puede encontrarse en cualquier punto muy cercano al origen y puede alcanzarse muy tempranamente rechazando la

## MATERIAL Y MÉTODO

---

hipótesis nula. No obstante, para las previsiones, y de acuerdo con las características de ambos fármacos, es más útil tener en cuenta la finalización temprana sin rechazar la hipótesis de igualdad. De esta manera el valor mínimo utilizado se corresponde al punto de corte interno de los brazos que delimitan la región de continuación con el eje de abscisas - vértice interno -.

Abreviaciones de los estimadores del tamaño muestral utilizados en texto y figuras:

Tamaño muestral	Teórico	Observado
Mínimo	Mín(t)t	Mín(t)
Percentil 50	P50(t)t	P50(t)
Media	E(t)t	E(t)
Percentil 90	P90(t)t	P90(t)
Máximo	Máx(t)t	Máx(t)

La probabilidad de finalización con un tamaño cercano a N mínima es máxima cuando la hipótesis nula es cierta ( $\theta=0$ ); en el caso de la N máxima es más probable la terminación cercana en ese punto cuando las diferencias reales son intermedias entre la hipótesis nula y la alternativa ( $0 < \theta < \theta_R$ ).

Intuitivamente se entiende que para la igualdad ( $\theta=0$ ) la trayectoria de Z y V vaya más o menos paralela y cerca del eje de abscisas, de tal manera que corte, generalmente, cerca del vértice interno. Si las diferencias son intermedias a las previstas ( $\theta=1/2*\theta_R$ ), lo más probable es que la trayectoria siga un camino intermedio entre los límites de uno de los brazos y requiera haber recorrido bastante distancia desde el origen para llegar cortar por uno de los límites.

## MATERIAL Y MÉTODO

---

Los resultados de estos estimadores,  $\text{Mín}(t)_t$ ,  $\text{P50}(t)_t$ ,  $\text{E}(t)_t$ ,  $\text{P90}(t)_t$  y  $\text{Máx}(t)_t$ , se presentarán globalmente y para cada uno de los valores de  $\theta$  estudiados. Los valores de  $\theta$  estudiados van entre  $-1.5$  y  $+1.5$  veces el valor de  $\theta_R$  con incrementos de  $0.25$ . Para estudiar los estimadores antes citados globalmente, se utilizarán sus valores promedios, a excepción de  $\text{Mín}(t)_t$  y  $\text{Máx}(t)_t$  puesto que estos son fijos y no varían en función de  $\theta$ .

Recordar que para datos binomiales  $\theta$  es el logaritmo neperiano del OR:

$$\theta = \log(\text{OR}) = \log ( P_e / (1-P_e) ) - \log ( P_c / (1-P_c) )$$

En la tabla siguiente se presentan las equivalencias de los distintos valores de  $\theta$  usados para diseñar el estudio TIM con medidas de proporciones y OR.

Theta (fracciones)	Theta	OR	P(con)	P(exp)	Diferencias absolutas	Diferencias relativas
3/2 (+ThetaR)	0.473	0.623	10.00%	6.48%	-3.52%	-35.22%
5/4 (+ThetaR)	0.394	0.674	10.00%	6.97%	-3.03%	-30.29%
(+ThetaR)	0.315	0.730	10.00%	7.50%	-2.50%	-25.00%
3/4 (+ThetaR)	0.236	0.790	10.00%	8.07%	-1.94%	-19.35%
1/2 (-ThetaR)	0.158	0.854	10.00%	8.67%	-1.33%	-13.31%
1/4 (+ThetaR)	0.079	0.924	10.00%	9.31%	-0.69%	-6.87%
0 (+ThetaR)	0.000	1.000	10.00%	10.00%	0.00%	0.00%
1/4 (-ThetaR)	-0.079	1.082	10.00%	10.73%	0.73%	7.30%
1/2 (-ThetaR)	-0.158	1.171	10.00%	11.51%	1.51%	15.10%
3/4 (-ThetaR)	-0.236	1.267	10.00%	12.34%	2.34%	23.40%
(-ThetaR)	-0.315	1.370	10.00%	13.21%	3.21%	32.10%
5/4 (-ThetaR)	-0.394	1.482	10.00%	14.14%	4.14%	41.40%
3/2 (-ThetaR)	-0.473	1.604	10.00%	15.13%	5.13%	51.30%

### Comprobación del cumplimiento de las especificaciones teóricas mediante simulaciones de estudios

De acuerdo con la información revisada se descartaron los métodos que no optimizaran el tamaño muestral para supuestos de  $\theta = 0$ . También se descartaron los métodos basados en las funciones de gasto de alfa por no haber programas para realizar el diseño en el momento del diseñar el estudio TIM.

Las simulaciones para la comprobación de las especificaciones de los distintos modelos fueron realizadas para los modelos secuenciales de límites continuos. Se usó el programa PEST v3.0 para calcular las rectas que delimitan la región de continuación de los modelos triangular, RPST no truncado, RPST truncando a 4000, 4500, 5000, 5100, 5200, 5300, 5305, 5400, 5500, 6000, 7000 y 10000 pacientes.

La construcción de los límites de la región de continuación es sencilla a partir de los datos de la siguiente tabla y teniendo en cuenta los siguientes puntos:

Las pendientes de los modelos RPST son idénticas para todas las rectas, pero cambian los signos de las ordenadas en el origen.

El vértice se puede calcular despejando de cualquiera de las 4 rectas que parten del origen de  $V$  ( $V=0$ ), el valor de  $V$  para que  $Z$  sea 0.

La recta perpendicular al eje de coordenadas con valor de  $V$  máximo corresponde al valor de  $V$  equivalente al punto de truncamiento (para modelos RPST truncados).

El punto que marca en el límite de truncamiento el cambio entre las hipótesis de igualdad y alternativas, correspondería a la prolongación de una recta paralela a los límites partiendo desde el punto de origen ( $Z=0$ ,  $V=0$ ).

Para el modelo triangular la pendiente de la recta externa del brazo superior es distinta que la del la recta interna del brazo superior; dado que es un modelo simétrico es fácil delimitar el brazo inferior.

## MATERIAL Y MÉTODO

---

Para una mejor comprensión se recomienda fijarse en las figuras de los distintos métodos en introducción y en la construcción de la región de continuación del estudio TIM que se describe en resultados.

<i>Modelos</i>	<i>Origen</i>	<i>Pendiente brazo superior externo</i>	<i>Pendiente brazo superior interno</i>	<i>Valor máximo de V</i>
Triangular	±13,52620	0,11073	0,33221	122,1460
RPST 4000	±12,48450	0,22050		79,8437
RPST 4500	±9,36122	0,22174		89,8242
RPST 5000	±8,75598	0,22290		99,8046
RPST 5100	±8,68426	0,22312		101,8000
RPST 5200	±8,62182	0,22332		103,7960
RPST 5300	±8,56705	0,22353		105,7920
RPST 5305	±8,56449	0,22354		105,8920
RPST 5400	±8,51873	0,22372		107,7890
RPST 5500	±8,47585	0,22391		109,7850
RPST 6000	±8,31997	0,22476		119,7650
RPST 7000	±8,16376	0,22607		139,7260
RPST 10000	±8,04064	0,22791		199,6090
RPST	±8,01358	0,22858		----

Los puntos de truncamiento del modelo RPST truncado se han diseñado teniendo en cuenta el mínimo posible, N justo superior a N fija -4000-, y como máximo un valor alto que optimice algunos estimadores pero que fuera plausible -10000-. Entre valores de 4000 y 6000 se han diseñado cada 500 y después se escogió un punto intermedio, 7000, entre 6000 y 10000.

## MATERIAL Y MÉTODO

---

En una primera presentación del protocolo a los Comités Éticos se valoró realizar el estudio secuencial con un poder del 90%, situándose el punto de truncamiento mínimo en unos 5305 pacientes. Por este motivo, aunque posteriormente por previsiones de problemas en las tasas de reclutamientos se rebajó el poder al 80%, se decidió mantener, una vez estudiados las opciones, un modelo truncando a 5305 pacientes, cambiando de esta manera lo menos posible el diseño, el cual, por otra parte, no era demasiado conocido por parte de las entidades evaluadoras.

Aunque se usaron todos los modelos citados para las simulaciones, con el fin de simplificar la exposición nos referiremos a partir de ahora solamente a los modelos triangular, RPST, y RPST truncando a 4500, 5305 y 7000. Los modelos RPST con puntos de truncamiento que no se desarrollaran en esta tesis, presentan resultados intermedios a los modelos que si se presentaran y no aportan información distinta o relevante adicional. En los listados anexos están los datos de todos los modelos simulados.

Se recogieron también datos correspondientes a valores teóricos de tamaño muestral y de la función de CO para la confrontación posterior con las simulaciones.

Se simularon 10000 estudios para cada uno de los modelos y distintos valores de theta. Con esas simulaciones se consigue una precisión de  $\pm 0.98\%$  en el punto de varianza máxima ( $p = 0.5$ ). Para estimar alrededor de  $p=0,05$  la precisión es de  $\pm 0,43\%$  y para  $p=0,20$  se sitúa en un  $\pm 0,78\%$ . Las simulaciones programaron en Fortran 90 de NAG <sup>181</sup>.

Para calcular valores mínimos, máximos, medias y percentiles de tamaño muestral y ahorro de los estudios simulados se utilizó el paquete estadístico SAS <sup>182</sup>. Los listados con los resultados teóricos y simulados se han realizados usando SPSSWIN v6.0.1 <sup>183</sup>\$. Ver anexo para los listados de programas de simulaciones y análisis. Para las figuras de esta tesis se ha utilizado Excel 97.

### *Tamaño muestral*

Con el fin de poder realizar comparaciones entre valores teóricos y los observados en las simulaciones, para éstas se han utilizado los mismos índices:

Valor mínimo observado - Mín(t) -

Valor máximo observado - Máx(t) -

Percentiles 50 y 90, respectivamente P50(t) y P90(t)

Valor promedio E(t)

Al igual que con los cálculos teóricos, la N mínima y N máxima simuladas no son buenos estimadores para las previsiones. Estas definiciones no se corresponden con la N mínima y la N máxima de las simulaciones por los siguientes motivos:

Para las simulaciones no debemos perder de vista que en caso de observar un valor de theta muy distinto al esperado, la trayectoria podría tener mucha pendiente y cortar la región de continuación muy cerca del origen, incluso antes de necesitar valores de V cercanos al valor del vértice interno; este hecho puede suceder siendo cierta la hipótesis de igualdad -  $\Theta=0$  - con una probabilidad de 0.05 –error de tipo I -.

El valor de N máxima del diseño es una aproximación puesto que en realidad los puntos de truncamiento, o vértices externos para el modelo triangular, se deben corresponder a valores de V y no de N. Como V, aparte de n, depende de la proporción de aparición global de los sucesos, al aumentar esta proporción, V requiere mayor tamaño muestral para poder mantener constante su valor. De esta manera se consigue mantener el poder aunque las especificaciones de proporciones del diseño difieran de las observadas.

Para finalizar,  $N$  mínima y máxima corresponden a los valores mínimos y máximos tras 10000 simulaciones para cada tipo de modelo y difícilmente se alcanzarán esos límites “reales” mediante simulaciones.

Al igual que con los valores teóricos, para los valores observados los estimadores  $\text{Mín}(t)$ ,  $P50(t)$ ,  $E(t)$ ,  $P90(t)$  y  $\text{Máx}(t)$  se presentarán globalmente y para cada uno de los valores de  $\theta$  estudiados. En este caso sí que se utilizará sus valores promedios para la evaluación global de  $\text{Mín}(t)$  y  $\text{Máx}(t)$  puesto que no son fijos y varían en función de  $\theta$ .

### *Ahorro en el número de pacientes*

El ahorro de los métodos secuenciales se ha cuantificado con las simulaciones desde dos puntos de vista. En primer lugar se ha estudiado el porcentaje de reducción en el número de pacientes, y posteriormente que porcentaje de estudios necesitan menor tamaño muestral que un diseño convencional. Estos cálculos se han realizado para los distintos modelos secuenciales y para distintos valores de theta.

Dado que, al igual que sucede con las estimaciones directas de tamaño muestral, los valores observados varían en función de la theta, se han calculado los valores promedio para cada estimador ( $P50(t)$ ,  $P90(t)$  y  $E(t)$ ). Recordamos que dicho promedio está calculado a partir de todos los valores simulados de entre  $\theta = -1,5 * \theta_R$  a  $\theta = +1,5 * \theta_R$ , en incrementos de 0,25. Además, se han resumido los valores calculando el promedio en el porcentaje de estudios que ahorran  $N$ .

Por otro lado, la cuantificación del ahorro absoluto en número de pacientes puede inferirse directamente a partir de los valores teóricos o simulados teniendo en cuenta que la  $N$  requerida para un estudio convencional es aproximadamente de 4000 pacientes. De todas maneras, estos resultados se hallan en el anexo S.

### *Valoración de las especificaciones estadísticas*

Se ha comprobado el comportamiento de los distintos modelos respecto a la CO tanto en las previsiones teóricas como con las comprobaciones de los valores observados mediante simulaciones. La característica operante se expresa como la probabilidad de aceptar la hipótesis de igualdad en función de los valores de estimación de diferencias entre ambos grupos  $\theta$ .

$$P(\text{aceptar } H_0 \mid \theta)$$

Cuando  $\theta$  tiene el valor de igualdad ( $\theta=0$ ), solamente se debe rechazar la hipótesis nula al nivel de protección de error primera especie seleccionado ( $\alpha$ ), habitualmente el 5%. En cambio cuando las diferencias reales son las buscadas en la hipótesis de trabajo, ( $\theta = \theta_R$ ), el porcentaje de estudios en los que se pueda rechazar la hipótesis de igualdad no debe ser inferior al poder predefinido en el estudio  $100 \cdot (1-\beta)$ . Desviaciones de estos valores indican un mal comportamiento y desaconsejarían su utilización. Información más amplia está explicada en la introducción.

### *Estudio TIM*

#### **Diseño**

Se seleccionó el modelo RPST truncando a 5305 y se procedió a implementarlo en el estudio TIM. Las características estadísticas seleccionadas fueron  $\alpha=0.05$  bilateral y  $\beta=0.2$ . La diferencia clínicamente relevante se estableció en un 2.5% (porcentaje absoluto) para un valor de base en el tratamiento experimental del 10%. Los detalles generales del método secuencial de límites continuos están descritos en introducción y los específicos del estudio en resultados.

#### **Estructuras y comités. Flujo de los datos y controles de calidad**

De acuerdo con la metodología secuencial del estudio, se requería disponer los datos validados en el momento de efectuarse una inspección, sobretodo la información que pudiera hacer variar la inclusión de los pacientes y la validación de los end-points. Para tal fin se estableció en el protocolo la formación de un Comité de Supervisión constituido por un cardiólogo y un neurólogo independientes a las tareas de reclutamiento del estudio.

La monitorización del estudio fue llevada a cabo por una empresa de servicios con reconocido prestigio internacional –actualmente Phenix Verum ITEM-Pharma-. Los monitores del estudio se encargaron de comprobar la calidad y fiabilidad de los datos cumplimentados por los investigadores. En caso de detectarse errores, incongruencias o datos no cumplimentados, se cursaron aclaraciones a los investigadores y no se remitieron al Comité de Supervisión hasta la completa resolución.

Una vez los cuadernos de los pacientes llegaban al Comité de Supervisión eran recomprobados para asegurar que los datos que se utilizarían en la inspección secuencial fueran correctos.

Posteriormente, los cuadernos de recogida de datos llegaban a la Unidad de Estadística, se procedía a la entrada de los datos en dos bases de datos paralelas y por dos técnicos distintos. Para evitar posibles errores en este proceso, se confrontaban ambas entradas con listados y se procedía a la corrección de los datos hasta conseguir que éstos fueran idénticos en las dos bases.

Una nueva medida de calidad consistió en la elaboración de controles de búsqueda de errores e inconsistencias de los datos. Una vez identificadas, se procesaban formularios de aclaraciones que eran enviados a la empresa de monitorización para su corrección por parte del investigador responsable. Finalmente se introducían los datos corregidos de nuevo y por duplicado con posterior confrontación.

Adicionalmente se realizó otro control de los datos una vez finalizado el reclutamiento de estudio. Durante tres semanas todos los cuadernos fueron revisados por los monitores del estudio en búsqueda de errores, y se cursaron las últimas aclaraciones.

### **Monitorización secuencial. Criterios de finalización del estudio**

Para terminar el estudio cruzando la región de continuación cerca del vértice interno se requeriría información correspondiente a valores de  $V > 38$ , lo que implica valores de tamaño muestral de alrededor de 1900 pacientes. Así, se decidió realizar la primera inspección secuencial a los 1000 pacientes reclutados, válidos para el análisis de eficacia.

## MATERIAL Y MÉTODO

---

Además, dado que diferencias grandes a favor de un tratamiento u otro eran muy improbables, con el tamaño muestral mínimo prefijado nos asegurábamos que en el hipotético caso de terminación prematura rechazando la hipótesis nula, se tendría una precisión aceptable para la estimación del tamaño del efecto comparado.

Las inspecciones sucesivas se acordó realizarlas cada 250 nuevas inclusiones validadas por el comité de supervisión.

Sólo los casos considerados válidos por el Comité de Supervisión fueron usados para la monitorización y, consecuentemente, para el análisis secuencial. En caso de que se detectaran casos no válidos después de su inclusión en la monitorización secuencial, se excluyeron del siguiente corte.

Una vez que la trayectoria secuencial sobrepasara los límites de la región de continuación, se daría por finalizada la inclusión de nuevos casos. Los casos ya incluidos tras la finalización serían incluidos, excepto aquellos casos que no cumplieran con los requisitos de la supervisión.

Al declarar la finalización del estudio con todos los casos monitorizados, la realización de la reevaluación de validez introdujo algunos cambios. Para esta nueva evaluación de la trayectoria secuencial consiguiente, se consideraron en cada inspección los casos válidos incluidos en cada corte. Finalmente los resultados se contrastaron con los obtenidos en primera instancia tanto en la monitorización como en el análisis de la variable principal.

### **Análisis estadístico**

Dada la planificación de revisar nuevamente los cuadernos de recogida de datos, no se realizó ningún análisis hasta después de cerrar el periodo de revisión. Se hizo de esta manera para evitar que los resultados pudieran influir de manera indirecta en las

## MATERIAL Y MÉTODO

---

decisiones. De todas maneras, la posibilidad era remota por que los monitores eran ciegos respecto a la asignación de A o B. Además los sobres con la correspondencia código a tratamiento se mantuvieron cerrados hasta después del análisis final. Una vez finalizada la revisión, fueron analizados para eficacia los datos de los pacientes validados de forma cruda y ajustando por el método secuencial. Se realizó un análisis adicional de la variable principal con los datos incluidos en el momento de la última inspección para confrontarlos con los definitivos.

Dado que las variables secundarias están muy correlacionadas con la variable principal, se realizó un ajuste para evitar sesgos. Toda la metodología de análisis está explicada en la introducción y desarrollada en Whitehead J. 1992 {2316}

### ***BÚSQUEDA DEL TAMAÑO MUESTRAL ÓPTIMO***

#### ***DISEÑOS TEÓRICOS***

##### **DISEÑO CLÁSICO**

El diseño convencional está planteado en el apartado anterior siendo el tamaño muestral necesario es de alrededor de 4000 pacientes. En relación con este diseño hay que recordar que para poder llegar a alguna conclusión, hay que haber reclutado totalmente el tamaño muestral previsto para las características del diseño. Por este motivo los valores de  $P_{50}(t)$ ,  $P_{90}(t)$ ,  $E(t)$ ,  $Mín(t)$  y  $Máx(t)$  deben tener siempre el mismo valor, el de  $N$  fija.

##### **DISEÑOS SECUENCIALES ÓPTIMOS**

Se estudia la posible aplicación de un diseño secuencial con el objetivo primordial de optimizar el tamaño muestral.

A continuación se analizan los distintos métodos secuenciales y sus posibles ventajas y desventajas para su aplicación al estudio TIM:

- métodos secuenciales por grupos
- intervalos de confianza repetidos
- restricción estocástica
- funciones de gasto de alfa
- métodos basadas en la función de verosimilitud

### *FUNCIONES DE GASTO DE ALFA*

Para un estudio como el TIM interesa testar las hipótesis a dos colas y con la posibilidad de finalización temprana para cualquiera de las dos hipótesis, nula y alternativa. A continuación se muestran los tamaños muestrales en función de 5 modelos distintos de funciones para terminación temprana bajo cualquiera de las dos hipótesis, del valor del parámetro de forma, y de las especificaciones del estudio TIM.

Para fijar el tamaño de  $N$  para cada inspección debemos dividir la  $N$  máxima teórica por el número previsto de éstas. Debemos recordar que el programa EaSt sólo permite hasta 10 inspecciones, por lo que hemos descartado menos de 8 cortes para los cálculos teóricos puesto que saldrían paquetes de demasiado volumen en tamaño muestral. En nuestro caso nos salen paquetes de entre 480 y 772 pacientes.

Los valores de  $N$  máxima han oscilado entre 4804 y 7720,  $E(t)t$  bajo  $H_0$  ( $\theta=0$ ) entre 2630 y 2849 y  $E(t)t$  bajo  $H_1$  ( $\theta=\theta_R$ ) entre 2858 y 3142. Los valores de  $E(t)t$  han sido inferiores a la  $N$  fija pero no la  $N$  máxima, siempre superior. Para valores intermedios de  $\theta$  ( $\theta=\theta_R/2$ ), las previsiones muestrales son bastante similares a los teóricos bajo la hipótesis alternativa pero algo superiores.

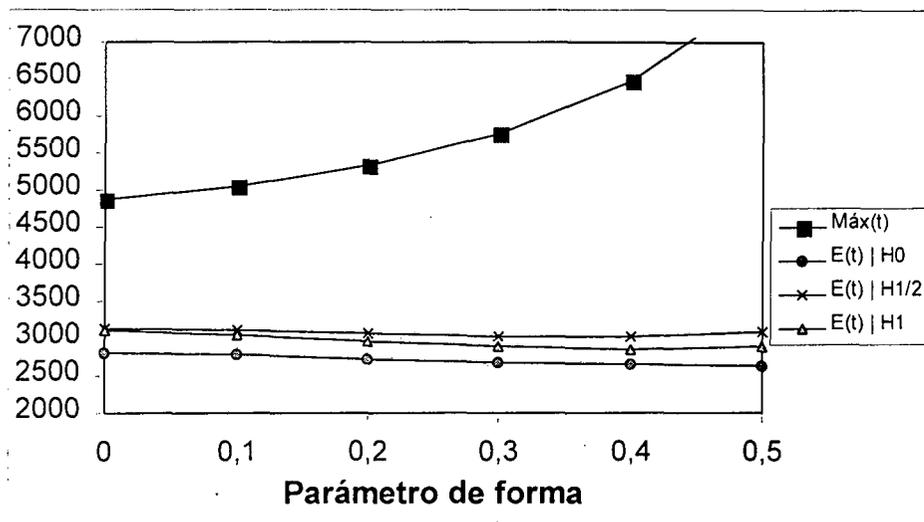
El parámetro de forma con valor 0 optimiza los valores esperados de  $N$  máxima, pero empeora los valores de  $N$  esperada bajo la hipótesis nula y bajo la hipótesis alternativa. Estas estimaciones van cambiando de manera que cuantos más nos acercamos al parámetro de forma 0.5 más se optimizan la  $N$  esperadas bajo  $H_0$  y  $H_1$  y empeora la  $N$  máxima. El número de inspecciones mejora muy discretamente los valores estimados para  $E(t)t$  bajo  $H_0$  y  $H_1$  y empeora la  $N$  máxima.

## RESULTADOS

Parámetro de forma	Número de inspecciones	N máxima	N esperada bajo $H_0$ ( $\theta=0$ )	N esperada bajo $H_{1/2}$ ( $\theta=\theta_R / 2$ )	N esperada bajo $H_1$ ( $\theta=\theta_R$ )
0	8	4804	2849	3176	3142
	9	4824	2839	3163	3123
	10	4862	2810	3140	3105
0,1	8	4987	2793	3127	3070
	9	5022	2785	3116	3052
	10	5043	2773	3103	3035
0,2	8	5231	2759	3094	3000
	9	5295	2799	3069	2977
	10	5335	2719	3057	2961
0,3	8	5651	2721	3068	2935
	9	5680	2704	3048	2914
	10	5756	2674	3024	2894
0,4	8	6325	2679	3059	2896
	9	6421	2662	3043	2875
	10	6482	2650	3030	2858
0,5	8	7306	2688	3116	2921
	9	7546	2653	3093	2900
	10	7720	2630	3080	2884

## RESULTADOS

---



### ***MÉTODOS BASADOS EN LÍMITES CONTINUOS***

Estos métodos permiten una amplia variedad de diseños secuenciales. De entre los posibles diseños tenemos:

- Triangular
- Razón de probabilidades secuencial RPST
- RPST truncado a distintos puntos
- Restricted Procedure (RP)
- Open Top Design (OTD)

Los diseños RP y OTD no son adecuados dado que la posibilidad de aceptación final de la  $H_0$  es alta, y por la forma de la región no se termina antes que un diseño clásico.

Así pues, haremos una valoración con más profundidad solamente con los tres primeros modelos citados: triangular, RPST y RPST truncando a 4500, 5305, y 7000 pacientes.

### OTROS DISEÑOS

#### *RESTRICCIÓN ESTOCÁSTICA*

Este método fue diseñado para optimizar decisiones de aceptación o rechazo de una hipótesis aplicado a controles de calidad, no aportando la posibilidad de estimación puntual ni de intervalos de confianza. Por este motivo su aplicabilidad en el campo del ensayo clínico es muy limitada y no se ajusta a las necesidades de nuestro estudio.

#### *INTERVALOS DE CONFIANZA REPETIDOS*

Esta técnica en el fondo esta basada en el método secuencial por grupos y conlleva el mismo problema por el cual no es un diseño óptimo para este ensayo clínico: si los intervalos de confianza que se han ido calculando en las distintas inspecciones no excluyen el 0, es decir no se puede rechazar la  $H_0$ , se requerirá reclutar el total de la muestra.

#### *MÉTODOS SECUENCIALES POR GRUPOS*

En estos métodos se van realizando inspecciones intermedias hasta que el valor de significación es menor que el valor de  $\alpha'$  calculado para una inspección dada. En ese momento se finaliza el estudio y se concluye superioridad de uno de los grupos. En caso de no poder concluir superioridad en ninguno de los cortes intermedios, será necesario reclutar toda la muestra. Dado que para calcular el tamaño muestral se utiliza un alfa igual o más estricto que el diseño convencional, si no es posible finalizar el estudio hasta el último corte el tamaño muestral requerido ha sido igual o superior a un diseño convencional.

Por consiguiente no son métodos aconsejables a menos que se esperen diferencias superiores en magnitud a las del diseño establecido, ya que nunca reducirán el tamaño muestral final si no se rechaza la hipótesis nula en alguna inspección previa a la última.

### ***EQUIVALENCIA TERAPÉUTICA***

Para el diseño de equivalencia terapéutica hay diversas fórmulas para su cálculo y en función de éstas el tamaño puede variar bastante. Así, si escogemos la fórmula que se propone en <sup>19</sup> equivalente a M&C {}, con las especificaciones:

Protección frente al error de primera especie  $-\alpha-$  de 0.05 (bilateral)

Poder 0.8

Proporciones iguales de eventos en ambos tratamientos situadas en 0.10

Rango de aceptación de equivalencia 0.025

el tamaño muestral asciende a 4522, lo que implica unos requerimientos de alrededor del 13% superiores al tamaño muestral clásico.

Si aplicamos una fórmula aproximada para analizarlo de acuerdo al diseño doblemente unilateral de Schuirmann o por intervalos de confianza <sup>62</sup>, harían falta 4934 pacientes (23% más que en la aproximación clásica).

El problema sigue siendo el mismo que en un diseño clásico, ya que se requiere siempre todo el tamaño muestral para poder concluir equivalencia o no equivalencia, teniendo la desventaja añadida de que se requiere mayor tamaño que en un estudio convencional.

### ***EQUIVALENCIA TERAPÉUTICA MEDIANTE METODOLOGÍA SECUENCIAL***

En el momento del diseño del estudio todavía no se había publicado el diseño para realizar un estudio con este método. Whitehead propone un diseño triangular bilateral con alfa convencional, y beta de 0.05, o incluso más restrictivo de 0.025. En este caso aunque no se realizó esta aproximación por no concurrencia en el tiempo, además se penaliza el tamaño muestral respecto al diseño secuencial convencional.

### ***MÉTODOS BAYESIANOS***

Aunque hoy en día la aplicación de las técnicas bayesianas está en auge, el problema principal es la falta de programas para la aplicación práctica de estos modelos y que las propuestas secuenciales con posible aplicación se publicaron 2 años después a la fecha del diseño del estudio TIM<sup>175,176</sup>

## RESULTADOS

---

### *ESTIMACIONES TEÓRICAS DE TAMAÑO MUESTRAL*

Los resultados teóricos se presentan inicialmente globales y a continuación se describen en función de distintos valores de theta.

### GLOBALES

#### *N mínima –Min(t)- y N máxima –Máx(t)-*

La N mínima más favorable corresponde al modelo RPST seguido de los modelos RPST truncando a valores lejanos, Triangular y RPST truncando cerca a N fija. Las diferencias entre los valores extremos, RPST y RPST 4500, son de 359 pacientes. En todos los casos los valores son inferiores a la N fija.

La N máxima es más desfavorable, por orden, para los modelos RPST, RPST truncando lejos de N fija y Triangular. Las diferencias entre los valores extremos son de varios miles y superiores en todos los casos a la N fija.

Modelo	N mínima
RPST	1757
RPST 7000	1810
RPST 5305	1920
TRIANGULAR	2041
RPST 4500	2116

Modelo	N máxima
RPST 4500	4500
RPST 5305	5305
TRIANGULAR	6119
RPST 7000	7000
RPST	---

## RESULTADOS

---

### *P50(t)t, P90(t)t y E(t)t*

Globalmente, como promedio de los valores para todas las thetas, el P50(t)t mejor es para el modelo RPST, seguido de los modelos RPST truncando de lejos hasta 5305, Triangular y RPST truncando cerca. Las diferencias entre los valores extremos son de alrededor de 600 pacientes.

Para el P90(t)t, los valores apuntan un comportamiento mejor para el modelo triangular (3404), y los RPST con punto de truncamiento cercano a N fija que presentan valores inferiores a 4000. Los modelos más perjudicados, de menos a más, son RPST, y RPST7000 con valores entre 4341 y 4626.

El modelo RPST con 2234 inclusiones es el diseño más favorable para E(t)t, seguido de RPST7000 con 2270, y Triangular con 2272. De todas maneras, las diferencias en tamaño con el más favorable son pequeñas y únicamente son mayores del centenar con el RPST4500 (151).

Modelos	Promedio P50(t)t
RPST	1814.9
RPST7000	1895.3
RPST5305	2071.9
TRIANGULAR	2145.1
RPST4500	2404.4

Modelos	Promedio P90(t)t
TRIANGULAR	3404.1
RPST4500	3872.2
RPST5305	4275.7
RPST	4340.6
RPST7000	4626.2

Modelos	Promedio E(t)t
RPST	2233.7
RPST7000	2270.8
TRIANGULAR	2272.9
RPST5305	2287.9
RPST4500	2384.9

### EN FUNCIÓN DE THETA

Para facilitar la explicación, aunque se presentan las figuras con las distintas estimaciones comparadas de los tamaños de muestra para valores de entre  $-1.5*Tr$  y  $+1.5*Tr$ , nos centraremos en los valores de  $\theta=0$ ,  $0.5*Tr$  y  $Tr$ . Los valores concretos de todas las opciones pueden verse en los apéndices.

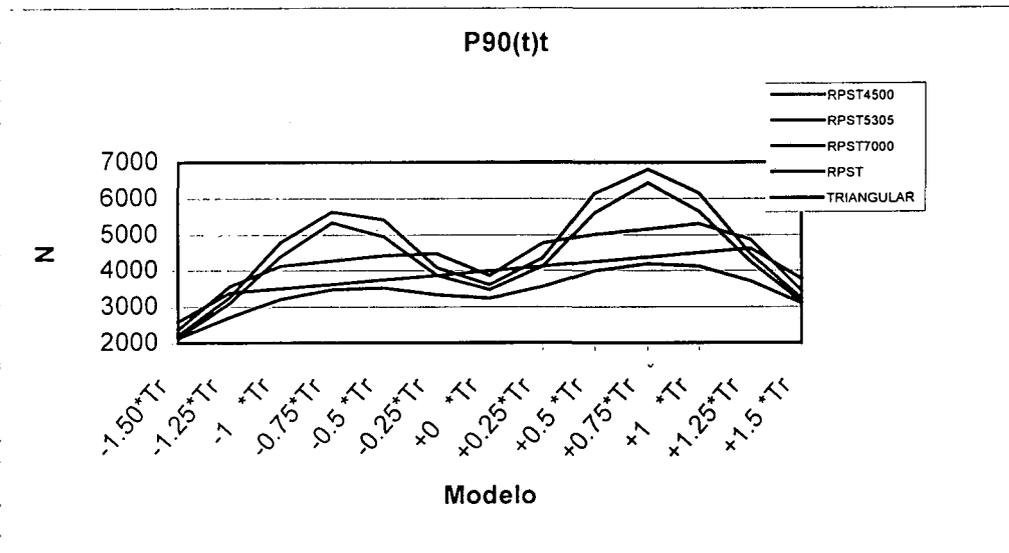
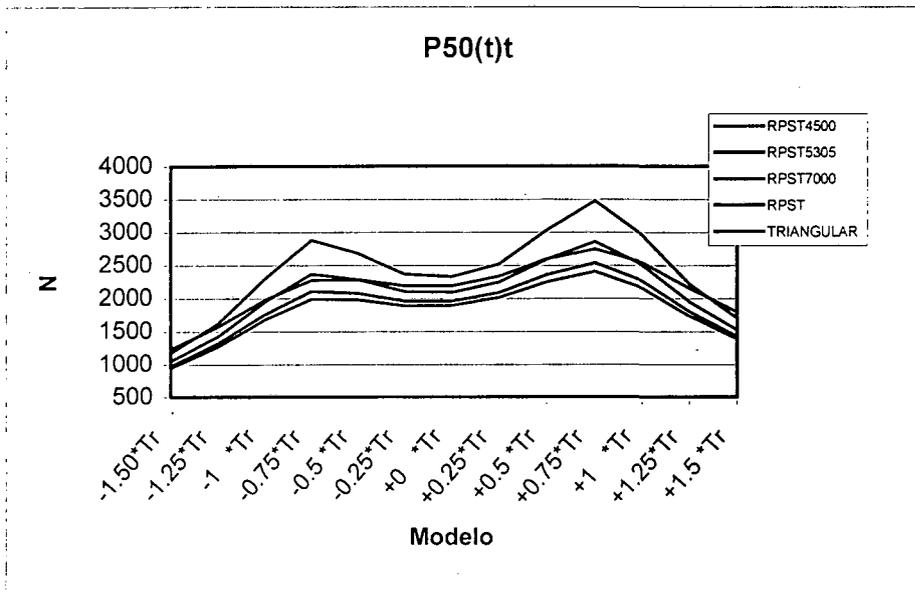
Los valores negativos de theta, como se puede observar en las figuras, son simétricos respecto los valores positivos de ésta, pero las estimaciones son siempre algo más favorables. Este hecho es debido a que los modelos secuenciales respetan el poder del diseño, tanto en un sentido como en otro, y la variabilidad es menor en el lado de valores negativos de theta.

Como resumen, las estimaciones más favorables son para valores de  $\theta=0$ , ya que de esta manera la probabilidad de que la trayectoria transcurra cercana al eje de coordenadas y corte los límites cerca del vértice interno es alta en estos modelos. También se podría dar el caso que Theta fuera grande y de esta manera la trayectoria tendría una gran pendiente y podría encontrarse con los límites externos pero parece muy poco probable a priori.

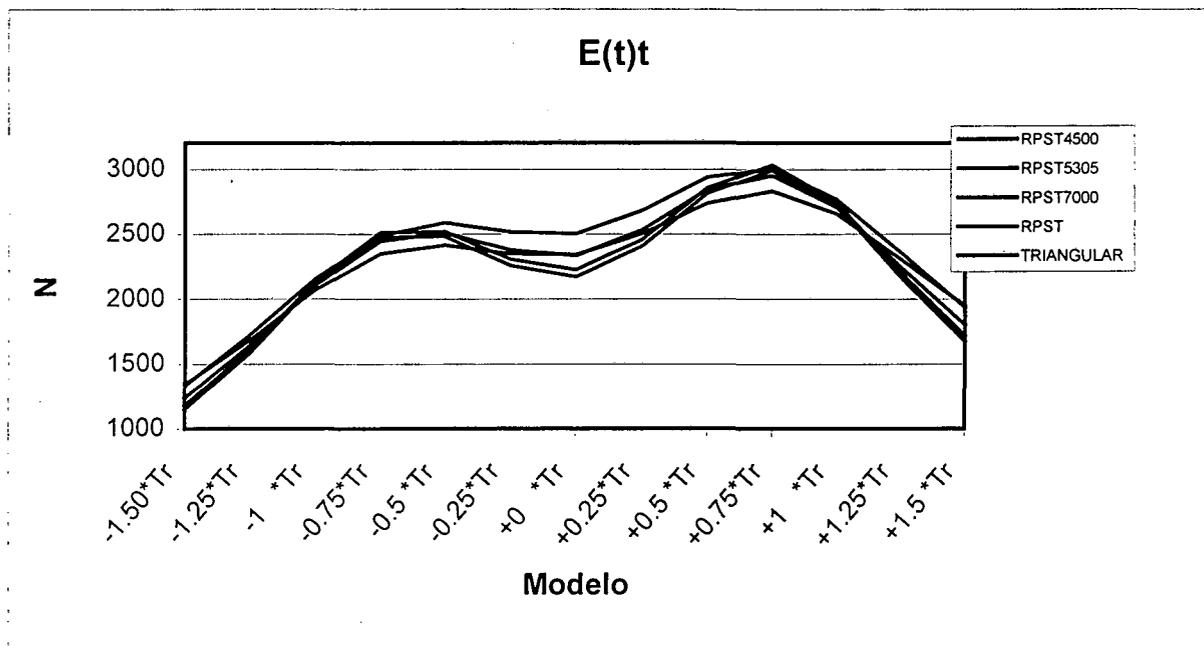
Los valores Theta intermedios entre 0 y theta R, empeoran las previsiones de los estimadores ya que es probable que la trayectoria transcurra por el centro de uno de los dos brazos sin tocar los límites.

Para valores de  $\theta = \theta_R$ , mejora el  $P50(t)$  y  $E(t)$  en relación con Thetas intermedias, pero empeora  $P90(t)$ .

RESULTADOS



# RESULTADOS



## RESULTADOS

---

En las tablas siguientes se muestran las estimaciones para los valores teóricos de  $\theta=0$ ,  $\theta=\theta_R$ , diferencia relevante a buscar, y el valor medio entre  $\theta=0$  y  $\theta=\theta_R$  ( $\theta=\theta_R/2$ ).

Se comentan a continuación las predicciones teóricas sobre las hipótesis más probables de theta, entre 0 y  $\theta_R$ , para los valores positivos. Los valores negativos son similares y casi simétricos pero con valores más favorables puesto que la varianza es menor (ver tablas completas en apéndice y figuras).

Los modelos más favorables para  $P50(t)t$  corresponden, por orden, a los modelos RPST, RPST truncando a valores lejanos del tamaño de N fija, el triangular y RPST truncado a N bajas. Esta consideración se confirma en todo el abanico de posibilidades de  $\theta$ . Las diferencias entre RPST y RPST4500 varían entre 431 para  $\theta=0$  y 799 para  $\theta = \theta_R$ .

Theta	Estimador	Modelo	Estimación
+0 *T	P50(t)t	RPST	1894
+0 *T	P50(t)t	RPST7000	1958
+0 *T	P50(t)t	RPST5305	2091
+0 *T	P50(t)t	TRIANGULAR	2192
+0 *T	P50(t)t	RPST4500	2325
+0.5 *T	P50(t)t	RPST	2248
+0.5 *T	P50(t)t	RPST7000	2356
+0.5 *T	P50(t)t	RPST5305	2591
+0.5 *T	P50(t)t	TRIANGULAR	2593
+0.5 *T	P50(t)t	RPST4500	3037
+1 *T	P50(t)t	RPST	2157
+1 *T	P50(t)t	RPST7000	2264
+1 *T	P50(t)t	RPST5305	2503
+1 *T	P50(t)t	TRIANGULAR	2545
+1 *T	P50(t)t	RPST4500	2956

## RESULTADOS

---

La estimación de  $P_{90}(t)_t$  varia en función de la Theta:

Para  $\Theta = 0$  los modelos más favorables son, por orden, triangular, RPST y RPST truncando lejos a cerca; las diferencias entre el mejor y peor modelo superan los 700 pacientes.

Para el resto de las opciones,  $\theta = 0.5 * T_r$  y  $\theta = T_r$ , se mantiene el orden, Triangular, RPST truncando cercano a N fija, RPST, y RPST truncando lejos; las diferencias entre los valores extremos superan las 2000 inclusiones.

Theta	Estimador	modelo	Estimación
+0 *T	$P_{90}(t)_t$	TRIANGULAR	3229
+0 *T	$P_{90}(t)_t$	RPST	3471
+0 *T	$P_{90}(t)_t$	RPST7000	3610
+0 *T	$P_{90}(t)_t$	RPST5305	3859
+0 *T	$P_{90}(t)_t$	RPST4500	3992
+0.5 *T	$P_{90}(t)_t$	TRIANGULAR	3988
+0.5 *T	$P_{90}(t)_t$	RPST4500	4245
+0.5 *T	$P_{90}(t)_t$	RPST5305	5005
+0.5 *T	$P_{90}(t)_t$	RPST	5599
+0.5 *T	$P_{90}(t)_t$	RPST7000	6144
+1 *T	$P_{90}(t)_t$	TRIANGULAR	4116
+1 *T	$P_{90}(t)_t$	RPST4500	4500
+1 *T	$P_{90}(t)_t$	RPST5305	5305
+1 *T	$P_{90}(t)_t$	RPST	5652
+1 *T	$P_{90}(t)_t$	RPST7000	6148

## RESULTADOS

---

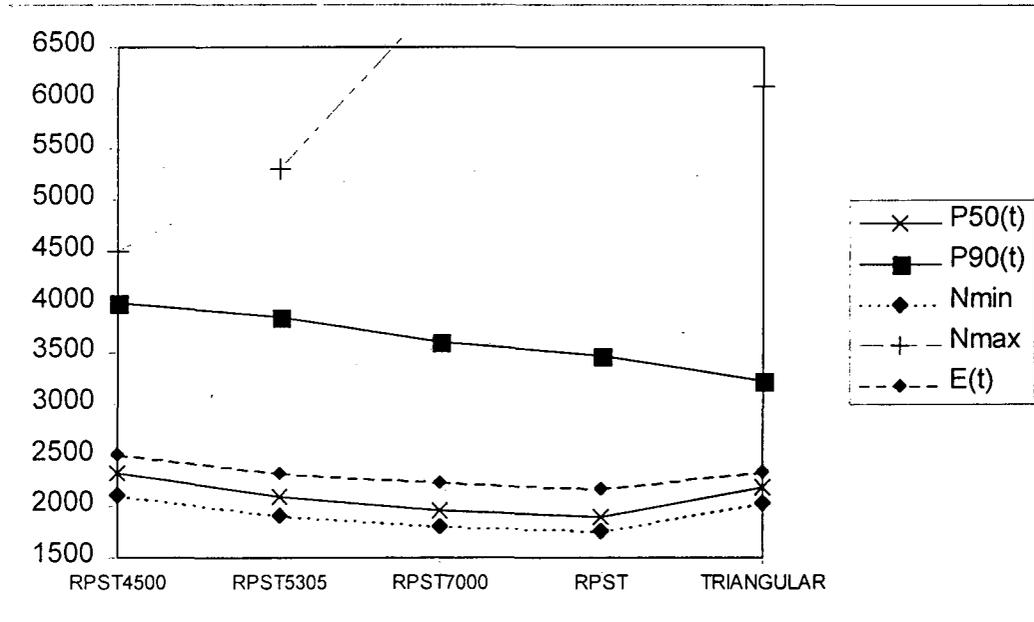
Los valores teóricos indican que los mejores modelos en función de theta son, para  $\theta = 0$ , RPST, RPST truncando lejos hasta 5305, Triangular, RPST truncando cerca; para diferencias intermedias, Triangular, RPST, RPST truncando a valores intermedios, RPST truncando a valores altos, RPST truncando a valores bajos; y para  $\theta = \theta_R$ , Triangular, RPST truncando a valores intermedios, RPST, y RPST al resto de valores.

Para la estimación de  $E(t)$  el modelo triangular es el que presenta mejores propiedades para theta distintas a 0. Las diferencias en las distintas estimaciones entre modelos son en general no muy grandes y menores a 100 pacientes.

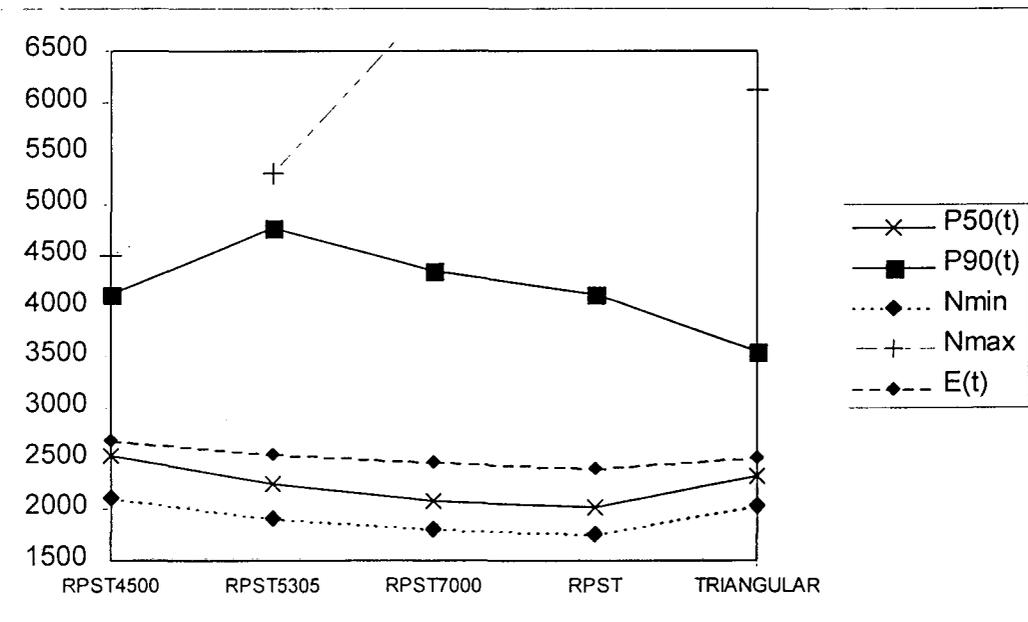
Theta	Estimador	modelo	Estimación
+0 *T	E(t)	RPST	2173
+0 *T	E(t)	RPST7000	2228
+0 *T	E(t)	RPST5305	2332
+0 *T	E(t)	TRIANGULAR	2342
+0 *T	E(t)	RPST4500	2502
+0.5 *T	E(t)	TRIANGULAR	2740
+0.5 *T	E(t)	RPST	2815
+0.5 *T	E(t)	RPST5305	2846
+0.5 *T	E(t)	RPST7000	2859
+0.5 *T	E(t)	RPST4500	2938
+1 *T	E(t)	TRIANGULAR	2659
+1 *T	E(t)	RPST5305	2705
+1 *T	E(t)	RPST	2724
+1 *T	E(t)	RPST7000	2744
+1 *T	E(t)	RPST4500	2770

# RESULTADOS

Theta=0

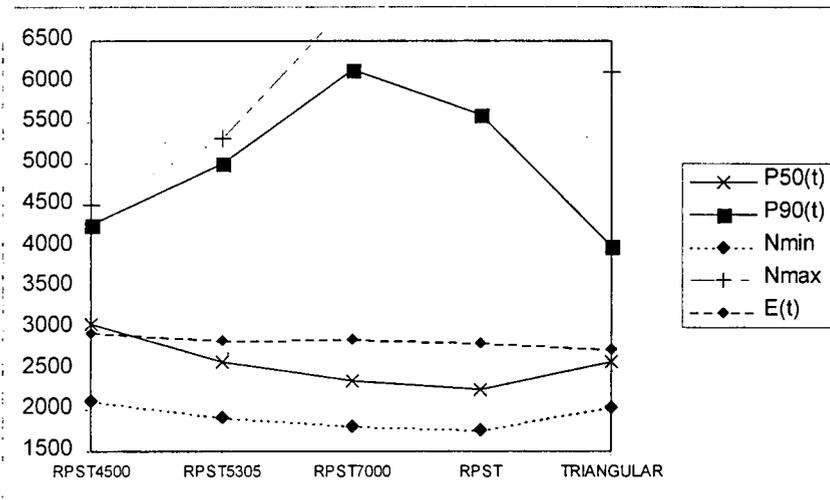


Theta = 0.25

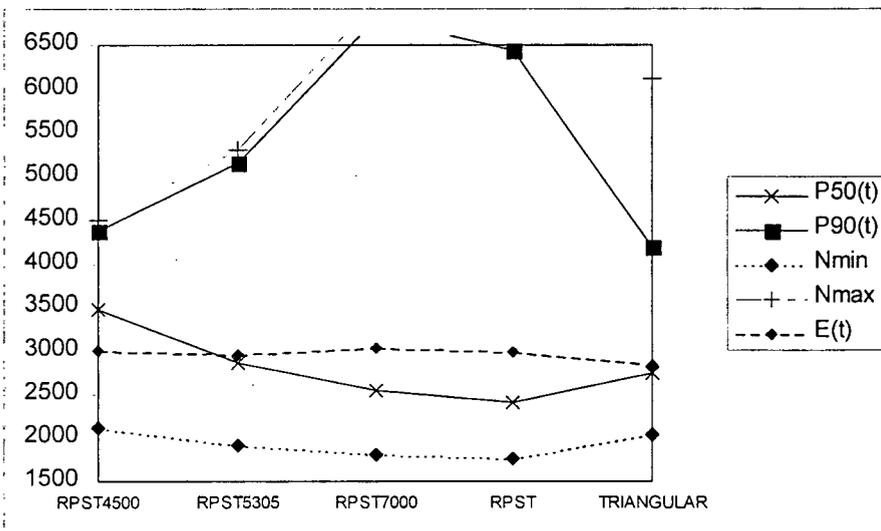


# RESULTADOS

Theta = 0.5

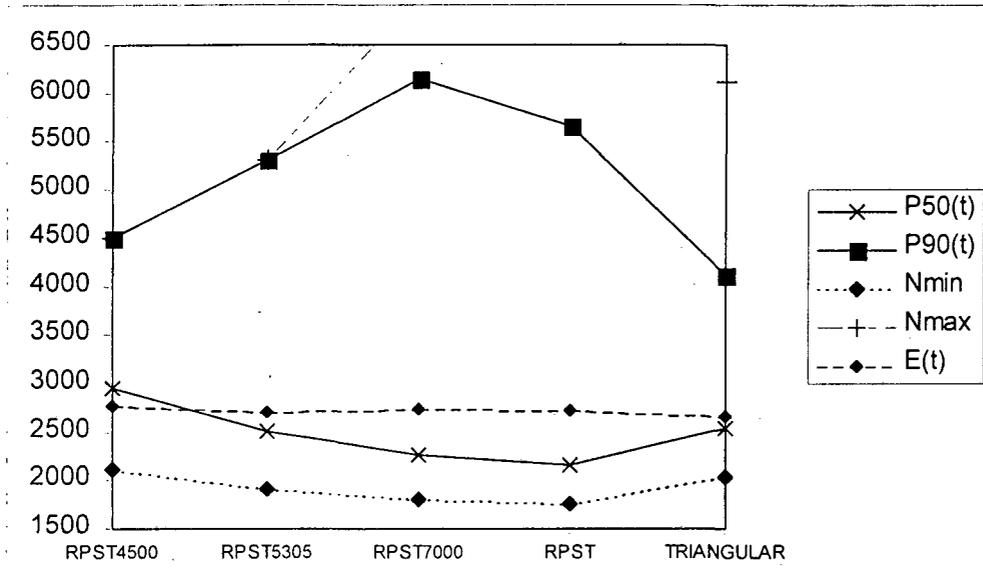


Theta = 0.75

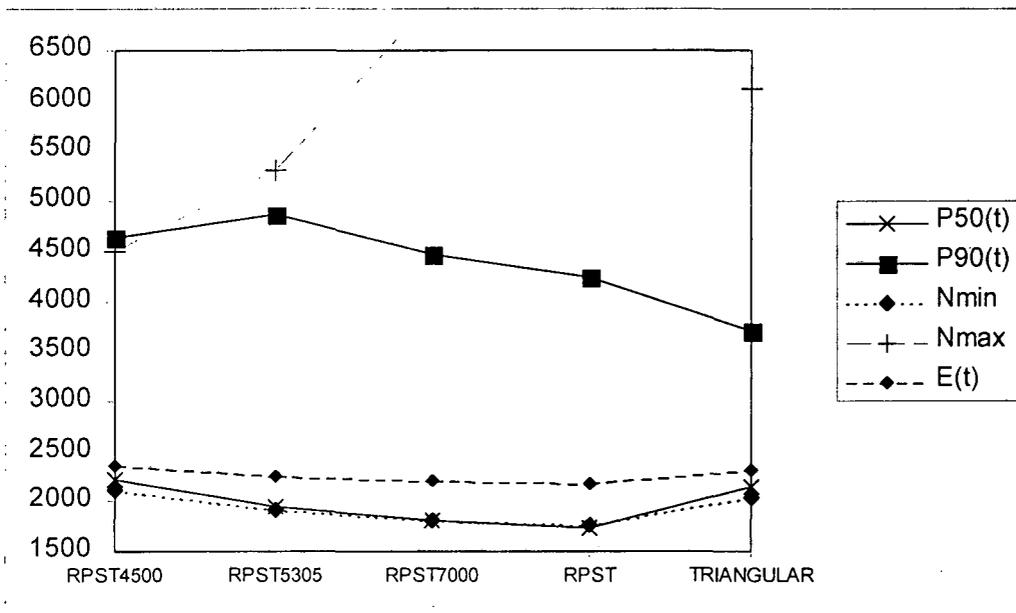


# RESULTADOS

Theta = Theta R



Theta = 1.25



## DESCRIPCIÓN DE LAS SIMULACIONES

### GLOBALES

Dado que los valores varían mucho en función de la theta, al igual que se hizo con los valores teóricos, se presentan primeramente los resultados promedio de todas las thetas de cada modelo y posteriormente para cada valor de theta estudiada.

#### *N mínima y N máxima*

Para Mín(t), los modelos más favorables son por orden el modelo RPST, seguido de los RPST con punto de truncamiento de mayor a menor tamaño de N y el Triangular. Los valores mínimos han oscilado entre 181 y 446 pacientes.

En el caso de Máx(t), la relación está invertida respecto a Mín(t), pero el modelo triangular se sitúa en una zona intermedia entre RPST 4500 y 5305, y las diferencias entre modelos son muy superiores. Éstas son de alrededor de 400 entre el modelo más favorable, RPST 4000, y el Triangular, aumentando a 700 con RPST 5305 y hasta 12000 con el modelo RPST.

Modelos Mín(t)	Promedio
RPST	180.8
RPST7000	186.9
RPST5305	201.8
RPST4500	239.8
TRIANGULAR	421.4

Modelos Máx(t)	Promedio
RPST4500	4642.5
TRIANGULAR	5056.3
RPST5305	5338.5
RPST7000	6845.8
RPST	16733.5

## RESULTADOS

---

### *P50(t), P90(t) y E(t)*

Al igual que sucede con los cálculos teóricos, los mejores estimadores para las previsiones son el P50(t), como medida de tendencia central, y el P90(t) como medida de valor extremo; el valor medio está afectado por valores extremos superiores y es sistemáticamente mayor al P50(t).

Globalmente, como promedio de los valores para todas las  $\theta$ s, los modelos más favorables para P50(t) corresponden a RPST y RPST truncando lejos de N fija. Las diferencias con los modelos Triangular y RPST truncando a N intermedias son de poca magnitud, unas 200 inclusiones.

El mejor comportamiento para P90(t) corresponde al modelo Triangular seguido de RPST truncado a 4500, 5305, triangular, y finalmente el RPST 7000 que está en último lugar; los valores oscilan entre 3429 y 4573.

Para E(t), los valores apuntan al también al modelo Triangular como el más favorable. Las diferencias entre los distintos modelos no llegan a 200 pacientes.

Modelos P50(t)	Promedio
RPST	1935,5
RPST7000	1984,3
RPST5305	2124,2
TRIANGULAR	2173,4
RPST4500	2406,6

Modelo P90(t)	Promedio
TRIANGULAR	3429,0
RPST4500	3875,5
RPST5305	4213,3
RPST	4482,5
RPST7000	4572,9

## RESULTADOS

---

Modelos E(t)	Promedio
TRIANGULAR	2305.2
RPST7000	2396.6
RPST5305	2413.2
RPST	2422.7
RPST4500	2508.7

### EN FUNCIÓN DE THETA

Como resumen, las estimaciones más favorables son para valores de  $\Theta=0$ , ya que de esta manera la probabilidad de que la trayectoria transcurra cercana al eje de coordenadas y corte los límites cerca del vértice interno es alta. También se podría dar el caso que  $\Theta$  fuera grande y de esta manera la trayectoria tendría una gran pendiente y podría encontrarse con los límites externos pero parece muy poco probable dada la semejanza, a priori entre ambos fármacos.

Los valores  $\Theta$  intermedios entre 0 y  $\Theta_R$ , empeoran las previsiones de los estimadores ya que es probable que la trayectoria transcurra por el centro de uno de los dos brazos sin tocar los límites.

Para valores de  $\Theta = \Theta_R$ , mejora el  $P50(t)$  y  $media(t)$  en relación con  $\Theta$  intermedias, pero empeora  $P90(t)$ .

## RESULTADOS

---

### N mínima y N máxima

Aunque los resultados de la N mínima y la N máxima no tienen demasiado interés, se explicarán únicamente con finalidad descriptiva.

En todos los casos el modelo con Mín(t) superior fue el triangular, y el más favorable el RPST, a excepción de valores de  $\theta = \theta_R$  aunque en tal caso la diferencia era mínima con el mejor modelo (38 pacientes).

Theta	Estimación	modelo	Estimador
+0 *Tr	Mín(t)	RPST	192
+0 *Tr	Mín(t)	RPST 7000	206
+0 *Tr	Mín(t)	RPST 5305	274
+0 *Tr	Mín(t)	RPST 4500	316
+0 *Tr	Mín(t)	TRIANGULAR	588
+0.5 *Tr	Mín(t)	RPST	186
+0.5 *Tr	Mín(t)	RPST 5305	226
+0.5 *Tr	Mín(t)	RPST 7000	234
+0.5 *Tr	Mín(t)	RPST 4500	262
+0.5 *Tr	Mín(t)	TRIANGULAR	514
+1 *Tr	Mín(t)	RPST 5305	176
+1 *Tr	Mín(t)	RPST 7000	214
+1 *Tr	Mín(t)	RPST	214
+1 *Tr	Mín(t)	RPST 4500	262
+1 *Tr	Mín(t)	TRIANGULAR	466

En la tabla de Máx(t) se observa que siempre el modelo más penalizante es el modelo abierto (RPST), y el que menores valores obtiene es el modelo RPST truncando a 4500 pacientes. Las diferencias son sobretodo marcadas al contrastar cualquier modelo con el diseño abierto, siendo las diferencias entre aproximadamente 5000 y 10000 pacientes en función de  $\theta$ .

## RESULTADOS

Theta	Estimación	modelo	Estimador
+0 *Tr	Máx(t)	RPST 4500	4656
+0 *Tr	Máx(t)	RPST 5305	5228
+0 *Tr	Máx(t)	TRIANGULAR	5254
+0 *Tr	Máx(t)	RPST 7000	6694
+0 *Tr	Máx(t)	RPST	12924
+0.5 *Tr	Máx(t)	RPST 4500	4928
+0.5 *Tr	Máx(t)	TRIANGULAR	5588
+0.5 *Tr	Máx(t)	RPST 5305	5706
+0.5 *Tr	Máx(t)	RPST 7000	7404
+0.5 *Tr	Máx(t)	RPST	18004
+1 *Tr	Máx(t)	RPST 4500	5140
+1 *Tr	Máx(t)	TRIANGULAR	5892
+1 *Tr	Máx(t)	RPST 5305	5996
+1 *Tr	Máx(t)	RPST 7000	7666
+1 *Tr	Máx(t)	RPST	18046

### *P50(t), P90(t) y E(t)*

Los modelos más favorables para P50(t) corresponden, por orden, a los diseños RPST, RPST truncando a 7000, 5305, triangular y RPST 4500. Esta consideración se confirma en todo el abanico de posibilidades de Theta.

Theta	Estimación	modelo	Estimador
+0 *Tr	P50(t)	RPST	2034
+0 *Tr	P50(t)	RPST 7000	2066
+0 *Tr	P50(t)	RPST 5305	2194
+0 *Tr	P50(t)	TRIANGULAR	2246
+0 *Tr	P50(t)	RPST 4500	2434
+0 *Tr	P50(t)	CLASICO	3962
+0.5 *Tr	P50(t)	RPST	2428
+0.5 *Tr	P50(t)	RPST 7000	2502
+0.5 *Tr	P50(t)	TRIANGULAR	2636
+0.5 *Tr	P50(t)	RPST 5305	2700
+0.5 *Tr	P50(t)	RPST 4500	3106
+0.5 *Tr	P50(t)	CLASICO	3962
+1 *Tr	P50(t)	RPST	2316
+1 *Tr	P50(t)	RPST 7000	2356
+1 *Tr	P50(t)	RPST 5305	2546
+1 *Tr	P50(t)	TRIANGULAR	2554
+1 *Tr	P50(t)	RPST 4500	2872
+1 *Tr	P50(t)	CLASICO	3962

## RESULTADOS

---

La estimación de  $P_{90}(t)$  varia en función de la  $\theta$ , aunque cabe destacar el buen comportamiento del triangular en todo el abanico y que para diferencias para  $\theta=0$  son relativamente pequeñas entre los distintos modelos, 700 pacientes. Para valores de  $\theta$  intermedios, las diferencias entre los modelos pueden superar varios miles.

Theta	Estimación	Modelo	Estimador
+0 *Tr	$P_{90}(t)$	TRIANGULAR	3257
+0 *Tr	$P_{90}(t)$	RPST	3519
+0 *Tr	$P_{90}(t)$	RPST 7000	3676
+0 *Tr	$P_{90}(t)$	RPST 5305	3822
+0 *Tr	$P_{90}(t)$	RPST 4500	3890
+0 *Tr	$P_{90}(t)$	CLASICO	3962
+0.5 *Tr	$P_{90}(t)$	CLASICO	3962
+0.5 *Tr	$P_{90}(t)$	TRIANGULAR	4010
+0.5 *Tr	$P_{90}(t)$	RPST 4500	4350
+0.5 *Tr	$P_{90}(t)$	RPST 5305	5006
+0.5 *Tr	$P_{90}(t)$	RPST	5769
+0.5 *Tr	$P_{90}(t)$	RPST 7000	6056
+1 *Tr	$P_{90}(t)$	CLASICO	3962
+1 *Tr	$P_{90}(t)$	TRIANGULAR	4115
+1 *Tr	$P_{90}(t)$	RPST 4500	4584
+1 *Tr	$P_{90}(t)$	RPST 5305	5256
+1 *Tr	$P_{90}(t)$	RPST	5853
+1 *Tr	$P_{90}(t)$	RPST 7000	5940

## RESULTADOS

---

Para la estimación de  $E(t)$  los resultados son similares a los hallados con el  $P50(t)$  en el sentido de que todos los modelos son mejores que un diseño convencional a  $N$  fija pero el modelos más favorecido es, en general, el triangular.

Theta	Estimación	modelo	Estimador
+0 *Tr	$E(t)$	RPST	2307,7
+0 *Tr	$E(t)$	RPST 7000	2377,7
+0 *Tr	$E(t)$	TRIANGULAR	2401,3
+0 *Tr	$E(t)$	RPST 5305	2474,5
+0 *Tr	$E(t)$	RPST 4500	2643,1
+0 *Tr	$E(t)$	CLASICO	3962,0
+0.5 *Tr	$E(t)$	TRIANGULAR	2786,9
+0.5 *Tr	$E(t)$	RPST 5305	3050,6
+0.5 *Tr	$E(t)$	RPST 7000	3069,3
+0.5 *Tr	$E(t)$	RPST	3117,1
+0.5 *Tr	$E(t)$	RPST 4500	3127,7
+0.5 *Tr	$E(t)$	CLASICO	3962,0
+1 *Tr	$E(t)$	TRIANGULAR	2671,2
+1 *Tr	$E(t)$	RPST 7000	2866,1
+1 *Tr	$E(t)$	RPST 5305	2867,7
+1 *Tr	$E(t)$	RPST 4500	2930,3
+1 *Tr	$E(t)$	RPST	2961,7
+1 *Tr	$E(t)$	CLASICO	3962,0



### *COMPROBACIÓN MEDIANTE SIMULACIONES*

Para comprobar la validez de los distintos diseños se han comparado los resultados teóricos y los observados con las simulaciones. Los resultados se presentan mediante figuras para facilitar la comprensión.

En general se observa muy buena concordancia entre los valores previstos y los valores observados tras las simulaciones para los percentiles 50, 90 y promedio, respectivamente  $P50(t)$ ,  $P90(t)$  y  $E(t)$ . Este hecho se puede comprobar en todos los distintos modelos, siendo el triangular el que más se ajusta a los valores previstos.

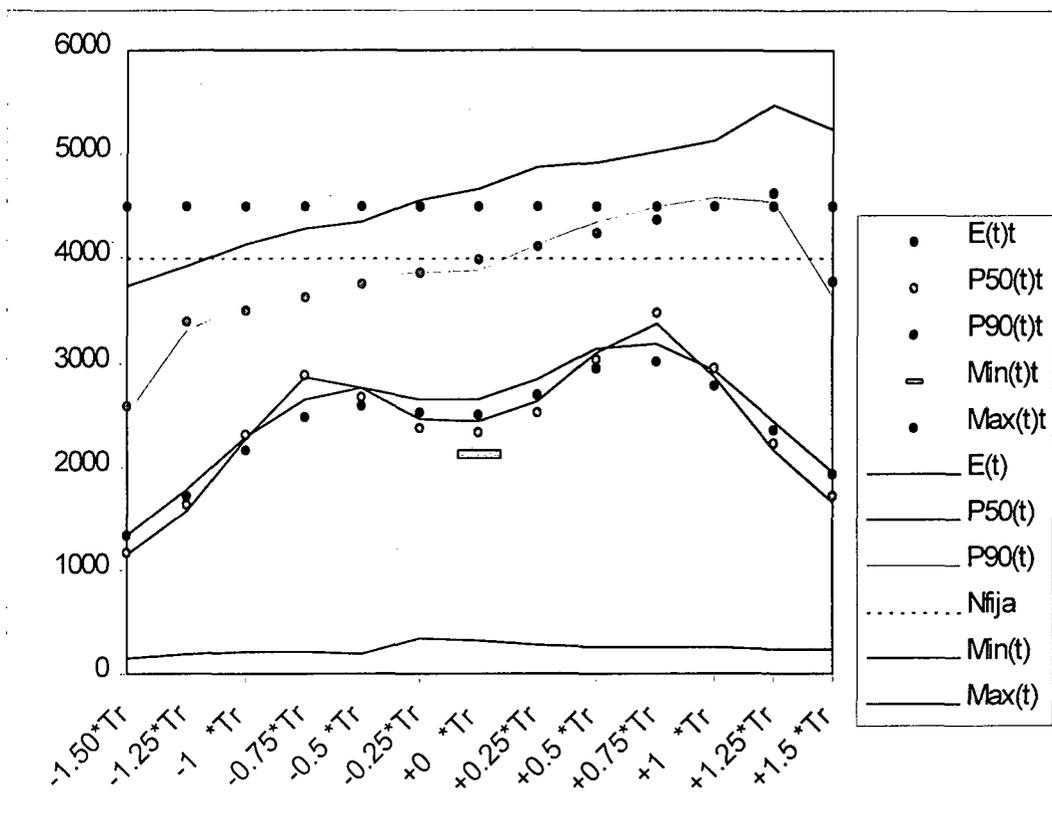
La única excepción se observa en los valores de  $N$  máxima. Los valores teóricos tienen el valor fijo del punto de truncamiento o del vértice y son constantes en todos los valores de  $\theta$ . En cambio las simulaciones indican que a medida que aumenta  $\theta$  se requerirá mayor tamaño muestral superando las previsiones, en general, aproximadamente con  $\theta \geq 0$ .

En las siguientes figuras se podrá comprobar la semejanza entre los valores teóricos y las correspondientes simulaciones en función de distintos valores de  $\theta$ .

Las distintas series se distinguen por el color, mostrándose siempre los valores correspondientes a las simulaciones mediante una línea continua y los esperados con puntos. En el caso del valor mínimo teórico, solamente se muestra su representación para  $\theta = 0$ , ya que es el valor del supuesto corte en el vértice interno. También se representa en línea discontinua el tamaño muestral que se requeriría en un estudio convencional con el tamaño muestral predefinido.

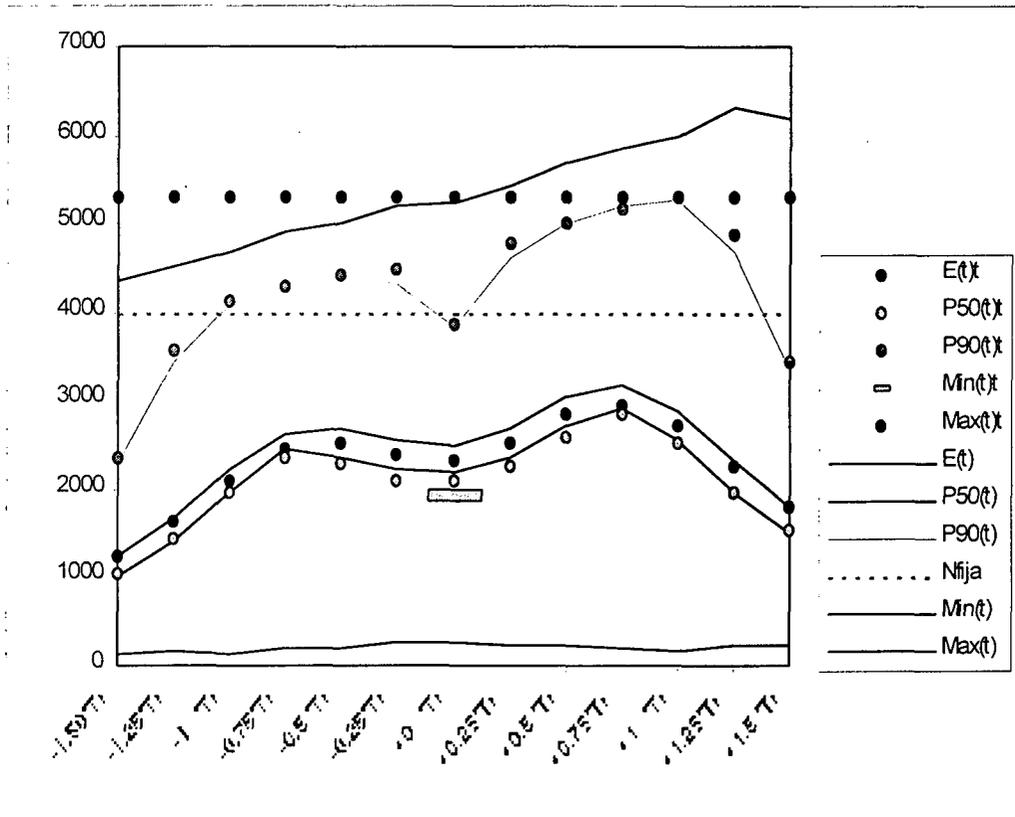
# RESULTADOS

RPST4500



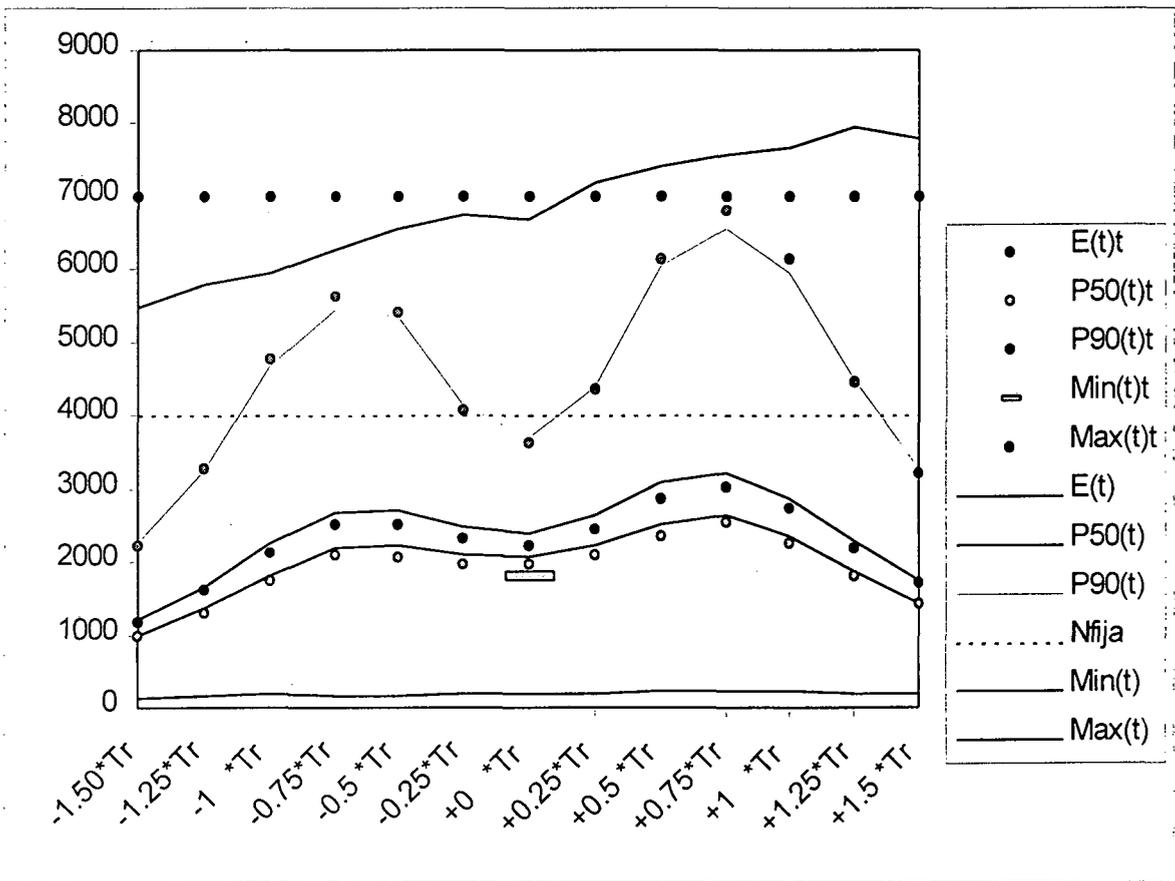
# RESULTADOS

RPST5305



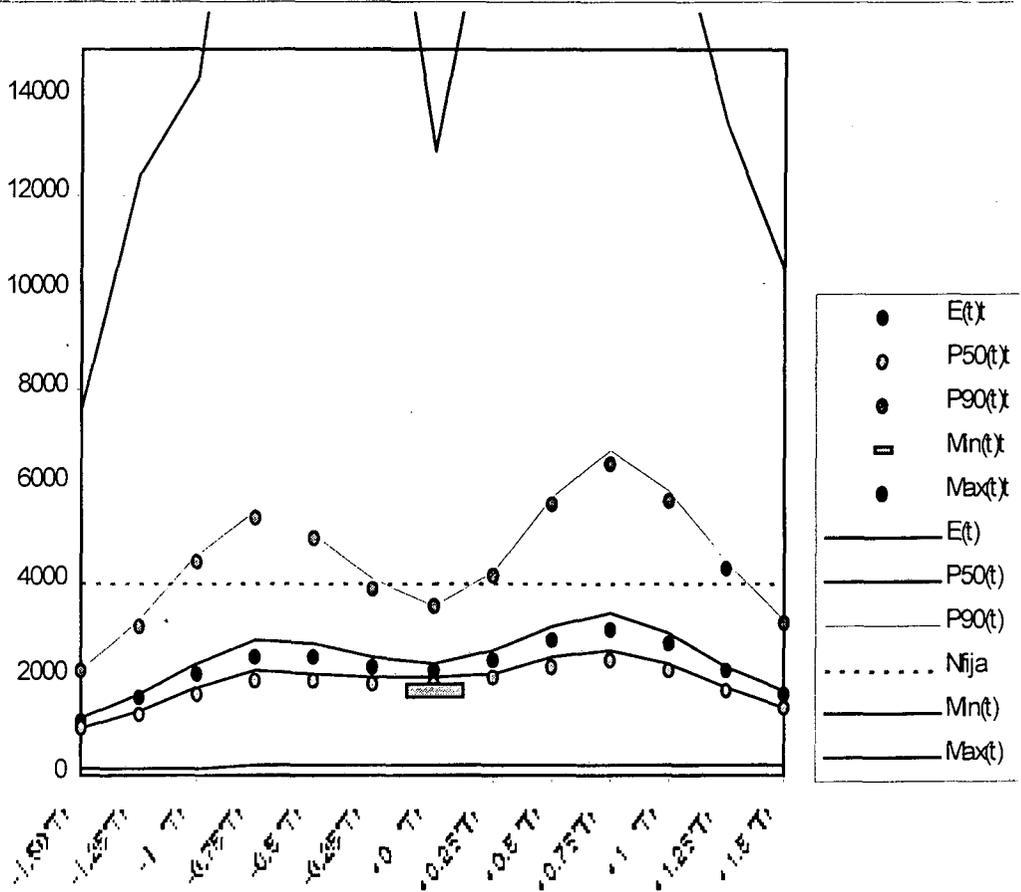
# RESULTADOS

RPST7000



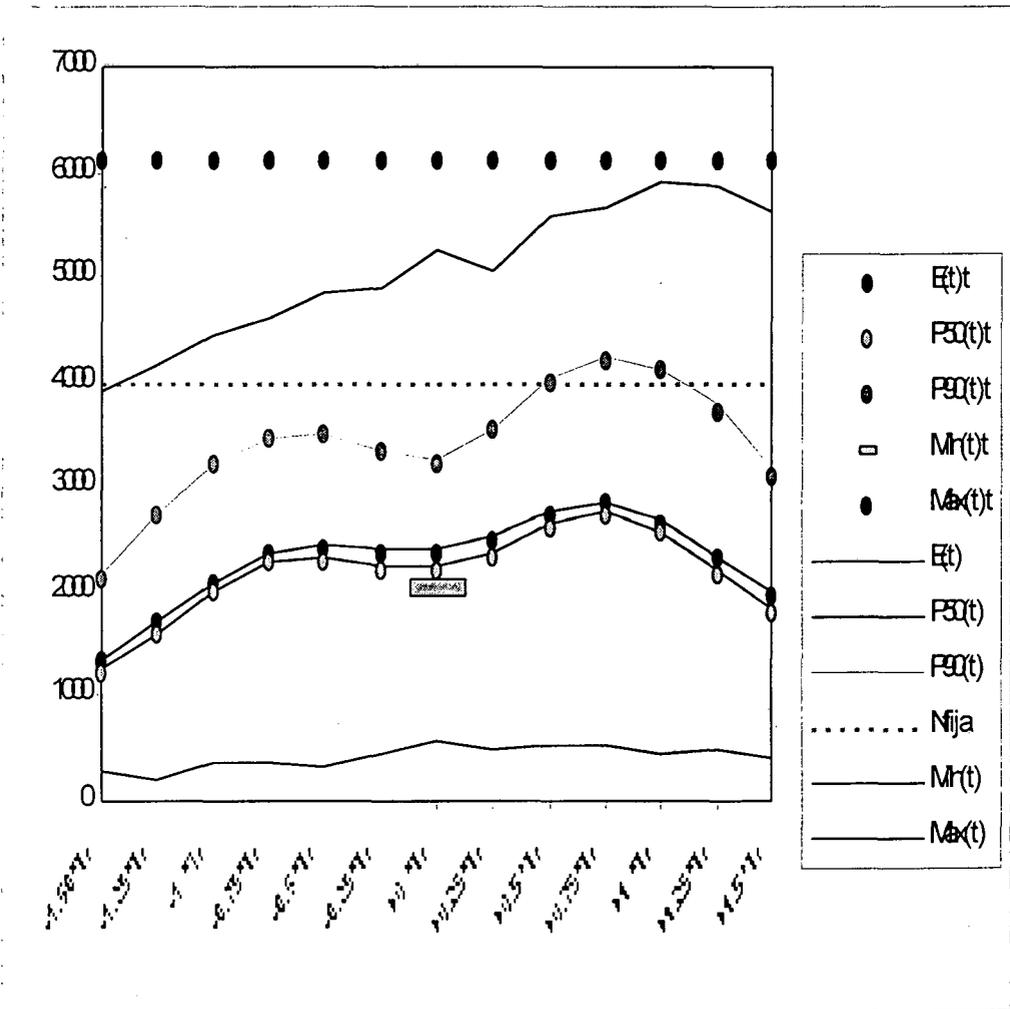
# RESULTADOS

RPST



# RESULTADOS

Triangular



### *AHORRO DE LOS MÉTODOS SECUENCIALES*

El ahorro de los métodos secuenciales se ha cuantificado con las simulaciones desde dos puntos de vista. En primer lugar se ha estudiado el porcentaje de reducción en el número de pacientes, y posteriormente que porcentaje de estudios necesitan menor tamaño muestral que un diseño convencional.

### **AHORRO: REDUCCIONES EN EL NÚMERO DE PACIENTES**

#### GLOBAL

Tomando los valores promedio para evaluar un ahorro general de los métodos secuenciales podemos concluir que al menos en un 50% de los estudios el ahorro es superior al 40% y puede llegar al 50%. Evidentemente, no todos los modelos se comportan igual y las diferencias son más evidentes al estudiar el  $P90(t)$ : al menos en un 10% de los estudios se puede pasar de sólo ahorrar un 14% a requerir 14% más muestra que en estudio convencional. Los valores promedio de  $E(t)$  indican que en un el ahorro global oscila entre un 37% y un 42%.

El triangular mostró ser el modelo mejor para los estimadores  $E(t)$  y  $P90(t)$ , mientras que para  $P50(t)$  fue el RPST. Las diferencias entre modelos son del 12% para  $P50(t)$ , 28% para  $P90(t)$  y sólo del 5% para  $E(t)$ .

Modelos $P50(t)$	Promedio
RPST	51.6%
RPST7000	50.4%
RPST5305	46.9%
TRIANGULAR	45.7%
RPST4500	39.8%

## RESULTADOS

---

Modelo P90(t)	Promedio
TRIANGULAR	14,3%
RPST4500	3,1%
RPST5305	-5,3%
RPST	-12,1%
RPST7000	-14,3%

Modelos E(t)	Promedio
TRIANGULAR	42,4%
RPST7000	40,1%
RPST5305	39,7%
RPST	39,4%
RPST4500	37,3%

## RESULTADOS

---

### *Por thetas*

En las tablas siguientes se muestran los valores observados en función de distintas diferencias entre los tratamientos. Al igual que en el resto de los parámetros estudiados, se muestran sólo las tablas con los resultados representativos. Para los valores negativos e intermedios de  $\theta$  los resultados son similares y casi simétricos pero con valores más favorables puesto que la varianza es menor (ver tablas completas en apéndice y figuras).

Los modelos más favorables para  $P50(t)$  corresponden, por orden, a los modelos RPST, RPST truncando a valores lejanos del tamaño de  $N$  fija, el triangular y RPST truncado a  $N$  bajas. Esta consideración se confirma en todo el abanico de posibilidades de  $\theta$ . Las diferencias entre RPST y RPST4500 son del varían entre 10% para  $\theta=0$  y 16,9 para  $\theta = 0,5*\theta_R$ .

Theta	Estimación	modelo	Ahorro (dif %)
+0 *Tr	P50(t)	RPST	49,2%
+0 *Tr	P50(t)	RPST 7000	48,4%
+0 *Tr	P50(t)	RPST 5305	45,2%
+0 *Tr	P50(t)	TRIANGULAR	43,9%
+0 *Tr	P50(t)	RPST 4500	39,2%
+0.5 *Tr	P50(t)	RPST	39,3%
+0.5 *Tr	P50(t)	RPST 7000	37,5%
+0.5 *Tr	P50(t)	TRIANGULAR	34,1%
+0.5 *Tr	P50(t)	RPST 5305	32,5%
+0.5 *Tr	P50(t)	RPST 4500	22,4%
+1 *Tr	P50(t)	RPST	42,1%
+1 *Tr	P50(t)	RPST 7000	41,1%
+1 *Tr	P50(t)	RPST 5305	36,4%
+1 *Tr	P50(t)	TRIANGULAR	36,2%
+1 *Tr	P50(t)	RPST 4500	28,2%

## RESULTADOS

La estimación de  $P90(t)$  varia en función de la Theta:

Para  $\theta = 0$  los modelos más favorables son, por orden, triangular, RPST y RPST truncando lejos a cerca; las diferencias entre el mejor y peor modelo superan el 15%.

Para el resto de las opciones,  $\theta = 0,5 * \theta_R$  y  $\theta = \theta_R$ , se mantiene el orden, Triangular, RPST truncando cercano a N fija, RPST, y RPST truncando lejos; las diferencias entre los valores extremos superan el 45%.

Theta	Estimación	Modelo	Ahorro (dif %)
+0 *Tr	$P90(t)$	TRIANGULAR	18,6%
+0 *Tr	$P90(t)$	RPST	12,0%
+0 *Tr	$P90(t)$	RPST 7000	8,1%
+0 *Tr	$P90(t)$	RPST 5305	4,5%
+0 *Tr	$P90(t)$	RPST 4500	2,8%
+0,5 *Tr	$P90(t)$	TRIANGULAR	-0,3%
+0,5 *Tr	$P90(t)$	RPST 4500	-8,8%
+0,5 *Tr	$P90(t)$	RPST 5305	-25,2%
+0,5 *Tr	$P90(t)$	RPST	-44,2%
+0,5 *Tr	$P90(t)$	RPST 7000	-51,4%
+1 *Tr	$P90(t)$	TRIANGULAR	-2,9%
+1 *Tr	$P90(t)$	RPST 4500	-14,6%
+1 *Tr	$P90(t)$	RPST 5305	-31,4%
+1 *Tr	$P90(t)$	RPST	-46,3%
+1 *Tr	$P90(t)$	RPST 7000	-48,5%

Los valores teóricos indican que los mejores modelos en función de theta son, para  $\theta = 0$ , RPST, RPST truncando lejos hasta 5305, Triangular, RPST truncando cerca; para diferencias intermedias, Triangular, RPST, RPST truncando a valores intermedios, RPST truncando a valores altos, RPST truncando a valores bajos; y para

## RESULTADOS

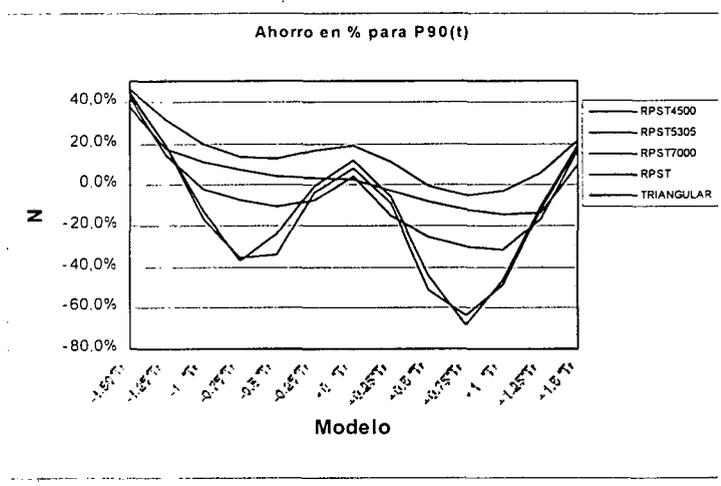
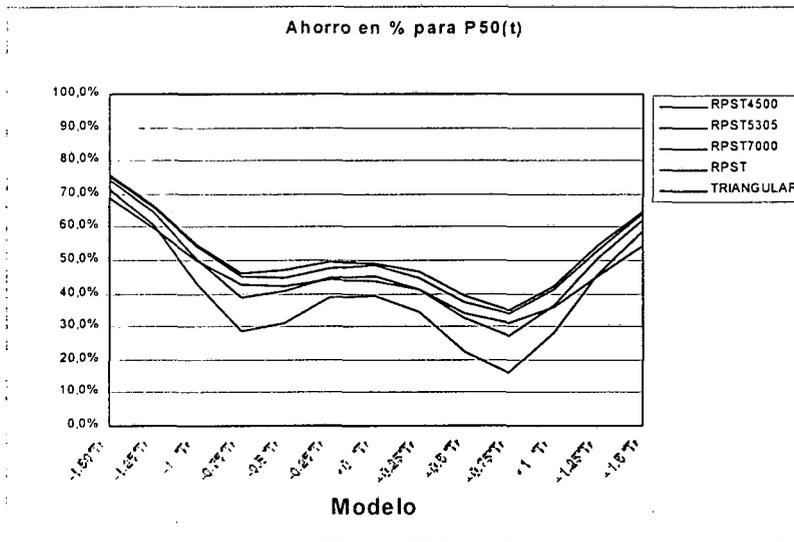
---

Theta = Theta R, Triangular, RPST truncando a valores intermedios, RPST, y RPST al resto de valores.

Para la estimación de  $E(t)$  el modelo triangular es el que presenta mejores propiedades para theta distintas a 0. Las diferencias en las distintas estimaciones entre modelos son en general no muy grandes y menores a 100 pacientes.

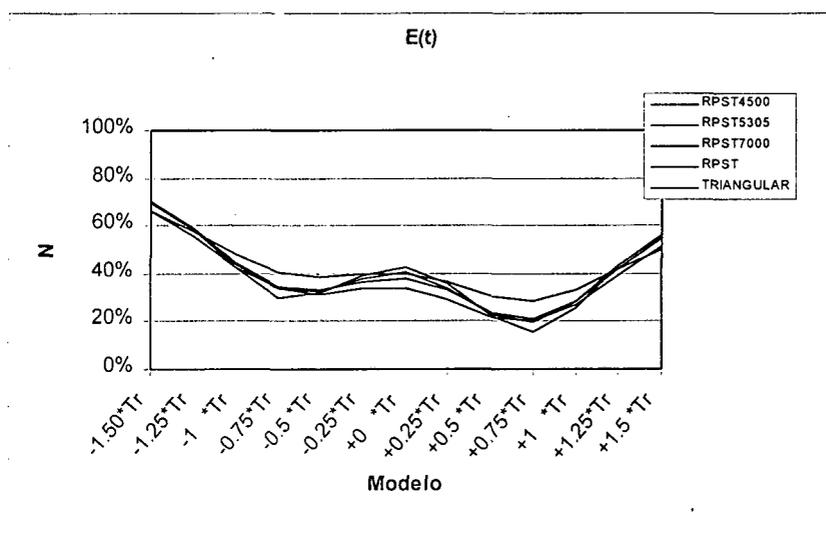
Theta	Estimación	modelo	Ahorro (dif %)
+0 *Tr	E(t)	RPST	42,3%
+0 *Tr	E(t)	RPST 7000	40,6%
+0 *Tr	E(t)	TRIANGULAR	40,0%
+0 *Tr	E(t)	RPST 5305	38,1%
+0 *Tr	E(t)	RPST 4500	33,9%
+0.5 *Tr	E(t)	TRIANGULAR	30,3%
+0.5 *Tr	E(t)	RPST 5305	23,7%
+0.5 *Tr	E(t)	RPST 7000	23,3%
+0.5 *Tr	E(t)	RPST	22,1%
+0.5 *Tr	E(t)	RPST 4500	21,8%
+1 *Tr	E(t)	TRIANGULAR	33,2%
+1 *Tr	E(t)	RPST 7000	28,3%
+1 *Tr	E(t)	RPST 5305	28,3%
+1 *Tr	E(t)	RPST 4500	26,7%
+1 *Tr	E(t)	RPST	26,0%

# RESULTADOS



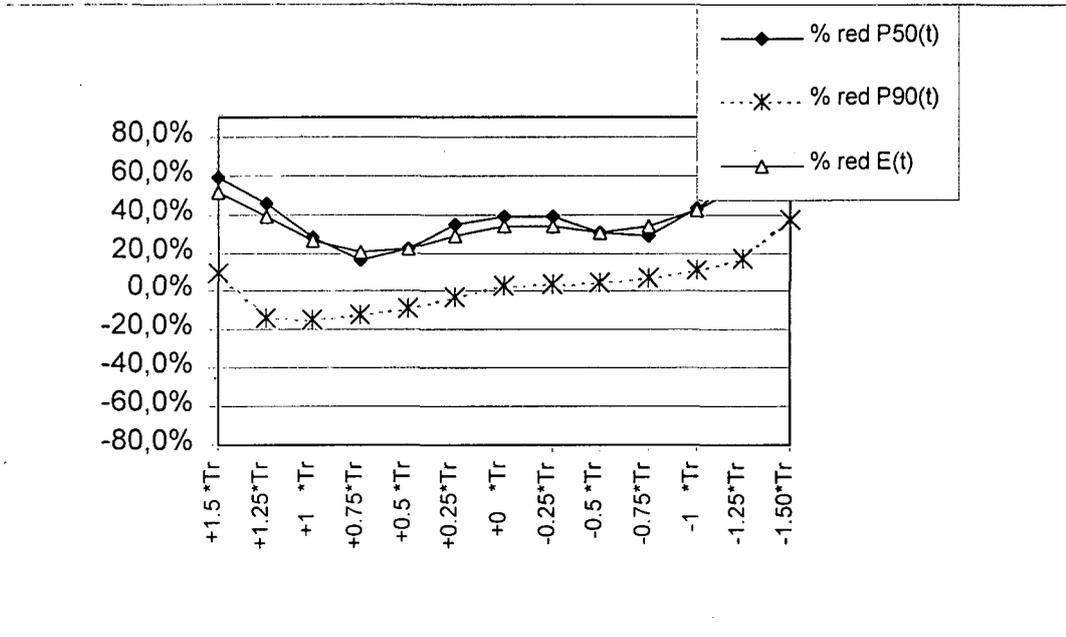
# RESULTADOS

---

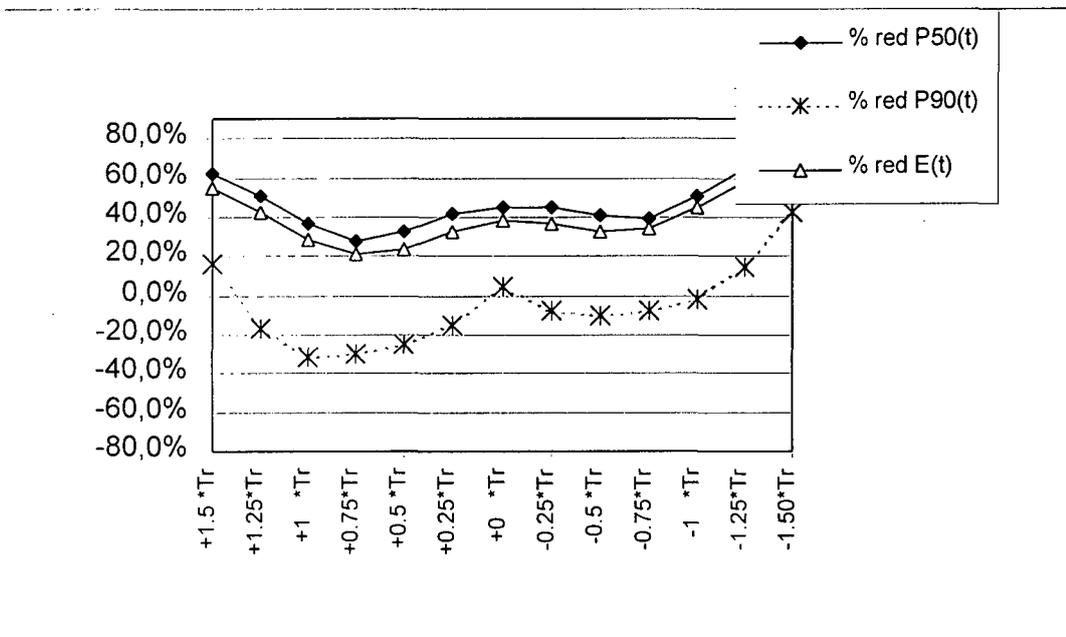


# RESULTADOS

## RPST 4500

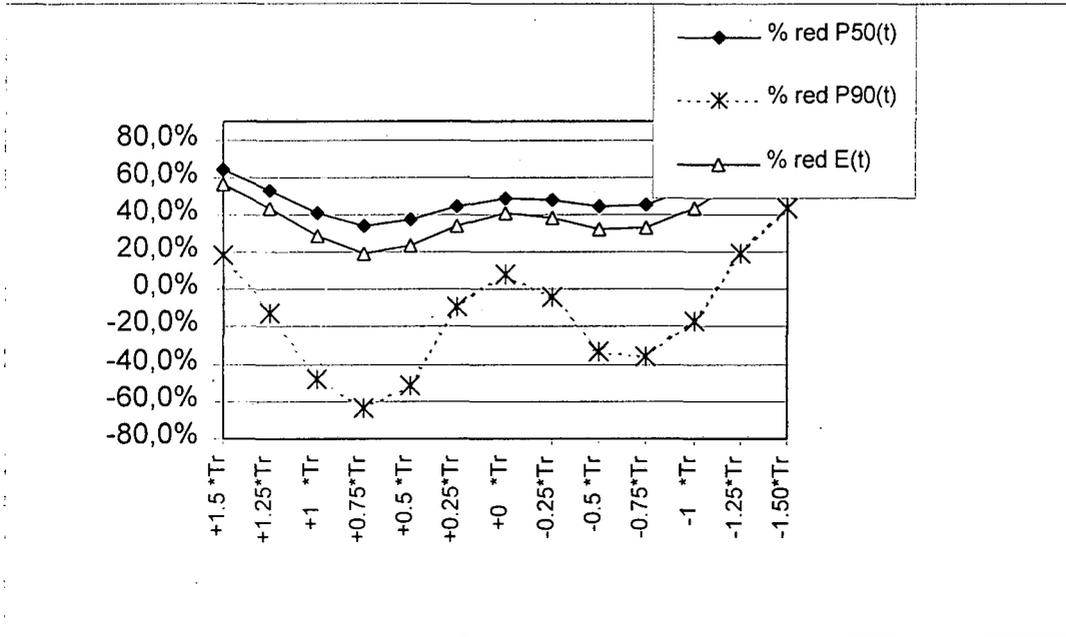


## RPST 5305

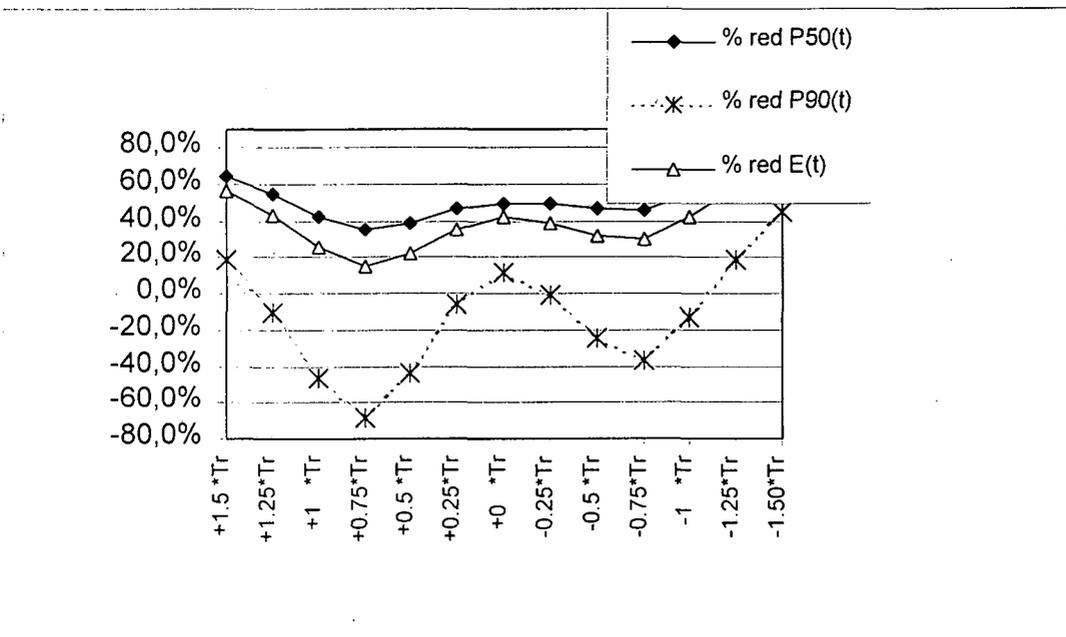


# RESULTADOS

## RPST 7000

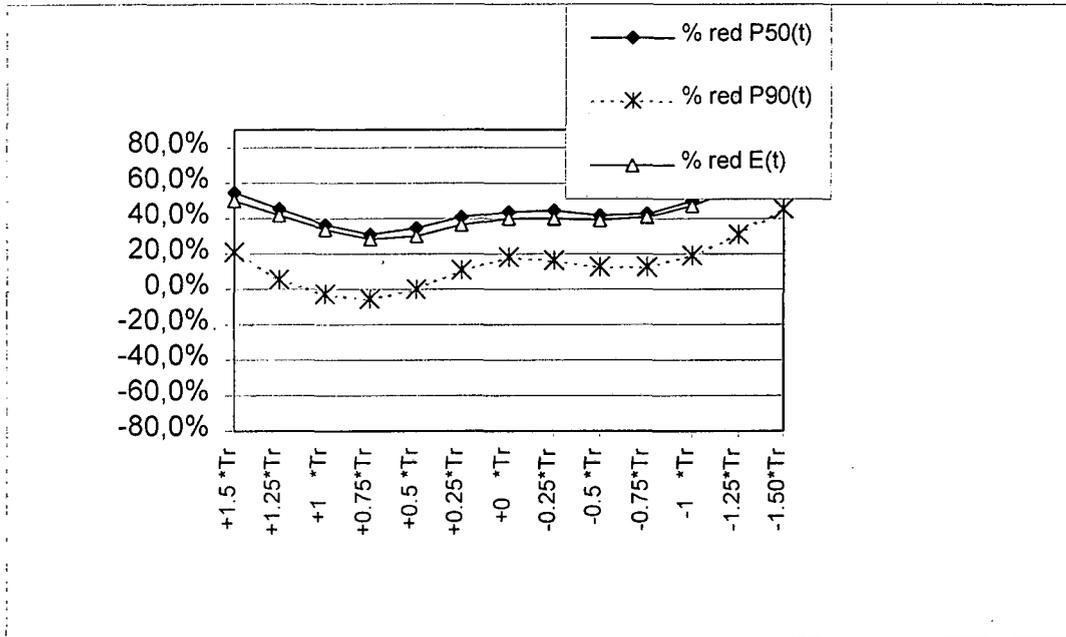


## RPST



# RESULTADOS

Triangulat



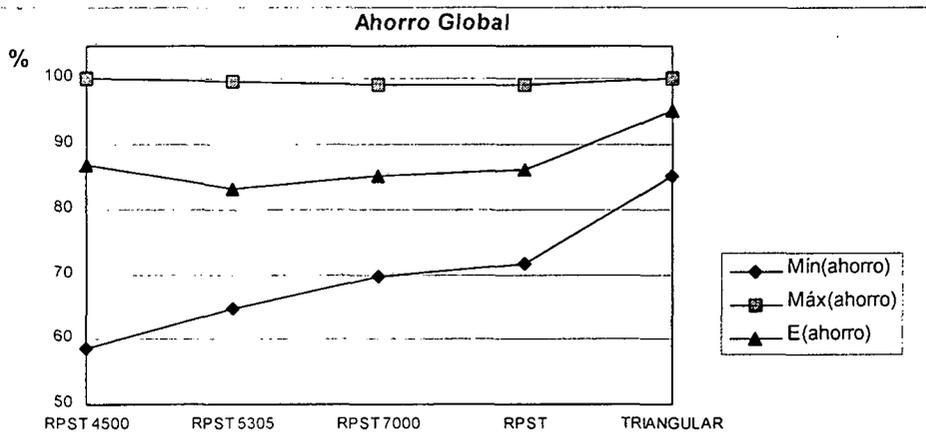
**AHORRO: PORCENTAJE DE ESTUDIOS QUE NO REQUIEREN EL TAMAÑO MUESTRAL CONVENCIONAL**

**GLOBAL**

En las estimaciones promedio, se puede ver que como en promedio en todos los modelos se supera el 80% de estudios concluidos sin llegar a requerir el tamaño muestral necesitado para un estudio convencional.

En la tabla siguiente se indican el porcentaje de estudios en los que se ha podido ahorrar menos y más pacientes para alguna theta determinada. El modelo RPST 4500 presenta muy buen ahorro en promedio, pero es el que presenta el porcentaje de ahorro más bajo para alguno de los casos.

Modelo	E(ahorro)	Min(ahorro)	Máx(ahorro)
TRIANGULAR	95,1	85,1	100,0
RPST 4500	86,9	58,6	100,0
RPST	86,1	71,7	99,0
RPST 7000	85,1	69,7	99,0
RPST 5305	83,1	64,7	99,6



## RESULTADOS

---

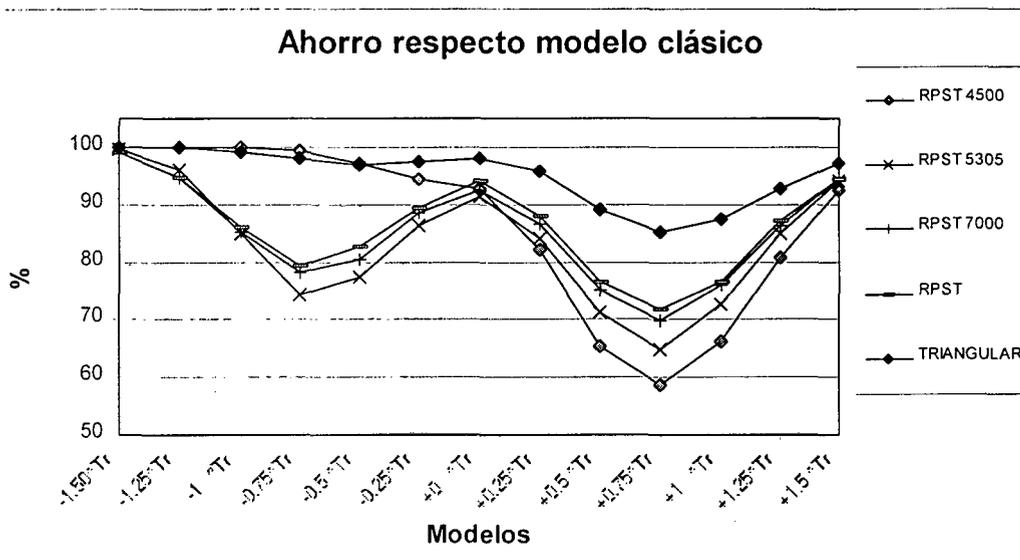
### *POR THETAS*

En general, los dos modelos que presentan el mejor son el modelo triangular y el RPST. En la tabla siguiente se presentan los valores para  $\theta = 0$ ,  $\theta = 0.5 * \theta_R$  y  $\theta = \theta_R$ ; los valores para el resto de posibilidades de  $\theta$  están detallados en el apéndice y podrán observarse en la figura siguiente.

Theta	Modelo	AHORRO
+0 *Tr	TRIANGULAR	98.1
+0 *Tr	RPST	94.1
+0 *Tr	RPST 4500	92.7
+0 *Tr	RPST 7000	92.4
+0 *Tr	RPST 5305	91.2
+0.5 *Tr	TRIANGULAR	89.1
+0.5 *Tr	RPST	76.6
+0.5 *Tr	RPST 7000	75.0
+0.5 *Tr	RPST 5305	71.1
+0.5 *Tr	RPST 4500	65.4
+1 *Tr	TRIANGULAR	87.5
+1 *Tr	RPST	76.6
+1 *Tr	RPST 7000	76.1
+1 *Tr	RPST 5305	72.5
+1 *Tr	RPST 5000	71.2
+1 *Tr	RPST 4500	66.2

## RESULTADOS

En esta figura se comprueba que el ahorro no es simétrico alrededor de los valores de theta y que para valores negativos de theta la optimización es mayor. Este hecho se acusa sobretodo en el modelo RPST 4500 el cual está situado en la mejor posición para valores negativos de theta y empeora en gran medida con los valores positivos de theta. El diseño triangular es que acusa esta variación en menor medida.



### *CARACTERÍSTICA OPERANTE*

Todos los modelos secuenciales presentan un buen comportamiento a nivel de la CO teórica, hecho que se comprueba y constata mediante simulaciones. Aunque se presentan las tablas simplificadas para  $\theta = 0$ ,  $\theta = \theta_R$  y para  $\theta = 0.5 * \theta_R$ , esta afirmación se puede constatar en las figuras siguientes y los valores se detallan en el anexo correspondiente.

Para el error de tipo I se ha observado en las simulaciones un aumento máximo del 5 por mil para el modelo RPST y en ningún se ha constatado valores menores a los teóricos.

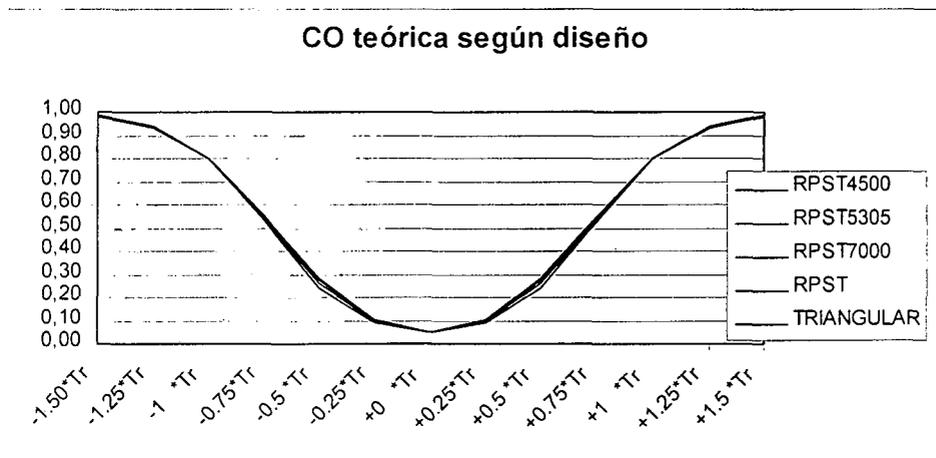
El diseño clásico ha experimentado un muy discreto aumento en el error de tipo II (0.6%) que no se ha producido en los modelos secuenciales. En todos éstos, el error  $\beta$  ha sido algo inferior respecto a las previsiones teóricas, entre 0.7% -RPST 5305- y 2% -RPST 7000-. En todos los modelos secuenciales se puede comprobar la simetría de la curva de CO a diferencia de lo que sucede con el modelo clásico. Éste presenta valores ligeramente superiores en la mitad de la curva correspondiente a negativos de theta, llegando hasta una disminución del error de tipo II de 8% para  $\theta = -1 * \theta_R$ .

A continuación se presentan las tablas resumidas de las estimaciones teóricas y las simulaciones

### *ESTIMACIONES TEÓRICAS*

## RESULTADOS

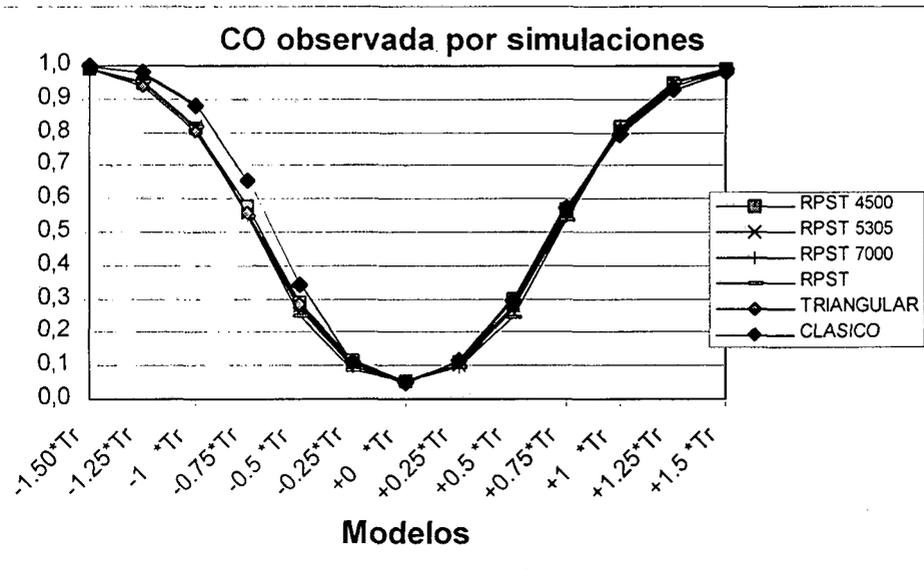
Theta	Modelos	P(rechazar H0   Theta)
+0 *T	RPST4500	0,05000
+0 *T	RPST5305	0,05000
+0 *T	RPST7000	0,05000
+0 *T	RPST	0,05000
+0 *T	TRIANGULAR	0,05000
+0.5 *T	RPST	0,24465
+0.5 *T	RPST7000	0,25900
+0.5 *T	RPST5305	0,27268
+0.5 *T	TRIANGULAR	0,28315
+0.5 *T	RPST4000 .	0,28802
+1 *T	TRIANGULAR	0,80001
+1 *T	RPST4500	0,80004
+1 *T	RPST5305	0,80010
+1 *T	RPST7000	0,80015
+1 *T	RPST	0,80016



# RESULTADOS

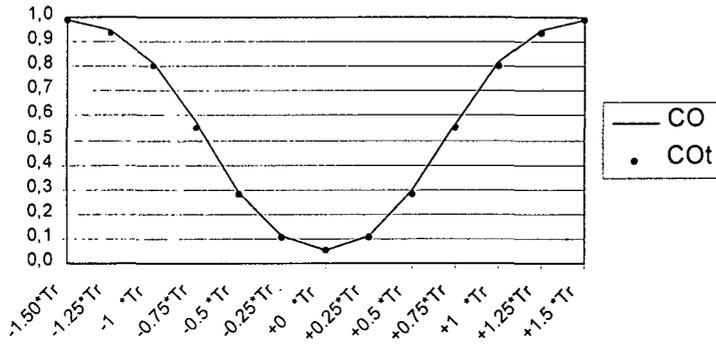
## ESTIMACIONES SIMULADAS

Theta	Modelo	Estimación P(aceptar H0   Theta)
+0 *Tr	CLASICO	0,050
+0 *Tr	RPST 4500	0,051
+0 *Tr	RPST 5305	0,051
+0 *Tr	TRIANGULAR	0,053
+0 *Tr	RPST 7000	0,054
+0 *Tr	RPST	0,055
+0,5 *Tr	RPST	0,245
+0,5 *Tr	RPST 7000	0,270
+0,5 *Tr	RPST 5305	0,278
+0,5 *Tr	TRIANGULAR	0,279
+0,5 *Tr	CLASICO	0,295
+0,5 *Tr	RPST 4500	0,297
+1 *Tr	CLASICO	0,794
+1 *Tr	RPST 5305	0,807
+1 *Tr	TRIANGULAR	0,807
+1 *Tr	RPST	0,811
+1 *Tr	RPST 4500	0,819
+1 *Tr	RPST 7000	0,820

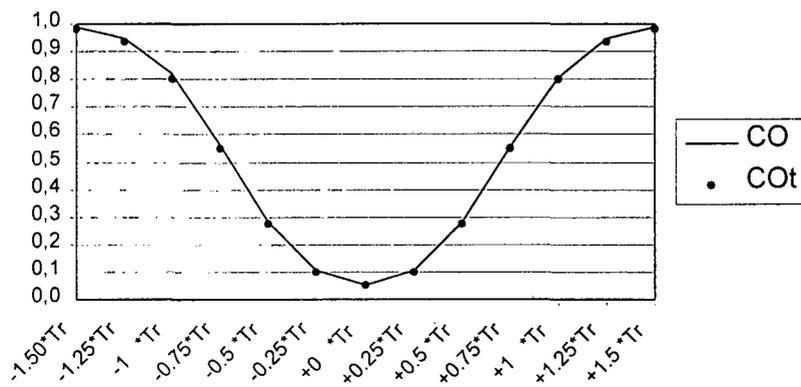


# RESULTADOS

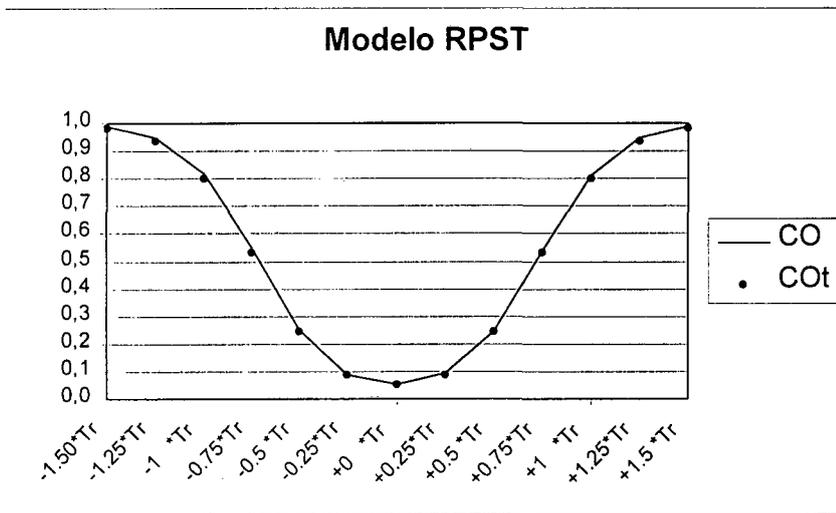
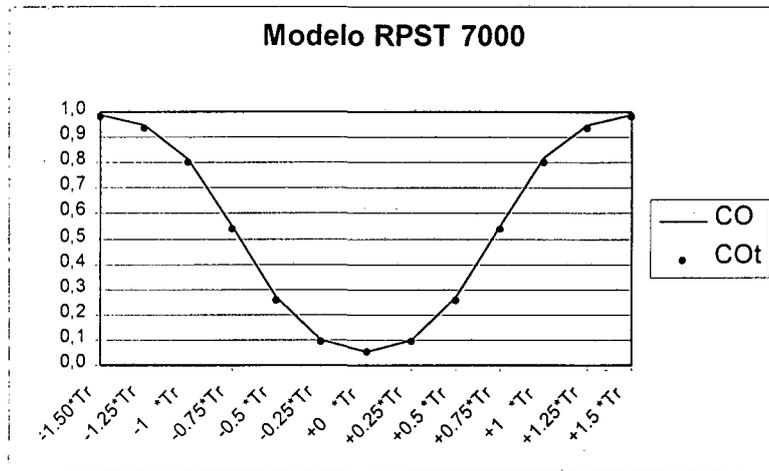
### Modelo RPST 4500



### Modelo RPST 5305

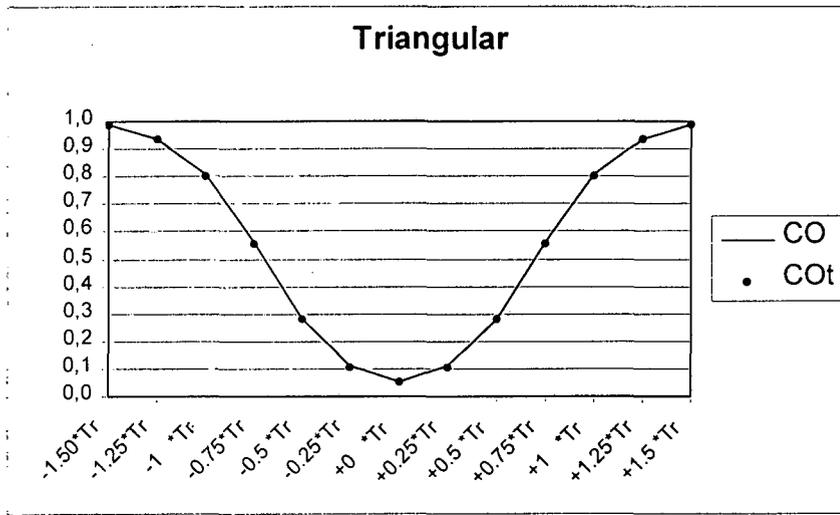


# RESULTADOS



# RESULTADOS

---



## ***DISEÑO, MONITORIZACIÓN Y ANÁLISIS ESTADÍSTICO DEL ESTUDIO TIM***

### ***DISEÑO***

El estudio TIM se diseñó usando el modelo secuencial de límites continuos RPST truncando al valor de 5305 pacientes. Las características estadísticas seleccionadas fueron  $\alpha=0.05$  bilateral y  $\beta=0.2$ . La diferencia clínicamente relevante se estableció en un 2.5% (porcentaje absoluto) para un valor de base para el tratamiento experimental del 10%; ( $\theta_R=0.3151$  y  $OR = 0.730$ ).

La región de continuación queda delimitada de la siguiente manera (ver tabla de valores de origen y pendiente en material y método):

Límites externos:

$$Z = 8.56449 + 0.22354*V; \text{ rechazo de } H_0 \text{ a favor de experimental}$$

$$Z = -8.56449 - 0.22354*V; \text{ rechazo de } H_0 \text{ a favor de control}$$

Límites internos: (a partir de  $V=38.31$ ); aceptación de  $H_0$

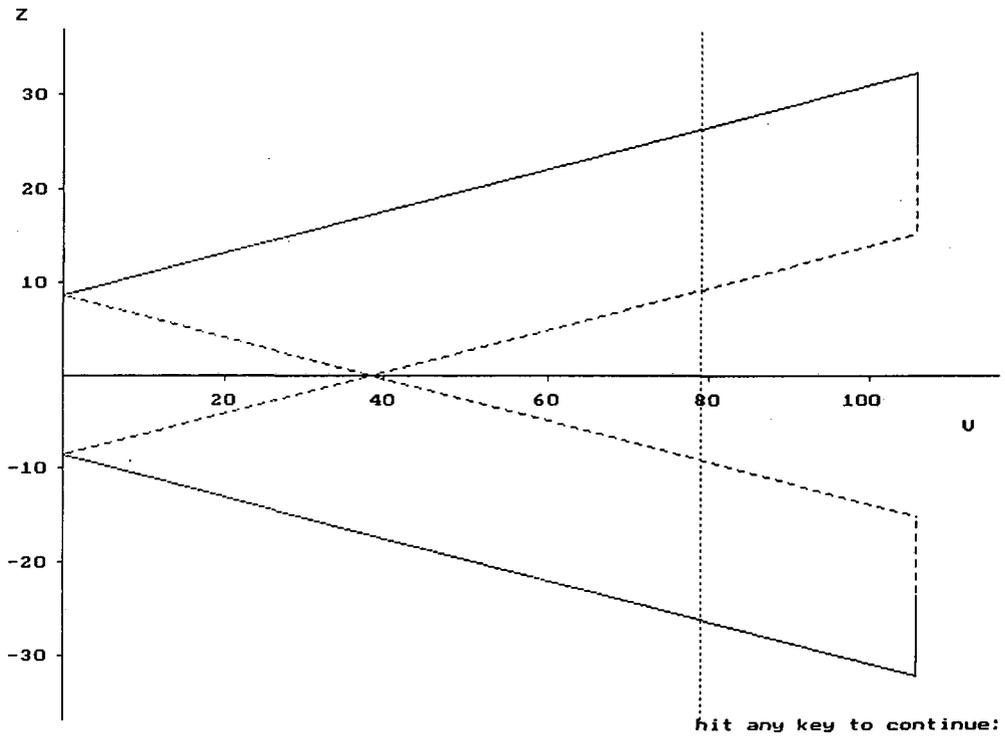
$$Z = 8.56449 - 0.22354*V$$

$$Z = -8.56449 + 0.22354*V$$

En límite truncamiento ( $V=105.892$ ) si  $Z$  está entre  $\pm 23.67$ , se aceptará la  $H_0$ ; en caso contrario  $H_A$

## RESULTADOS

---



Los detalles generales del método secuencial RPST truncado están ampliamente descritos en introducción. Las características específicas de CO, expectativas de tamaño muestral y ahorro de pacientes para este modelo, se desarrollan en los apartados anteriores de resultados.

*MONITORIZACIÓN*

**PRE-REVISIÓN OFICIAL**

Tras aproximadamente dos años de inclusiones se decidió realizar el primer corte secuencial. Las fechas concretas de realización de las distintas inspecciones, así como el número de pacientes con end-point y los porcentajes se especifican en la siguiente tabla.

Corte	Fecha	A			B			Total			A-B
		n	N	%	n	N	%	n	N	%	%
1	1-dic-94	58	466	12,45%	43	456	9,43%	101	922	10,95%	3,02%
2	1-jun-95	65	614	10,59%	57	617	9,24%	122	1231	9,91%	1,35%
3	27-nov-95	77	729	10,56%	67	718	9,33%	144	1447	9,95%	1,23%
4	20-may-96	92	872	10,55%	83	856	9,70%	175	1728	10,13%	0,85%
5	4-feb-97	107	1044	10,25%	101	1035	9,76%	208	2079	10,00%	0,49%

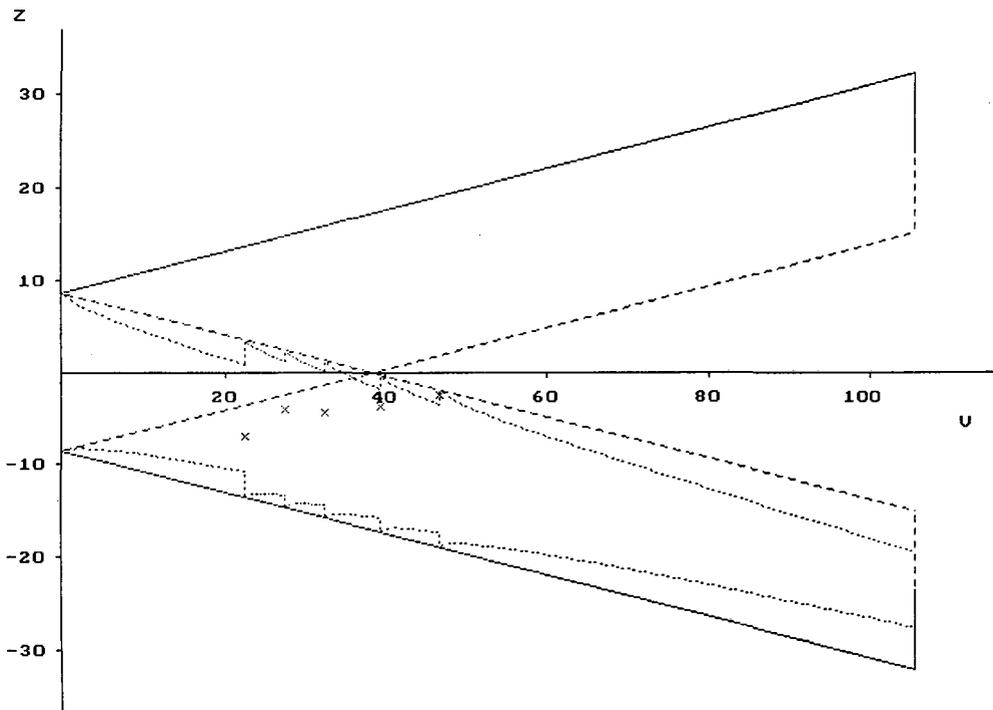
Llama la atención que, aunque el porcentaje global de eventos se ha mantenido más o menos constante con leves oscilaciones alrededor del 10%, las diferencias entre ambos grupos que inicialmente eran relativamente grandes (3.02% absoluto, 24.26% relativo) han ido paulatinamente evolucionando hacia la igualdad (0.49% absoluto, 4.78 relativo).

Los incrementos de información en términos de eventos, pacientes y tiempo han sido similares a los previstos y de similar magnitud a partir de la primera inspección. Recordamos que ésta se preestableció realizarla a las 1000 inclusiones. No obstante, el penúltimo y último corte se realizaron con más información, sobretudo en unidades de V: entre un 41% y un 53% más que en los cortes 2 y 3. El reclutamiento por mes ha ido variando: 41.97, 50.93, 36.20, 48.17 y 40.50 pacientes mes respectivamente en las inspecciones 1 a 5.

## RESULTADOS

En la tabla siguiente se presentan los distintos tipos de incrementos de información, los valores de la trayectoria secuencial y el límite más cercano con la corrección por monitorización no continua.

Corte	Fecha	Incrementos			Z'	Incremento de V	V	Corrección "Árbol de Navidad"	LÍMITE	FINAL
		E-P	N	Días						
1	1-dic-94	101	922	659	-6,95228	22,4840	22,4840	2,7644	0,7740	NO
2	1-jun-95	21	309	182	-4,14866	4,99325	27,4773	1,3027	1,1195	NO
3	27-nov-95	22	216	179	-4,45266	4,94016	32,4174	1,2958	0,0221	NO
4	20-may-96	31	281	175	-3,68981	6,90188	39,3193	1,5316	-1,7566	NO
5	4-feb-97	33	351	260	-2,54978	7,47820	46,7975	1,5943	-3,4909	STOP



hit any key to continue:

New Max U= 262.71

Stop the study and accept the null hypothesis

## RESULTADOS

### *Otras alternativas secuenciales*

Para otras alternativas, otros modelos secuenciales de límites continuos y dos ejemplos de diseños secuenciales basados en la función de gasto de alfa, se presentan el límite de finalización equivalente y la conclusión de finalización. Como se puede observar en las tablas y figuras siguientes, únicamente el caso del modelo RPST 4000 no se hubiera podido todavía concluir en la 5<sup>ta</sup> inspección.

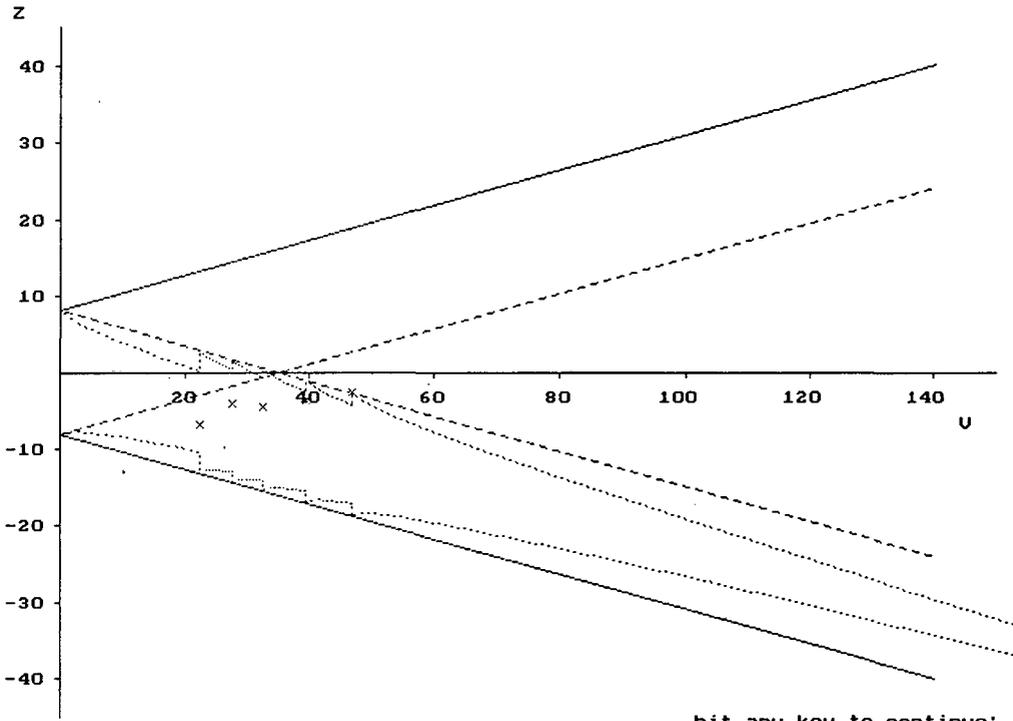
Corte	RPST4000		RPST4500		RPST5000		RPST6000	
	LIMITE	FINAL	LIMITE	FINAL	LIMITE	FINAL	LIMITE	FINAL
1	4.762	NO	1.611	NO	0,980	NO	0,502	NO
2	5,123	NO	1.966	NO	1,329	NO	0,841	NO
3	4,041	NO	0.877	NO	0,234	NO	-0,262	NO
4	2.283	NO	-0.889	NO	-1,540	NO	-2,049	NO
5	0,571	NO	-2.610	STOP	-3,269	STOP	-3,793	STOP

Corte	RPST7000		RPST10000		RPST		TRIANGULAR	
	LIMITE	FINAL	LIMITE	FINAL	LIMITE	FINAL	LIMITE	FINAL
1	0.316	NO	0.152	NO	0,110	NO	3.292	NO
2	0,649	NO	0,476	NO	0,430	NO	3,095	NO
3	-0.461	NO	-0.643	NO	-0.692	NO	1.461	NO
4	-2.257	NO	-2.452	NO	-2,506	NO	-1.068	NO
5	-4,010	STOP	-4.219	STOP	-4,278	STOP	-3.615	STOP

Corte	Parámetro 0.2		Parámetro 0.5	
	LIMITE	FINAL	LIMITE	FINAL
1	0	NO	-0.18569	NO
2	-0.13996	NO	-0.6047	NO
3	-0.26008	NO	-0.72332	NO
4	-0.86006	NO	-1.33505	NO
5	-2.783	STOP	-2.78444	STOP

# RESULTADOS

---

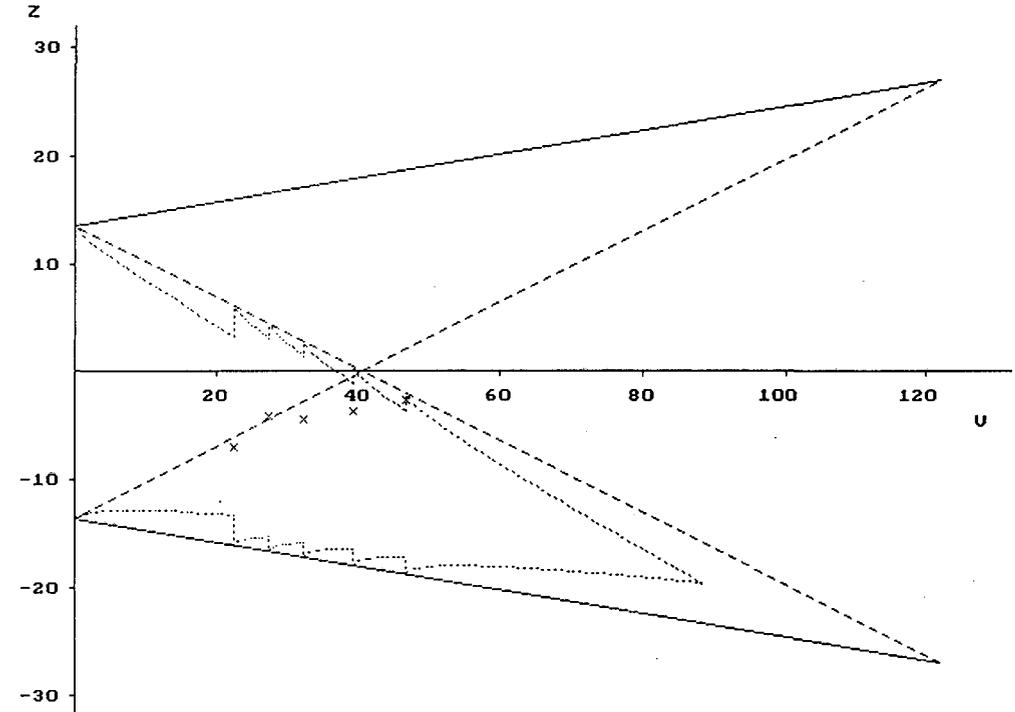


hit any key to continue:

New Max U= 236.69

Stop the study and accept the null hypothesis

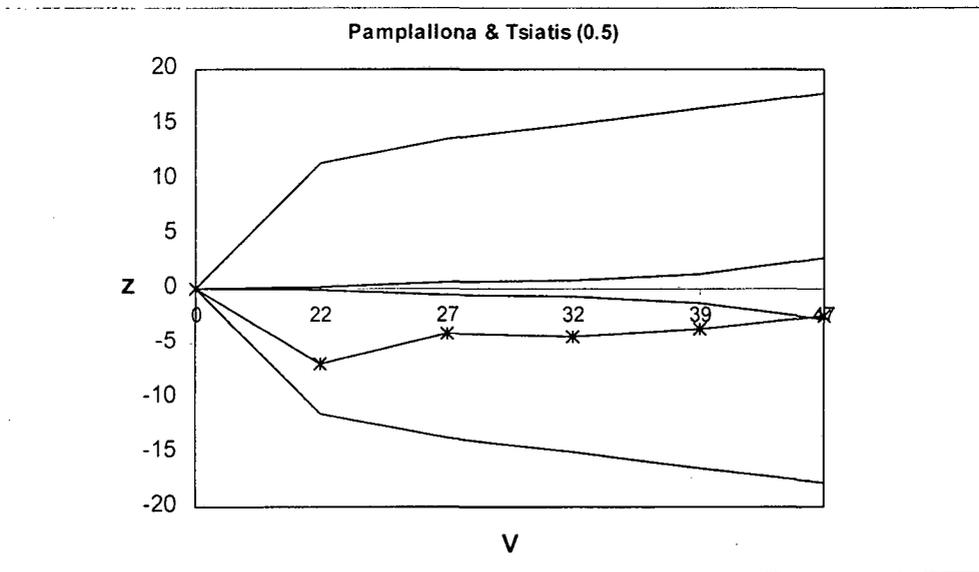
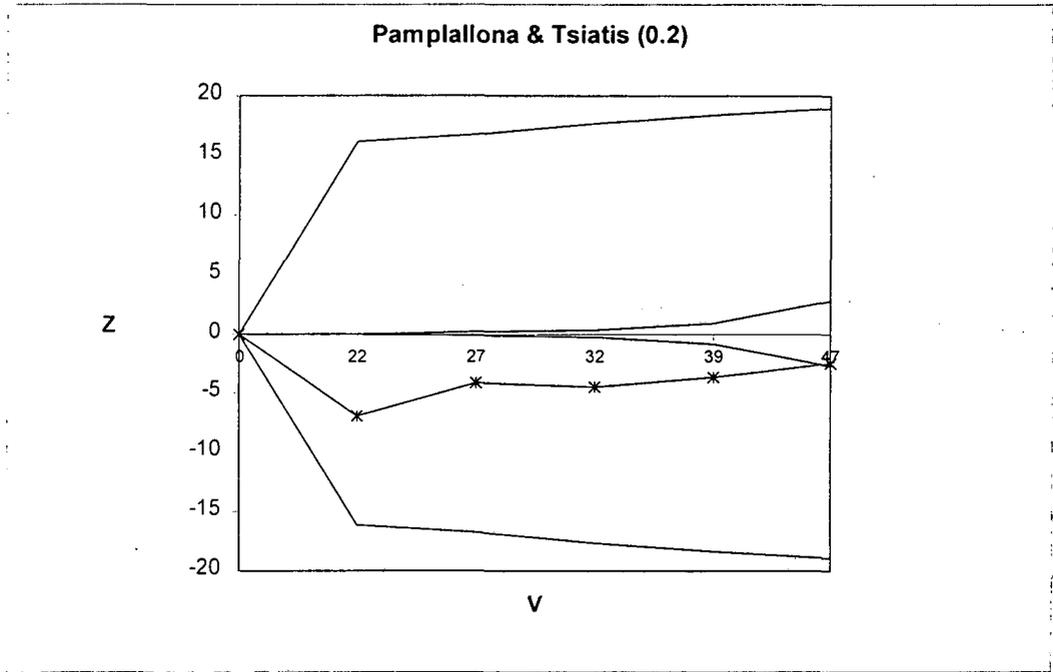
RESULTADOS



hit any key to continue:

New Max U= 88.55 Stop the study and accept the null hypothesis

# RESULTADOS



## POST- REVISIÓN

La reevaluación de los datos tras la revisión mostró unos resultados prácticamente superponibles a los previos (ver tabla siguiente):

Corte	Fecha	A			B			Total			A-B
		n	N	%	n	N	%	n	N	%	%
1	1-dic-94	53	448	11.83%	39	441	8.84%	92	889	10.35%	2.99%
2	1-jun-95	63	588	10.71%	53	589	9.00%	116	1177	9.87%	1.71%
3	27-nov-95	69	684	10.09%	61	681	8.96%	130	1365	9.52%	1.13%
4	20-may-96	85	818	10.39%	73	800	9.16%	158	1618	9.77%	1.27%
5	4-feb-97	105	1068	9.83%	99	1056	9.38%	204	2124	9.61%	0.46%

El porcentaje global de eventos ha seguido una línea descendiente, bajando del 10.35% en el primer corte hasta el 9.61% en el último. Las diferencias entre ambos grupos fueron disminuyendo: inicialmente eran del 2.99% (25.3% relativo) y al final sólo del 0.46% (4.6% relativo).

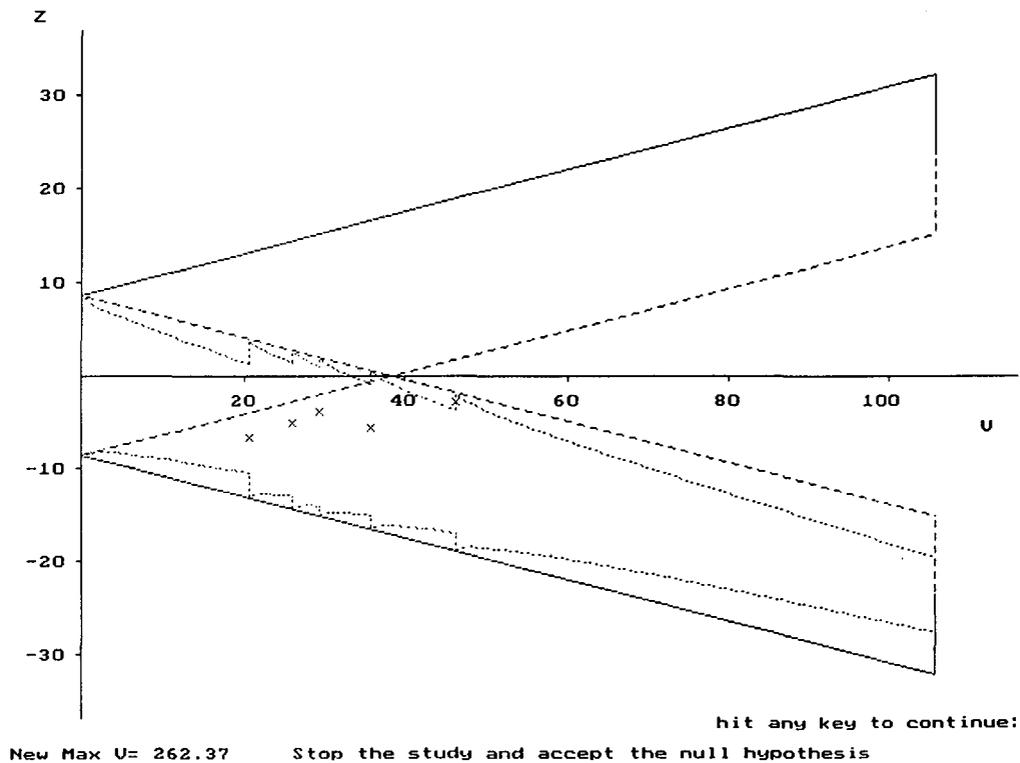
La principal diferencia radica en el tamaño muestral final. Aunque hubo algún caso considerado válido que al reevaluarse fue excluido, se observa un notable incremento producido por la inclusión de los casos en que los que no se había finalizado el seguimiento, y por lo tanto todavía no eran evaluables en la fecha de la última inspección.

Los incrementos de información de los distintos cortes secuenciales difieren poco a excepción de la última inspección. En este caso se observa un incremento del 39.86% en relación a los valores de V antes y después de la revisión, 7.48 y 10.46 respectivamente.

## RESULTADOS

De acuerdo con la nueva evaluación, el reclutamiento por mes fue en las 5 inspecciones: 40.47, 47.47, 31.51, 43.37 y 58.38

Corte	Fecha	Incrementos			Z'	Incremento de V	V	Corrección "Árbol de Navidad"	LÍMITE	FINAL
		E-P	N	Días						
1	1-dic-94	92	889	659	-6.638	20,6198	20,620	2,647	1,308	NO
2	1-jun-95	24	288	182	-5,049	5,52209	26,142	1,370	1,351	NO
3	27-nov-95	14	188	179	-3,857	3,26288	29,405	1,053	0,938	NO
4	20-may-96	28	253	175	-5,121	6,23801	35,643	1,456	-0,859	NO
5	4-feb-97	46	506	260	-2,424	10,45893	46,102	1,885	-3,627	STOP



## RESULTADOS

### *Otras alternativas secuenciales*

Los resultados son superponibles a los previos a la revisión y no hay ninguna diferencia en cuanto a los criterios de decisión de finalización. Todos los modelos estudiados finalizan en la 5 inspección a excepción del modelo RPST 4000.

Corte	RPST4000		RPST4500		RPST5000		RPST6000	
	LIMITE	FINAL	LIMITE	FINAL	LIMITE	FINAL	LIMITE	FINAL
0	12,485	NO	9,361	NO	8,756	NO	8,320	NO
1	5,290	NO	2,142	NO	1,512	NO	1,038	NO
2	5,350	NO	2,195	NO	1,559	NO	1,074	NO
3	4,948	NO	1,788	NO	1,149	NO	0,658	NO
4	3,169	NO	0,002	NO	-0,645	NO	-1,147	NO
5	0,434	NO	-2,747	STOP	-3,406	STOP	-3,927	STOP

Corte	RPST7000		RPST10000		RPST		TRIANGULAR	
	LIMITE	FINAL	LIMITE	FINAL	LIMITE	FINAL	LIMITE	FINAL
0	8,164	NO	8,041	NO	8,014	NO	13,526	NO
1	0,855	NO	0,694	NO	0,653	NO	4,029	NO
2	0,884	NO	0,713	NO	0,668	NO	3,472	NO
3	0,463	NO	0,286	NO	0,239	NO	2,705	NO
4	-1,350	NO	-1,539	NO	-1,590	NO	0,229	NO
5	-4,144	STOP	-4,352	STOP	-4,410	STOP	-3,675	STOP

Corte	Parámetro 0.2		Parámetro 0.5	
	LIMITE	FINAL	LIMITE	FINAL
0	0	NO	0	NO
1	0	NO	-0,13309	NO
2	-0,0861	NO	-0,51372	NO
3	-0,19537	NO	-0,53449	NO
4	-0,4337	NO	-1,09388	NO
5	-3,35852	STOP	-3,25524	STOP

## RESULTADOS

---

### ANÁLISIS

#### *Análisis principal*

No entraremos a comentar en detalle los resultados del análisis adicional con los datos previos a la revisión, pero se puede comprobar que son muy similares tanto la estimación del efecto como el valor de p del test de hipótesis.

El porcentaje de cambio de  $\theta$  entre antes y después de la revisión es de 3.2% (-0.129 y -0.125 respectivamente);

El valor de p varía de 0.582 a 0.549 (antes y después respectivamente).

De esta manera las conclusiones del estudio no difieren con la información aportada previa y posteriormente a la reevaluación de los pacientes e inclusión de los casos que se hallaban en seguimiento al dictar la finalización del periodo de inclusiones.

Los datos se describen en las siguientes tablas.

		A			B			Total			A-B
	Fecha	n	N	%	n	N	%	n	N	%	%
Bellaterra	4-feb-97	107	1044	10.25%	101	1035	9.76%	208	2079	10,00%	0,49%

Corte	Z	V
5	-2.550	46.797

Estimador	Estimación	SE (Est)	IC 95%	
			Inf	Sup
Theta mle ( $\hat{\theta}$ )	-0.054	0,146	-0.341	0,232
Theta mediana ( $\hat{\theta}_M$ )	-0.094	-	-0.460	0,208
Theta ajustada ( $\tilde{\theta}$ )	-0.129	0.166	-0.454	0,196

## RESULTADOS

---

Estimador	Estimación	IC 95%	
		Inf	Sup
OR (tetha mle)	0.947	0.711	1.261
OR (tetha mediana)	0.911	0.631	1.231
OR (tetha ajustada)	0.879	0.635	1.217

Los resultados de la estimación del efecto combinado se presentan en la siguiente tabla.

Tal como era esperable, la magnitud del efecto se aleja poco de 0 para  $\theta$  y de 1 para el OR, los cuales expresan la igualdad en una escala aditiva y multiplicativa respectivamente. Para evitar la infraestimación, al haber finalizado aceptando la hipótesis nula, tomaremos como estimador el estimador no sesgado de  $\theta$  ( $\tilde{\theta}$ ).

En cualquiera de los estimadores que utilizemos, el IC 95% contiene el valor de la igualdad confirmando la toma de decisión al cortar por los límites internos de la razón de continuación; el valor de p estimado fue de  $p = 0,582$ .

Estimador	Estimación	SE (Est)	IC 95%	
			Inf	Sup
Theta mle ( $\hat{\theta}$ )	-0.053	0.147	-0.341	0.236
Theta mediana ( $\theta_M$ )	-0.086	-	-0.456	0.216
Theta ajustada ( $\tilde{\theta}$ )	-0.125	0.168	-0.455	0.204

Estimador	Estimación	IC 95%	
		Inf	Sup
OR (tetha mle)	0.949	0.711	1.266
OR (tetha mediana)	0.917	0.634	1.241
OR (tetha ajustada)	0.882	0.634	1.227

*Análisis variables secundarias*

Se realizó un ajuste en las variables secundarias dada su correlación con la variable principal. En los resultados ajustados se observa un desplazamiento de los estimadores del efecto de las variables secundarias en el sentido de la principal. De esta manera hay un aumento del efecto en las variables que tienen el mismo sentido que la variable principal (AVC, Muerte) y una disminución de la variable que tenía el estimador en dirección opuesta al end-point (IAM).

Para los valores de  $p$ , los resultados lógicamente apuntan en el mismo sentido que para las estimaciones de  $\theta$ . A continuación se detallan en las tablas los resultados siguientes:

- Descripción de las variables y estimaciones de crudas al finalizar la monitorización
- Valores de  $\theta$ , y los parámetros necesarios para su cálculo,
- Estimación del efecto en forma de OR y de proporciones ajustados,
- Estimación de los valores de  $p$ .

Tabla xx: Resultados crudos al finalizar el estudio TIM con los datos post-revisión.

	A			B			Total			A-B
	N	N	%	n	N	%	n	N	%	%
<b>E-P</b>	105	1068	9.83%	99	1056	9.38%	204	2124	9.60%	0.46%
<b>IAM</b>	18	1068	1.69%	30	1056	2.84%	48	2124	2.26%	-1.16%
<b>AVC</b>	14	1068	1.31%	5	1056	0.47%	19	2124	0.89%	0.84%
<b>Muerte</b>	79	1068	7.40%	69	1056	6.53%	148	2124	6.97%	0.86%

## RESULTADOS

	Z	V	theta	Var(Th)	s2	r	W	sesgo	Var(Th ins)	Theta no sesgado		
			Z/V	1/V						Th ins	LI	LS
E-P	-2.424	46.102	-0.053	0.022	1,000	1,000	?	0,073	0,028	-0.125	-0,455	0,204
IAM	6.136	11.729	0.523	0.085	0,254	0,466	10,847	0,067	0,091	0.456	-0,135	1,047
AVC	-4.446	4.708	-0.945	0.212	0,102	0,291	4,294	0,066	0,218	-1.011	-1,926	-0,096
Muerte	-4.582	34.422	-0.133	0,029	0,747	0,840	33,446	0,071	0,035	-0,204	-0,572	0,164

	OR no sesgado			OR directo		
	OR	LI	LS	OR	LI	LS
E-P	0,882	0.634	1.227	0.949	0.711	1,266
IAM	1.577	0.873	2.848	1.687	0.952	2,990
AVC	0.364	0.146	0.908	0.389	0.158	0,960
Muerte	0.816	0.564	1.179	0.875	0.627	1,223

	A	B	p(dif)
E-P (~) Est	10,15%	9,06%	1.09%
E-P (~) LI	11.54%	7.64%	3.90%
E-P (~) LS	8.72%	10.49%	-1,77%
IAM (~) Est	1,77%	2,76%	-0,99%
IAM (~) LI	2.41%	2,11%	0,30%
IAM (~) LS	1.20%	3.34%	-2.14%
AVC (~) Est	1,31%	0,48%	0,83%
AVC (~) LI	1.55%	0.23%	1.32%
AVC (~) LS	0.94%	0.85%	0,09%
Muerte (~) Est	7,62%	6,31%	1,32%
Muerte (~) LI	8.77%	5,15%	3,62%
Muerte (~) LS	6.44%	7.50%	-1.06%

## RESULTADOS

---

### Variable principal

Valor de p valor ajustado				Valor de p crudo	
Directo		Test de Wald			
U	P	U	p	U	p
0,550466	0,582000	0,745528	0,455952	0,356973	0,721112

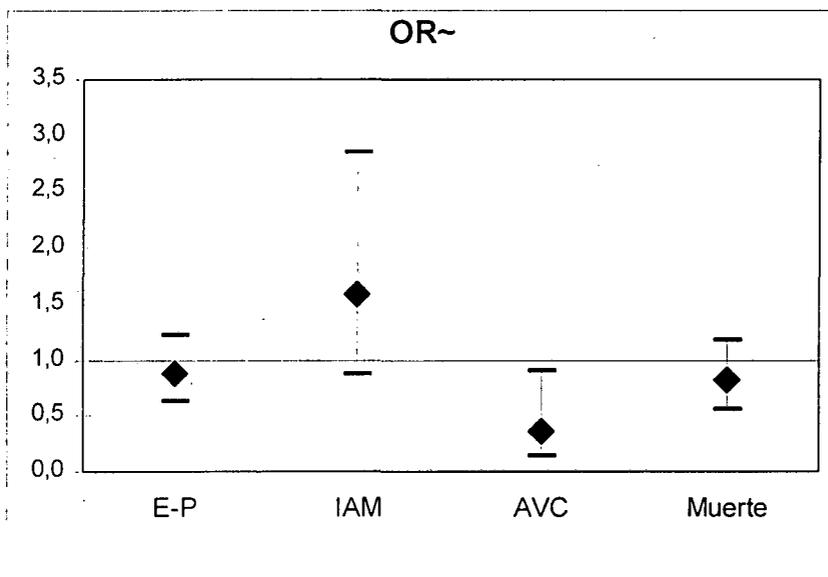
### Variabes secundarias

	Análisis ajustado (Test Wald)		Análisis crudo	
	U	p	U	p
IAM	1,5113184	0,1307074	1,791580	0,073200
AVC	2,1656131	0,0303407	2,049385	0,040424
Muerte	1,0852928	0,2777921	0,780974	0,434818

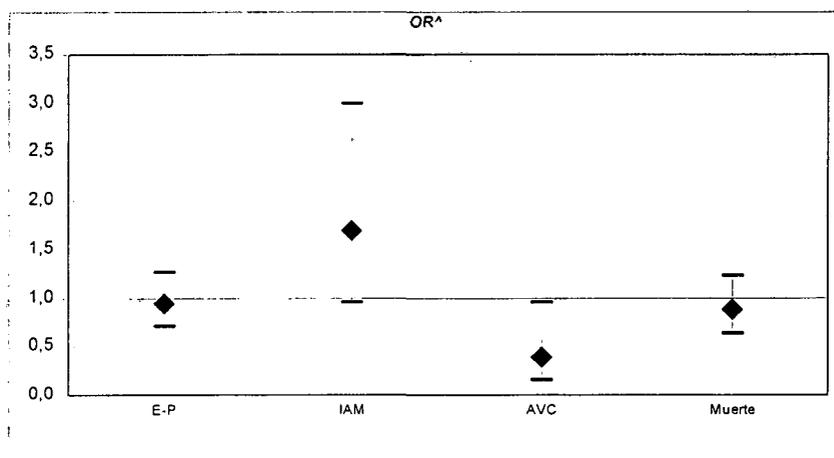
## RESULTADOS

---

Representación de los ORs ajustados por el análisis secuencial:



Representación de los ORs crudos (sin ajuste por la metodología secuencial):



*Análisis pre-revisión*

No entraremos a comentar en detalle los resultados del análisis adicional con los datos previos a la revisión, pero se puede comprobar que son muy similares tanto la estimación del efecto como el valor de p del test de hipótesis.

El porcentaje de cambio de  $\theta$  entre antes y después de la revisión es de 3.2% (-0.129 y -0.125 respectivamente);

El valor de p varía de 0.582 a 0.549 (antes y después respectivamente).

De esta manera las conclusiones del estudio no difieren con la información aportada previa y posteriormente a la reevaluación de los pacientes e inclusión de los casos que se hallaban en seguimiento al dictar la finalización del periodo de inclusiones.

Los datos se describen en las siguientes tablas.

	Fecha	A			B			Total			A-B
		n	N	%	n	N	%	n	N	%	%
Bellaterra	4-feb-97	107	1044	10.25%	101	1035	9,76%	208	2079	10,00%	0,49%

Corte	Z	V
5	-2.550	46.797

Estimador	Estimación	SE (Est)	IC 95%	
			Inf	Sup
Theta mle ( $\hat{\theta}$ )	-0.054	0.146	-0,341	0,232
Theta mediana ( $\theta_M$ )	-0.094	-	-0,460	0,208
Theta ajustada ( $\tilde{\theta}$ )	-0.129	0.166	-0,454	0,196

## RESULTADOS

---

Estimador	Estimación	IC 95%	
		Inf	Sup
OR (tetha mle)	0.947	0,711	1,261
OR (tetha mediana)	0.911	0,631	1,231
OR (tetha ajustada)	0.879	0,635	1,217

### DISCUSIÓN

#### *Planificación del diseño*

#### *Especificaciones del diseño*

#### **Selección de la variable principal**

En el estudio TIM se estableció de acuerdo a la información de ensayos clínicos publicados en patología cardiovascular {2341} y al consenso de grupos de estudio de la enfermedad cardiovascular {2376}.

De todos es sabido que el establecimiento de variables compuestas facilita la realización de los estudios al aumentar los eventos y disminuir el tamaño de muestra que se requiere. No obstante, al estudiar el efecto comparado entre tratamientos, podemos encontrarnos en distintas situaciones de magnitud y sentido en las variables que forman la variable compuesta:

Efectos de similar intensidad y del mismo sentido

Efectos de distinta intensidad distinta pero con el mismo sentido

Efectos de distinto sentido

Los casos a) y b) no suponen ningún problema al interpretar los resultados y presuponen una misma base fisiopatológica. En cambio, en caso de encontrarse con la tercera opción (c) las conclusiones acerca del estudio se complican, sobretodo cuando el efecto es muy marcado. Si una de esas variables es una variable dura como la mortalidad, entonces el peso clínico de dicha variable puede decantar las conclusiones. Si las variables son de similar importancia entonces difícilmente se podrá llegar a una conclusión clara. De esta manera, si nos encontramos con c)

debemos suponer que quizás los mecanismos biológicos implicados no son idénticos y que para futuros estudios es recomendable una profunda reevaluación a la hora de la selección de la variable principal. En algún ensayo clínico reciente el efecto fue para las mismas variables fue distinto en función de estratos de patología de inclusión {2377}.

### **Diferencias clínicamente relevantes**

A partir de los resultados publicados del ensayo clínico ISIS-2 {2341} en 1992, se delimita una magnitud en el efecto comparado de la aspirina con placebo. Dado que aspirina es superior a placebo en la prevención de eventos cardiovasculares en pacientes con sospecha de infarto, en cualquier nuevo ensayo el tratamiento experimental no puede compararse con placebo y el grupo de referencia debe ser aspirina.

En este caso, el límite aceptable de diferencia clínicamente relevante frente aspirina debe estar, en sentido negativo (favorable a aspirina), lo suficientemente lejos del valor esperado para el grupo placebo. En caso contrario si concluyera el estudio aceptando la hipótesis nula pero cerca de dicho límite, se estaría aceptando absurdamente igualdad con aspirina para un tratamiento de eficacia similar al placebo; los problemas éticos de aceptar dicha afirmación se remarcan por la gravedad de la patología estudiada.

Así para la variable estudiada la aceptación del límite negativo de superioridad, es decir, inferioridad, se tuvo en cuenta la diferencia esperada con placebo (28%) estableciéndose finalmente en un 25%.

La variabilidad depende de la incidencia de la variable principal para los dos grupos de tratamiento. En el estudio TIM se estimó la variabilidad basal en un 10% de acuerdo al estudio ISIS-2.

### **Diferencias esperadas**

Por datos de ensayos clínicos de triflusal {2375,2379} y de otros antiagregantes {2378} comparados con aspirina, aunque no era del todo descartable, en principio no eran esperables diferencias de gran magnitud. Así lo más operativo fue establecer unas diferencias relevantes que si eran superadas se aceptaría superioridad de uno de los dos tratamientos, asumiendo que existía una alta probabilidad de finalizar aceptando la hipótesis nula.

### **Protección frente al error de tipo I**

Aunque el objetivo inmediato de un ensayo clínico en fase III es la evaluación de la eficacia de una estrategia terapéutica, no debemos dejar de lado que en el fondo se intenta probar si dicho medicamento ofrecerá algún beneficio en el momento de exponer a éste la población general. Entonces es lógico pensar que a los futuros sujetos diana del fármaco evaluado se les debe proteger de estrategias ineficaces y riesgos derivados del uso de los medicamentos. Dicho nivel de protección se establece en un estudio concreto limitando el error de tipo I hasta un nivel aceptable, habitualmente un 5%.

A pesar de que los estudios a una sola cola no aumentan el riesgo para los futuros pacientes, su aplicación esta sujeta a controversias. Si se formulan correctamente las hipótesis, se puede descartar tratamientos ineficaces de tal manera de tal manera que sean aceptados sólo aquellos que sean al menos tan eficaces como el grupo de referencia, y manteniendo el error  $\alpha$  en 0,05.

La desventaja en todo caso será para el promotor: no habrá podido demostrar ser mejor y aunque pueda concluir que no es peor que el fármaco control, se encontrará con las dificultades de conquistar un mercado de pautas establecidas y más conocidas. Este hecho puede ser más notable si, además, no puede aportar ventajas claras alternativas.

La ventaja de estos estudios es que requieren menor tamaño muestral, en el caso habitual de  $\beta=0.2$  y  $\alpha=0.05$  la única diferencia en el cálculo es  $f(U_\alpha+U_\beta)^2$ :

$$\text{Para pruebas a una cola, } f(U_\alpha+U_\beta)^2 = 6,183$$

$$\text{Para pruebas a dos colas, } f(U_\alpha+U_\beta)^2 = 7,849$$

En un hipotético estudio como el TIM con un diseño convencional a una sola cola se hubiera requerido unos 840 pacientes menos (21% de reducción en el tamaño muestral):

$$6,183 / 7,849 = 0,79$$

$$4000 * 0,79 = 3160$$

En el caso de estudio TIM nos inclinamos evitar los problemas de los tests unilaterales y se estudiaron distintas opciones pero con el requisito de poder hacer tests bilaterales. Todos los grandes estudios se han realizado a dos colas {2378}.

### **Protección frente al error de tipo II**

En un estudio convencional los requerimientos de poder ( $1-\beta$ ) se acostumbran a establecer en alrededor de 0,8 (80%). La crítica que se puede hacer a estudios con poder bajo es evidente, pero todavía más relevante cuando el grupo de control es un fármaco activo. En estos casos, la posibilidad de demostrar que el fármaco experimental es realmente peor que el control, cuando en realidad así sea, será muy baja; como consecuencia, el riesgo de someter a futuros pacientes a tratamientos no

eficaces es alto. En el caso de comparación frente a placebo el problema es solamente para el promotor, puesto a diferencia del caso anterior, difícilmente se aceptarían para registro estudios negativos frente a placebo. Así, parece claro que no se deban diseñar estudios con un poder bajo y que en caso de sospechar cambios por aumento de la varianza o disminución del reclutamiento (en estudio de supervivencia) se recomiende proceder a una reevaluación del tamaño muestral. Es importante que dicha decisión no este motivada por la observación de las diferencias entre grupos y que siempre se realice de forma enmascarada.

Teniendo en cuenta estos puntos, la justificación de utilizar un poder no inferior al utilizado normalmente (80%) en el estudio TIM está clara. Dado que a priori las diferencias esperadas eran inferiores a las consideradas clínicamente relevantes, tal y como finalmente se ha podido constatar, no habría sido tolerable diseñar este estudio con un poder inferior. Además, al terminar aceptando la hipótesis nula, los resultados ofrecerían dudas razonables respecto si a uno de los tratamientos era superior que el otro. Indudablemente, métodos que se ajusten en función de la variabilidad de tal manera que el poder no se modifique (ej.: métodos secuenciales basados en límites continuos) son de gran utilidad y una posible alternativa a la reestimación del cálculo muestral. Así en el estudio TIM por el diseño utilizado el poder es mantenido en el 80% para las diferencias preestablecidas en la escala de  $\theta$  ( $\theta_R=0,3151 \Rightarrow OR=0,730$ ).

### *Selección del modelo óptimo basándose en las especificaciones teóricas*

#### **Aplicabilidad de los métodos**

test a dós colas

tamaño muestral definitivo

posibilidad de finalización prematura bajo H0 y H1

protección frente a los errores estadísticos

Logística y programas que lo apliquen

Reconocimiento por la comunidad científica

Ya se ha discutido previamente la importancia de poderse realizar test a dos colas. Aquellos modelos que sólo se pudieran realizar pruebas unilaterales se han desestimado.

El siguiente criterio que se usó para discriminar que diseños se podrían aplicar al estudio TIM es el tamaño muestral necesario para llevarse a cabo. Un diseño clásico requería unos 4000 pacientes con las especificaciones del estudio TIM. Ese tamaño, aunque esta en el rango de los estudios en patología cardiovascular {2378}, no deja de tener una envergadura tal que no sea justificable buscar alternativas.

Las distintas aproximaciones a la equivalencia terapéutica no son una solución porque el tamaño muestral en ninguna de las opciones válidas estudiadas, (test doblemente unilateral de Schuirman {1291}, intervalos de confianza, Hauck y Anderson {2293}) optimiza el tamaño muestral {2294,1303,1942,2270,2287,2292}.

Cuando se estudia el efecto comparado de aspirina versus placebo en la protección de eventos cardiovasculares observando los datos de los distintos ensayos clínicos publicados, vemos que las diferencias son sólo de relativa magnitud. Si nos fijamos en los estudios de comparación de un antiagregante contra aspirina, entonces el

## DISCUSIÓN

---

efecto es menor y para encontrar diferencias de dicha magnitud hace falta tamaños muestrales muy grandes {2378}. Además, como la mayoría de estos estudios no han demostrado superioridad, las previsiones de terminar aceptando la hipótesis nula en estudio como el TIM parecieron bastante razonables. Así sería interesante que el diseño seleccionado pueda optimizar el tamaño muestral desde este punto de vista.

Los métodos secuenciales por grupos {2326,536,226,2267,844,549,2031} si que son válidos cuando el objetivo es terminar antes si hay gran certeza para rechazar la hipótesis nula; como fue comentada previamente no optimizaran el tamaño muestral en caso contrario. Alguno de estos estudios utilizó este método como el GUSTO-IIb con 12000 pacientes, pero el objetivo es únicamente hacer alguna inspección, muy penalizada en términos de  $\alpha$  para no afectar a la última inspección, para obtener datos previos sin intención o pensamiento claro de finalización prematura.

El problema principal de los métodos secuenciales de intervalos de confianza repetidos {2258} y restricción estocástica {2304} es su aplicabilidad a los ensayos clínicos de demostración de eficacia. No hemos encontrado ningún ensayo clínico en patología cardiovascular donde se haya aplicado esta metodología.

La adaptación de los métodos bayesianos parece más plausible que los métodos anteriores, pero todavía falta por desarrollar los programas que sirvan de herramientas para su instauración. No obstante, hay cierta controversia y en el caso del estudio TIM los métodos que se hubieran podido aplicar se publicaran con posterioridad {2305,1974}.

En los últimos años ha habido gran número de trabajos que han explorado e implementado la metodología secuencial. Al igual que ha sucedido con los métodos bayesianos, los métodos secuenciales de límites continuos para equivalencia terapéutica {1925} y la aplicación concreta de los diseños de gasto de alfa mediante desarrollo de software {2300}, se han sucedido con posterioridad al inicio de nuestro estudio. No se ha encontrado diseños que utilicen el método de Whitehead

## DISCUSIÓN

---

seguramente por el poco tiempo transcurrido desde su publicación. Estudios basados en funciones con gasto de alfa para terminación prematura de  $H_0$  se han realizado muchos, ya que de hecho, estos métodos son una extensión de los diseños secuenciales por grupos.

Finalmente, la metodología utilizada inicialmente para la exploración a fondo mediante simulaciones fue la de los diseños de límites continuos. En nuestro caso se descartaron los métodos Restricted Procedure (RP) y Open Top Design (OTD) por no haber probabilidad de terminar prematuramente bajo la hipótesis nula, y el segundo, además, por ser unilateral. Una posible e interesante aplicación de estos dos modelos secuenciales es la monitorización de acontecimientos adversos.

En un estudio en que se estudie la eficacia de dos medicamentos, incluso con metodología convencional, se puede y es aconsejable realizar en algunos casos la monitorización de acontecimientos adversos de una manera que permita finalizar con criterios objetivos, y con posibilidad de realizar una estimación no sesgada por las inspecciones; los modelos RP y OTD son ideales para este objetivo.

Si los dos medicamentos tienen potencial desconocido, o conocido y grave, de producir acontecimientos adversos, nos inclinaríamos por el modelo RP que permite concluir en caso de haber una mayor incidencia en cualquiera de los dos grupos. Si el objetivo fuera descartar un nuevo tratamiento en frente de un control conocido y seguro, seleccionaríamos el diseño OTD.

## DISCUSIÓN

Comprobación mediante simulaciones de los modelos seleccionados

### Identificación de un error de el diseño

Es aconsejable realizar simulaciones cuando se diseña un estudio, sobretodo con métodos alternativos y cuando se utilizan programas de ordenador nuevos. En nuestro caso cuando inicialmente se diseñó el estudio TIM, se utilizó el programa PEST2 {2380} el cual incorporaba respecto a la anterior los modelos RPST truncados. Ante las estimaciones de tamaño muestral tan buenas, la decisión fue clara en escoger el modelo RPST truncando a 5305, e incluso se pudo asumir el diseño con un poder del 90%.

Al realizar las simulaciones de comprobación, las estimaciones que a priori eran tan optimistas no pudieron ser confirmadas. Afortunadamente la inmediata salida de la siguiente versión (PEST3 {2299}), permitió confrontar el error y corregir el diseño. El error se debía a un cambio de signo en la versión 2 que restringía el tamaño de la región de continuación en lugar de aumentarlo. En la siguiente tabla se muestran las estimaciones teóricas para el modelo RPST 5305 calculadas con la versión 2 con un 90% de poder y con la versión 3 del programa con un 80% y 90% de poder:

Diferencias ( $\theta$ )	$p(\text{exp})$	$p(\text{con})$	PEST2		PEST3			
			Poder 90%		Poder 90%		Poder 80%	
			$P50(t)t$	$P90(t)t$	$P50(t)t$	$P90(t)t$	$P50(t)$	$P90(t)$
+0 * $\theta_R$	0,100	0,100	1557	2457	4706	4706	2091	3859
+0,25* $\theta_R$	0,093	0,100	1610	2605	4855	4855	2243	4765
+0.5 * $\theta_R$	0,087	0,100	1649	2804	5005	5005	2591	5005
+0,75* $\theta_R$	0,081	0,100	1607	2885	5155	5155	2855	5155
+1 * $\theta_R$	0,075	0,100	1452	2773	5305	5305	2503	5305
+1,25* $\theta_R$	0,070	0,100	1234	2475	4375	5454	1949	4861
+1,5 * $\theta_R$	0,065	0,100	1026	2090	3297	5603	1525	3432

## DISCUSIÓN

---

Se observa como el modelo inicial tenía mejores expectativas incluso para el modelo equivalente con poder inferior (80%). Los aumentos en requerimientos muestrales fueron:

Bajo  $H_0$  ( $\theta=0$ ): 514 pacientes más (+32,6%)

Bajo  $H_1$  ( $\theta=\theta_R$ ): 1051 pacientes más (+72,4%)

Fruto de estos nuevos cálculos se procedió a estudiar de nuevo todos los modelos para comprobar que opciones había tras la versión nueva con las correcciones para los modelos RPST truncados.

### ***DISEÑO, MONITORIZACIÓN Y ANÁLISIS ESTADÍSTICO DEL ESTUDIO TIM***

#### ***DISEÑO***

El estudio TIM se diseñó usando el modelo secuencial de límites continuos RPST truncando al valor de 5305 pacientes. Las características estadísticas seleccionadas fueron  $\alpha=0.05$  bilateral y  $\beta=0.2$ . La diferencia clínicamente relevante se estableció en un 2.5% (porcentaje absoluto) para un valor de base para el tratamiento experimental del 10%; ( $\theta_R=0.3151$  y  $OR = 0.730$ ).

Los detalles generales del método secuencial RPST truncado están ampliamente descritos en introducción. Las características específicas de CO, expectativas de tamaño muestral y ahorro de pacientes para este modelo, se desarrollan en los apartados anteriores de resultados.

#### ***MONITORIZACIÓN***

##### **PRE-REVISIÓN OFICIAL**

Tras aproximadamente dos años de inclusiones se decidió realizar el primer corte secuencial. Las fechas concretas de realización de las distintas inspecciones, así como el número de pacientes con end-point y los porcentajes se especifican en la siguiente tabla.

## DISCUSIÓN

---

Llama la atención que, aunque el porcentaje global de eventos se ha mantenido más o menos constante con leves oscilaciones alrededor del 10%, las diferencias entre ambos grupos que inicialmente eran relativamente grandes (3.02% absoluto, 24.26% relativo) han ido paulatinamente evolucionando hacia la igualdad (0.49% absoluto, 4.78 relativo).

Los incrementos de información en términos de eventos, pacientes y tiempo han sido similares a los previstos y de similar magnitud a partir de la primera inspección. Recordamos que ésta se preestableció realizarla a las 1000 inclusiones. No obstante, el penúltimo y último corte se realizaron con más información, sobretudo en unidades de V: entre un 41% y un 53% más que en los cortes 2 y 3. El reclutamiento por mes ha ido variando: 41.97, 50.93, 36.20, 48.17 y 40.50 pacientes mes respectivamente en las inspecciones 1 a 5.

En la tabla siguiente se presentan los distintos tipos de incrementos de información, los valores de la trayectoria secuencial y el límite más cercano con la corrección por monitorización no continua.

Corte	Fecha	Incrementos			Z'	Incremento de V	V	Corrección "Árbol de Navidad"	LÍMITE	FINAL
		E-P	N	Días						
1	1-dic-94	101	922	659	-6.95228	22,4840	22,4840	2,7644	0,7740	NO
2	1-jun-95	21	309	182	-4.14866	4,99325	27,4773	1,3027	1,1195	NO
3	27-nov-95	22	216	179	-4.45266	4,94016	32,4174	1,2958	0,0221	NO
4	20-may-96	31	281	175	-3.68981	6,90188	39,3193	1,5316	-1,7566	NO
5	4-feb-97	33	351	260	-2.54978	7,47820	46,7975	1,5943	-3,4909	STOP

## DISCUSIÓN

---

### *Otras alternativas secuenciales*

Para otras alternativas, otros modelos secuenciales de límites continuos y dos ejemplos de diseños secuenciales basados en la función de gasto de alfa, se presentan el límite de finalización equivalente y la conclusión de finalización. Como se puede observar en las tablas y figuras siguientes, únicamente el caso del modelo RPST 4000 no se hubiera podido todavía concluir en la 5<sup>ta</sup> inspección.

Corte	RPST4000		RPST4500		RPST5000		RPST6000	
	LIMITE	FINAL	LIMITE	FINAL	LIMITE	FINAL	LIMITE	FINAL
1	4.762	NO	1.611	NO	0,980	NO	0,502	NO
2	5.123	NO	1.966	NO	1,329	NO	0,841	NO
3	4.041	NO	0.877	NO	0,234	NO	-0,262	NO
4	2.283	NO	-0.889	NO	-1,540	NO	-2,049	NO
5	0.571	NO	-2.610	STOP	-3.269	STOP	-3,793	STOP

Corte	RPST7000		RPST10000		RPST		TRIANGULAR	
	LIMITE	FINAL	LIMITE	FINAL	LIMITE	FINAL	LIMITE	FINAL
1	0.316	NO	0.152	NO	0,110	NO	3.292	NO
2	0.649	NO	0.476	NO	0,430	NO	3.095	NO
3	-0.461	NO	-0.643	NO	-0,692	NO	1.461	NO
4	-2.257	NO	-2.452	NO	-2.506	NO	-1,068	NO
5	-4.010	STOP	-4.219	STOP	-4,278	STOP	-3,615	STOP

Corte	Parámetro 0.2		Parámetro 0.5	
	LIMITE	FINAL	LIMITE	FINAL
1	0	NO	-0.18569	NO
2	-0.13996	NO	-0.6047	NO
3	-0.26008	NO	-0.72332	NO
4	-0.86006	NO	-1.33505	NO
5	-2.783	STOP	-2.78444	STOP

## POST- REVISIÓN

La reevaluación de los datos tras la revisión mostró unos resultados prácticamente superponibles a los previos (ver tabla siguiente):

Corte	Fecha	A			B			Total			A-B
		n	N	%	n	N	%	n	N	%	%
1	1-dic-94	53	448	11,83%	39	441	8,84%	92	889	10,35%	2,99%
2	1-jun-95	63	588	10,71%	53	589	9,00%	116	1177	9,87%	1,71%
3	27-nov-95	69	684	10,09%	61	681	8,96%	130	1365	9,52%	1,13%
4	20-may-96	85	818	10,39%	73	800	9,16%	158	1618	9,77%	1,27%
5	4-feb-97	105	1068	9,83%	99	1056	9,38%	204	2124	9,61%	0,46%

El porcentaje global de eventos ha seguido una línea descendente, bajando del 10.35% en el primer corte hasta el 9.61% en el último. Las diferencias entre ambos grupos fueron disminuyendo: inicialmente eran del 2.99% (25.3% relativo) y al final sólo del 0.46% (4.6% relativo).

La principal diferencia radica en el tamaño muestral final. Aunque hubo algún caso considerado válido que al reevaluarse fue excluido, se observa un notable incremento producido por la inclusión de los casos en que los que no se había finalizado el seguimiento, y por lo tanto todavía no eran evaluables en la fecha de la última inspección.

## DISCUSIÓN

---

Los incrementos de información de los distintos cortes secuenciales difieren poco a excepción de la última inspección. En este caso se observa un incremento del 39.86% en relación a los valores de V antes y después de la revisión, 7.48 y 10.46 respectivamente.

De acuerdo con la nueva evaluación, el reclutamiento por mes fue en las 5 inspecciones: 40.47, 47.47, 31.51, 43.37 y 58.38

Corte	Fecha	Incrementos			Z'	Incremento de V	V	Corrección "Árbol de Navidad"	LÍMITE	FINAL
		E-P	N	Días						
1	1-dic-94	92	889	659	-6.638	20,6198	20,620	2,647	1,308	NO
2	1-jun-95	24	288	182	-5.049	5,52209	26,142	1,370	1,351	NO
3	27-nov-95	14	188	179	-3.857	3,26288	29,405	1,053	0,938	NO
4	20-may-96	28	253	175	-5.121	6,23801	35,643	1,456	-0,859	NO
5	4-feb-97	46	506	260	-2.424	10,45893	46,102	1,885	-3,627	STOP

## DISCUSIÓN

---

### *Otras alternativas secuenciales*

Los resultados son superponibles a los previos a la revisión y no hay ninguna diferencia en cuanto a los criterios de decisión de finalización. Todos los modelos estudiados finalizan en la 5 inspección a excepción del modelo RPST 4000.

Corte	RPST4000		RPST4500		RPST5000		RPST6000	
	LIMITE	FINAL	LIMITE	FINAL	LIMITE	FINAL	LIMITE	FINAL
0	12.485	NO	9.361	NO	8,756	NO	8,320	NO
1	5.290	NO	2.142	NO	1,512	NO	1,038	NO
2	5.350	NO	2.195	NO	1,559	NO	1,074	NO
3	4.948	NO	1.788	NO	1,149	NO	0,658	NO
4	3.169	NO	0.002	NO	-0,645	NO	-1,147	NO
5	0,434	NO	-2.747	STOP	-3,406	STOP	-3,927	STOP

Corte	RPST7000		RPST10000		RPST		TRIANGULAR	
	LIMITE	FINAL	LIMITE	FINAL	LIMITE	FINAL	LIMITE	FINAL
0	8,164	NO	8,041	NO	8,014	NO	13,526	NO
1	0,855	NO	0,694	NO	0,653	NO	4,029	NO
2	0,884	NO	0,713	NO	0,668	NO	3,472	NO
3	0,463	NO	0,286	NO	0,239	NO	2,705	NO
4	-1,350	NO	-1,539	NO	-1,590	NO	0,229	NO
5	-4,144	STOP	-4,352	STOP	-4,410	STOP	-3,675	STOP

Corte	Parámetro 0.2		Parámetro 0.5	
	LIMITE	FINAL	LIMITE	FINAL
0	0	NO	0	NO
1	0	NO	-0.13309	NO
2	-0.0861	NO	-0.51372	NO
3	-0.19537	NO	-0.53449	NO
4	-0.4337	NO	-1.09388	NO

## DISCUSIÓN

---

5	-3,35852	STOP	-3.25524	STOP
---	----------	------	----------	------

## ANÁLISIS

### *Análisis principal*

No entraremos a comentar en detalle los resultados del análisis adicional con los datos previos a la revisión, pero se puede comprobar que son muy similares tanto la estimación del efecto como el valor de p del test de hipótesis.

El porcentaje de cambio de  $\theta$  entre antes y después de la revisión es de 3.2% (-0.129 y -0.125 respectivamente);

El valor de p varía de 0.582 a 0.549 (antes y después respectivamente).

De esta manera las conclusiones del estudio no difieren con la información aportada previa y posteriormente a la reevaluación de los pacientes e inclusión de los casos que se hallaban en seguimiento al dictar la finalización del periodo de inclusiones.

Los datos se describen en las siguientes tablas.

		A			B			Total			A-B
	Fecha	n	N	%	n	N	%	n	N	%	%
<b>Bellaterra</b>	4-feb-97	107	1044	10.25%	101	1035	9.76%	208	2079	10,00%	0,49%

Corte	Z	V
5	-2,550	46,797

## DISCUSIÓN

Estimador	Estimación	SE (Est)	IC 95%	
			Inf	Sup
Theta mle ( $\hat{\theta}$ )	-0.054	0,146	-0,341	0,232
Theta mediana ( $\theta_M$ )	-0.094	-	-0,460	0,208
Theta ajustada ( $\tilde{\theta}$ )	-0.129	0.166	-0,454	0,196

Estimador	Estimación	IC 95%	
		Inf	Sup
OR (tetha mle)	0.947	0.711	1,261
OR (tetha mediana)	0.911	0.631	1,231
OR (tetha ajustada)	0.879	0.635	1,217

Los resultados de la estimación del efecto combinado se presentan en la siguiente tabla.

Tal como era esperable, la magnitud del efecto se aleja poco de 0 para  $\theta$  y de 1 para el OR, los cuales expresan la igualdad en una escala aditiva y multiplicativa respectivamente. Para evitar la infraestimación, al haber finalizado aceptando la hipótesis nula, tomaremos como estimador el estimador no sesgado de  $\theta$  ( $\tilde{\theta}$ ).

En cualquiera de los estimadores que utilizemos, el IC 95% contiene el valor de la igualdad confirmando la toma de decisión al cortar por los límites internos de la razón de continuación; el valor de p estimado fue de  $p = 0,582$ .

Estimador	Estimación	SE (Est)	IC 95%	
			Inf	Sup
Theta mle ( $\hat{\theta}$ )	-0.053	0.147	-0,341	0,236
Theta mediana ( $\theta_M$ )	-0.086	-	-0,456	0,216
Theta ajustada ( $\tilde{\theta}$ )	-0.125	0.168	-0,455	0,204

Estimador	Estimación	IC 95%	
		Inf	Sup
OR (tetha mle)	0.949	0.711	1,266
OR (tetha mediana)	0.917	0.634	1,241
OR (tetha ajustada)	0.882	0,634	1,227

*Análisis variables secundarias*

Se realizó un ajuste en las variables secundarias dada su correlación con la variable principal. En los resultados ajustados se observa un desplazamiento de los estimadores del efecto de las variables secundarias en el sentido de la principal. De esta manera hay un aumento del efecto en las variables que tienen el mismo sentido que la variable principal (AVC, Muerte) y una disminución de la variable que tenía el estimador en dirección opuesta al end-point (IAM).

Para los valores de  $p$ , los resultados lógicamente apuntan en el mismo sentido que para las estimaciones de  $\theta$ . A continuación se detallan en las tablas los resultados siguientes:

Descripción de las variables y estimaciones de crudas al finalizar la monitorización

Valores de  $\theta$ , y los parámetros necesarios para su cálculo,

Estimación del efecto en forma de OR y de proporciones ajustados,

Estimación de los valores de  $p$ .

Tabla xx: Resultados crudos al finalizar el estudio TIM con los datos post-revisión.

	A			B			Total			A-B
	N	N	%	n	N	%	n	N	%	%
E-P	105	1068	9.83%	99	1056	9.38%	204	2124	9,60%	0,46%
IAM	18	1068	1.69%	30	1056	2.84%	48	2124	2,26%	-1.16%
AVC	14	1068	1.31%	5	1056	0.47%	19	2124	0,89%	0,84%
Muerte	79	1068	7.40%	69	1056	6.53%	148	2124	6,97%	0,86%

	Z	V	theta	Var(Th)	s2	r	W	sesgo	Var(Th ins)	Theta no sesgado		
			Z/V	1/V						Th ins	LI	LS
E-P	-2.424	46.102	-0.053	0.022	1,000	1,000	?	0,073	0,028	-0.125	-0,455	0,204
IAM	6.136	11.729	0.523	0.085	0,254	0,466	10,847	0,067	0,091	0.456	-0,135	1,047
AVC	-4.446	4.708	-0.945	0.212	0,102	0,291	4,294	0,066	0,218	-1,011	-1,926	-0,096
Muerte	-4.582	34.422	-0.133	0.029	0,747	0,840	33,446	0,071	0,035	-0.204	-0,572	0,164

## DISCUSIÓN

---

	OR no sesgado			OR directo		
	OR	LI	LS	OR	LI	LS
E-P	0.882	0.634	1.227	0.949	0.711	1.266
IAM	1.577	0.873	2.848	1.687	0.952	2.990
AVC	0.364	0.146	0.908	0.389	0.158	0.960
Muerte	0.816	0.564	1.179	0.875	0.627	1.223

	A	B	p(dif)
E-P (~) Est	10,15%	9,06%	1,09%
E-P (~) LI	11,54%	7,64%	3,90%
E-P (~) LS	8,72%	10,49%	-1,77%
IAM (~) Est	1,77%	2,76%	-0,99%
IAM (~) LI	2,41%	2,11%	0,30%
IAM (~) LS	1,20%	3,34%	-2,14%
AVC (~) Est	1,31%	0,48%	0,83%
AVC (~) LI	1,55%	0,23%	1,32%
AVC (~) LS	0,94%	0,85%	0,09%
Muerte (~) Est	7,62%	6,31%	1,32%
Muerte (~) LI	8,77%	5,15%	3,62%
Muerte (~) LS	6,44%	7,50%	-1,06%

### Variable principal

Valor de p valor ajustado				Valor de p crudo	
Directo		Test de Wald			
U	P	U	p	U	p
0.550466	0.582000	0.745528	0.455952	0.356973	0.721112

VARIABLES SECUNDARIAS

	Análisis ajustado (Test Wald)		Análisis crudo	
	U	p	U	p
IAM	1.5113184	0,1307074	1,791580	0,073200
AVC	2.1656131	0.0303407	2,049385	0,040424
Muerte	1.0852928	0,2777921	0,780974	0,434818

Representación de los ORs ajustados por el análisis secuencial:

Representación de los ORs crudos (sin ajuste por la metodología secuencial):

*Análisis pre-revisión*

No entraremos a comentar en detalle los resultados del análisis adicional con los datos previos a la revisión, pero se puede comprobar que son muy similares tanto la estimación del efecto como el valor de p del test de hipótesis.

El porcentaje de cambio de  $\theta$  entre antes y después de la revisión es de 3.2% (-0.129 y -0.125 respectivamente);

El valor de p varía de 0.582 a 0.549 (antes y después respectivamente).

De esta manera las conclusiones del estudio no difieren con la información aportada previa y posteriormente a la reevaluación de los pacientes e inclusión de los casos que se hallaban en seguimiento al dictar la finalización del periodo de inclusiones.

Los datos se describen en las siguientes tablas.

## DISCUSIÓN

---

	Fecha	A			B			Total			A-B
		n	N	%	n	N	%	n	N	%	%
Bellaterra	4-feb-97	107	1044	10.25%	101	1035	9,76%	208	2079	10,00%	0,49%

Corte	Z	V
5	-2.550	46.797

Estimador	Estimación	SE (Est)	IC 95%	
			Inf	Sup
Theta mle ( $\hat{\theta}$ )	-0.054	0,146	-0,341	0,232
Theta mediana ( $\hat{\theta}_M$ )	-0.094	-	-0,460	0,208
Theta ajustada ( $\tilde{\theta}$ )	-0.129	0,166	-0,454	0,196

Estimador	Estimación	IC 95%	
		Inf	Sup
OR (tetha mle)	0.947	0,711	1,261
OR (tetha mediana)	0.911	0,631	1,231
OR (tetha ajustada)	0.879	0,635	1,217

### CONCLUSIONES

#### *Respecto a los objetivos principales*

##### *Búsqueda de un diseño metodológico-estadístico teórico óptimo*

Los modelos secuenciales basados en límites continuos se mostraron como los diseños más idóneos para en el momento del diseño del estudio TIM.

La selección de modelo RPST 5305 se realizó considerando los buenos valores de los distintos estimadores (P50, P90 y E), la proximidad del vértice interno y la necesidad de variar lo menos posible el modelo previamente aprobado por comités éticos y entidades reguladoras.

Los métodos basados en gasto de alfa hubieran sido una alternativa clara a los métodos basados en límites continuos, si hubiera estado desarrollado software en el momento del diseño. No obstante, todavía queda por resolverse el análisis final con la estimación.

Las principales razones para desestimar los demás diseños fueron:

Requerimientos mayores de tamaño muestral mayor, como el clásico y equivalencia terapéutica.

Imposibilidad de finalización prematura bajo la hipótesis nula, como los métodos secuenciales por grupos.

Aplicación controvertida y falta de software: métodos bayesianos

Dificultades inherentes al método para su aplicación a ensayos clínicos para evaluación de eficacia comparada: restricción estocástica e intervalos de confianza repetidos.

## CONCLUSIONES

---

### *Comprobación de los supuestos teóricos previamente a la realización del estudio mediante simulaciones*

La comprobación permitió la identificación y corrección de un error en el programa utilizado para el diseño.

Con la evaluación posterior, las simulaciones han mostrado una buena adhesión de los modelos testados a las especificaciones teóricas.

### *Realización del diseño, monitorización del estudio y análisis estadístico de los resultados*

Los procesos de diseño a análisis se han realizado conforme a lo previsto en el plan estadístico. Se finalizó el estudio de acuerdo a las previsiones estimadas por  $P50(t)$ .

Los principales puntos y los problemas detectados han sido:

Durante el proceso de diseño: detección de un error en el software utilizado para el diseño

Durante la monitorización estadística: la depuración de los datos es un proceso lento y dinámico:

Hay un enlentecimiento, difícilmente evitable, del flujo de información entre la finalización de las inclusiones y su monitorización estadística

## CONCLUSIONES

---

La corrección de datos lleva consigo el cambio de situación de casos válidos a inválidos y viceversa.

Imposibilidad de realizar inspecciones ajustándose a calendario y que éstas sean de igual tamaño debido a importantes variaciones de velocidad de reclutamiento.

En estudios multicéntricos, aunque se puede intentar minimizar, si se finaliza cruzando la región de continuación es imposible de evitar un importante "overrunning".

En nuestro estudio no ha afectado al proceso de monitorización y acabamiento del estudio, pero es evidente que este hecho puede suceder en algún caso. Creemos que es necesario la realización y presentación de la monitorización con los datos en el momento de la inspección y con la información validada.

Durante el análisis:

La estimación por los distintos métodos no cambió significativamente los resultados ni conclusiones respecto a la variable principal.

Debido a la elevada correlación entre variable principal y secundaria se ajustaron mediante la técnica secuencial. Los resultados finales no varían respecto las conclusiones, pero hay una gran proximidad con la igualdad de los intervalos de confianza de dos variables secundarias que hubieran podido cambiar la significación.

El análisis definitivo debe efectuarse una vez finalizado completamente la monitorización estadística, debiendo esperar a la llegada de los casos incluidos en el estudio cuando se decidió finalizarlo.

### *Respecto a los objetivos secundarios*

La revisión de métodos alternativos al diseño clásico permitió identificar distintos modelos con mejores expectativas de tamaño muestral que el diseño clásico.

Las simulaciones comprobaron el cumplimiento de las especificaciones teóricas estadísticas de los modelos más favorables, una vez rediseñados los modelos RPST.

Los modelos secuenciales basados en límites continuos aportan la posibilidad de un ahorro importante en el número de pacientes respecto a modelos clásicos que globalmente se sitúa entre un 40% y un 50%. Las distintas opciones de estos modelos y las diferencias realmente esperadas modifican en estas expectativas

## ANEXOS

*DATOS TEÓRICOS PARA DISEÑOS SECUENCIALES*

Theta	Diseño	Med (n)	P90 (n)	E (n)
-1.50* Tr	TRIANGULAR	1236	2131	1342
	RPST 4000	1651	2907	1670
	RPST 4500	1175	2594	1332
	RPST 5000	1080	2421	1261
	RPST 5100	1069	2399	1253
	RPST 5200	1059	2380	1245
	RPST 5300	1050	2363	1238
	RPST 5305	1050	2362	1238
	RPST 5400	1042	2347	1232
	RPST 5500	1035	2333	1227
	RPST 6000	1009	2280	1206
	RPST 7000	982	2221	1183
	RPST 10000	958	2164	1160
	RPST	952	2148	1153
-1.25* Tr	TRIANGULAR	1572	2713	1689
	RPST 4000	2371	3009	2067
	RPST 4500	1622	3385	1726
	RPST 5000	1475	3676	1664
	RPST 5100	1456	3632	1658
	RPST 5200	1440	3593	1652
	RPST 5300	1426	3558	1647
	RPST 5305	1426	3556	1647
	RPST 5400	1414	3526	1643
	RPST 5500	1403	3497	1639
	RPST 6000	1361	3387	1627
	RPST 7000	1318	3263	1615
	RPST 10000	1279	3138	1604
	RPST	1269	3103	1589
-1 * Tr	TRIANGULAR	1980	3203	2069
	RPST 4000	3113	3113	2481
	RPST 4500	2301	3502	2156
	RPST 5000	2033	3891	2111
	RPST 5100	2001	3969	2108
	RPST 5200	1974	4047	2106
	RPST 5300	1949	4125	2105
	RPST 5305	1948	4128	2105
	RPST 5400	1927	4202	2105
	RPST 5500	1908	4280	2105

# ANEXOS

---

	RPST 6000	1837	4669	2113
	RPST 7000	1762	4785	2136
	RPST 10000	1696	4483	2182
	RPST	1678	4398	2120
-0.75* Tr	TRIANGULAR	2273	3471	2347
	RPST 4000	3219	3219	2808
	RPST 4500	2879	3622	2488
	RPST 5000	2485	4024	2447
	RPST 5100	2440	4104	2445
	RPST 5200	2400	4185	2445
	RPST 5300	2366	4265	2446
	RPST 5305	2364	4269	2446
	RPST 5400	2335	4346	2447
	RPST 5500	2308	4426	2450
	RPST 6000	2209	4829	2467
	RPST 7000	2106	5634	2512
	RPST 10000	2014	5472	2609
	RPST	1990	5331	2476

# ANEXOS

---

Theta	Diseño	Med (n)	P90 (n)	E (n)
-0.5 * Tr	TRIANGULAR	2286	3516	2416
	RPST 4000	3327	3327	2958
	RPST 4500	2678	3743	2591
	RPST 5000	2379	4159	2524
	RPST 5100	2344	4242	2518
	RPST 5200	2313	4326	2513
	RPST 5300	2286	4409	2510
	RPST 5305	2284	4413	2510
	RPST 5400	2262	4492	2507
	RPST 5500	2240	4575	2505
	RPST 6000	2161	4991	2505
	RPST 7000	2077	5417	2521
	RPST 10000	2002	5042	2568
	RPST	1982	4936	2482
-0.25* Tr	TRIANGULAR	2189	3336	2349
	RPST 4000	3300	3437	2975
	RPST 4500	2365	3867	2519
	RPST 5000	2170	4297	2412
	RPST 5100	2147	4383	2400
	RPST 5200	2125	4469	2388
	RPST 5300	2107	4476	2379
	RPST 5305	2106	4474	2378
	RPST 5400	2090	4434	2370
	RPST 5500	2075	4396	2362
	RPST 6000	2020	4250	2335
	RPST 7000	1959	4083	2307
	RPST 10000	1905	3917	2285
	RPST	1890	3869	2261
+0 * Tr	TRIANGULAR	2192	3229	2342
	RPST 4000	3145	3549	3013
	RPST 4500	2325	3992	2502
	RPST 5000	2150	3959	2375
	RPST 5100	2128	3923	2359
	RPST 5200	2109	3890	2345
	RPST 5300	2092	3861	2332
	RPST 5305	2091	3859	2332
	RPST 5400	2077	3834	2321
	RPST 5500	2064	3810	2311
	RPST 6000	2013	3717	2271
	RPST 7000	1958	3610	2228
	RPST 10000	1907	3502	2186
	RPST	1894	3471	2173
+0.25* Tr	TRIANGULAR	2331	3553	2502
	RPST 4000	3515	3661	3168

# ANEXOS

---

	RPST 4500	2519	4118	2683
	RPST 5000	2312	4576	2569
	RPST 5100	2286	4668	2556
	RPST 5200	2264	4759	2544
	RPST 5300	2244	4767	2533
	RPST 5305	2243	4765	2533
	RPST 5400	2226	4722	2524
	RPST 5500	2210	4682	2516
	RPST 6000	2151	4526	2486
	RPST 7000	2087	4348	2457
	RPST 10000	2029	4172	2434
	RPST	2013	4120	2408
+0.5 * Tr	TRIANGULAR	2593	3988	2740
	RPST 4000	3774	3774	3355
	RPST 4500	3037	4245	2938
	RPST 5000	2698	4717	2862
	RPST 5100	2658	4812	2855
	RPST 5200	2623	4906	2850
	RPST 5300	2592	5000	2846
	RPST 5305	2591	5005	2846
	RPST 5400	2565	5095	2843
	RPST 5500	2541	5189	2842
	RPST 6000	2451	5661	2841
	RPST 7000	2356	6144	2859
	RPST 10000	2271	5719	2913
	RPST	2248	5599	2815

# ANEXOS

---

Theta	Diseño	Med(n)	P90 (n)	E (n)
+0.75* Tr	TRIANGULAR	2744	4192	2834
	RPST 4000	3887	3887	3390
	RPST 4500	3476	4373	3004
	RPST 5000	3001	4859	2954
	RPST 5100	2946	4956	2953
	RPST 5200	2898	5053	2952
	RPST 5300	2857	5150	2953
	RPST 5305	2855	5155	2953
	RPST 5400	2820	5247	2955
	RPST 5500	2787	5345	2958
	RPST 6000	2667	5830	2979
	RPST 7000	2542	6802	3033
	RPST 10000	2432	6607	3150
	RPST	2403	6437	2990
+1 * Tr	TRIANGULAR	2545	4116	2659
	RPST 4000	4000	4000	3188
	RPST 4500	2956	4500	2770
	RPST 5000	2613	5000	2712
	RPST 5100	2572	5100	2709
	RPST 5200	2536	5200	2706
	RPST 5300	2504	5300	2705
	RPST 5305	2503	5305	2705
	RPST 5400	2476	5400	2705
	RPST 5500	2452	5500	2705
	RPST 6000	2360	6000	2715
	RPST 7000	2264	6148	2744
	RPST 10000	2179	5761	2804
	RPST	2157	5652	2724
+1.25* Tr	TRIANGULAR	2149	3708	2309
	RPST 4000	3241	4113	2826
	RPST 4500	2217	4627	2360
	RPST 5000	2016	5025	2275
	RPST 5100	1991	4965	2266
	RPST 5200	1969	4911	2258
	RPST 5300	1950	4863	2251
	RPST 5305	1949	4861	2251
	RPST 5400	1933	4820	2246
	RPST 5500	1918	4780	2241
	RPST 6000	1861	4630	2223
	RPST 7000	1801	4459	2208
	RPST 10000	1749	4290	2192
	RPST	1735	4242	2172
+1.5 * Tr	TRIANGULAR	1796	3097	1950
	RPST 4000	2399	4225	2427

## ANEXOS

---

RPST 4500	1707	3770	1935
RPST 5000	1570	3518	1833
RPST 5100	1553	3487	1820
RPST 5200	1539	3458	1809
RPST 5300	1526	3433	1799
RPST 5305	1525	3432	1799
RPST 5400	1514	3411	1790
RPST 5500	1504	3391	1782
RPST 6000	1467	3313	1752
RPST 7000	1427	3227	1718
RPST 10000	1392	3145	1685
RPST	1383	3122	1675

**DATOS SIMULADOS PARA DISEÑOS SECUENCIALES**

THETA	TIPO	NMEAN	NMIN	N50	N90	NMAX
+1.5 *Tr	Triangular	1992.31	400	1832	3166	5620
	RPST 4000	2491.96	378	2302	4136	4902
	RPST 4500	1939.93	226	1648	3618	5246
	RPST 5000	1856.06	228	1580	3457	5906
	RPST 5100	1831.94	176	1532	3407	5910
	RPST 5200	1821.85	192	1528	3382	6188
	RPST 5300	1811.69	232	1522	3339	6216
	RPST 5305	1804.57	232	1510	3350	6192
	RPST 5400	1808.56	242	1512	3364	6370
	RPST 5500	1818.38	202	1526	3426	6292
	RPST 6000	1764.06	206	1466	3285	6804
	RPST 7000	1754.49	194	1444	3279	7798
	RPST 10000	1729.57	172	1436	3226	10626
	RPST	1743.46	180	1422	3250	10410
	CLASICO	3962.00	3962	3962	3962	3962
+1.25*Tr	Triangular	2337.72	478	2188	3776	5862
	RPST 4000	2971.10	466	3090	4224	4858
	RPST 4500	2445.31	240	2172	4538	5464
	RPST 5000	2341.64	188	2012	4723	5846
	RPST 5100	2333.23	150	1996	4688	5952
	RPST 5200	2318.24	176	1984	4648	6150
	RPST 5300	2323.32	150	1954	4692	6224
	RPST 5305	2317.39	216	1974	4681	6326
	RPST 5400	2316.26	192	1982	4642	6198
	RPST 5500	2287.28	152	1914	4568	6326
	RPST 6000	2273.76	208	1890	4494	6874
	RPST 7000	2280.88	176	1876	4492	7942
	RPST 10000	2251.96	176	1810	4418	11020
	RPST	2273.57	190	1822	4421	13446
	CLASICO	3962.00	3962	3962	3962	3962
+1 *Tr	Triangular	2671.24	466	2554	4115	5892
	RPST 4000	3355.36	384	3818	4182	4718
	RPST 4500	2930.25	262	2872	4584	5140
	RPST 5000	2875.11	212	2630	4998	5720
	RPST 5100	2864.14	274	2562	5100	5804
	RPST 5200	2844.84	256	2540	5162	5890
	RPST 5300	2860.30	186	2538	5266	6204
	RPST 5305	2867.72	176	2546	5256	5996
	RPST 5400	2853.31	162	2498	5346	6326
	RPST 5500	2846.98	236	2488	5420	6244
RPST 6000	2856.40	218	2430	5786	6892	

## ANEXOS

---

	RPST 7000	2866.14	214	2356	5940	7666
	RPST 10000	2931.08	170	2290	5980	10988
	RPST	2961.66	214	2316	5853	18046
	CLASICO	3962.00	3962	3962	3962	3962
+0.75*Tr	Triangular	2862.54	530	2752	4220	5646
	RPST 4000	3528.55	508	3796	4084	4628
	RPST 4500	3186.13	248	3364	4486	5030
	RPST 5000	3187.99	262	3058	4946	5550
	RPST 5100	3152.79	146	2978	5022	5672
	RPST 5200	3164.47	206	2936	5110	5818
	RPST 5300	3147.67	206	2876	5202	6024
	RPST 5305	3173.73	206	2910	5200	5854
	RPST 5400	3178.05	200	2880	5290	6232
	RPST 5500	3140.00	212	2828	5356	6058
	RPST 6000	3184.99	224	2748	5794	6604
	RPST 7000	3217.95	220	2650	6554	7562
	RPST 10000	3296.83	206	2608	6516	10562
	RPST	3386.36	222	2600	6745	28290
	CLASICO	3962.00	3962	3962	3962	3962

# ANEXOS

---

THETA	TIPO	NMEAN	NMIN	N50	N90	NMAX
+0.5 *Tr	Triangular	2786.87	514	2636	4010	5588
	RPST 4000	3481.66	598	3668	3956	4372
	RPST 4500	3127.65	262	3106	4350	4928
	RPST 5000	3083.74	236	2820	4762	5414
	RPST 5100	3058.67	264	2774	4838	5416
	RPST 5200	3026.89	240	2696	4916	5656
	RPST 5300	3038.07	206	2688	5006	5710
	RPST 5305	3050.64	226	2700	5006	5706
	RPST 5400	3064.73	238	2696	5082	5760
	RPST 5500	3037.96	284	2634	5166	5866
	RPST 6000	3062.90	228	2604	5535	6336
	RPST 7000	3069.28	234	2502	6056	7404
	RPST 10000	3070.80	242	2414	5690	10120
	RPST	3117.09	186	2428	5769	18004
	CLASICO	3962.00	3962	3962	3962	3962
+0.25*Tr	Triangular	2537.53	474	2352	3557	5056
	RPST 4000	3307.52	694	3462	3808	4390
	RPST 4500	2846.43	286	2620	4124	4882
	RPST 5000	2729.07	340	2402	4444	5202
	RPST 5100	2708.01	212	2386	4482	5416
	RPST 5200	2713.19	262	2378	4560	5412
	RPST 5300	2687.60	282	2350	4556	5524
	RPST 5305	2688.75	244	2342	4600	5430
	RPST 5400	2704.74	284	2366	4626	5680
	RPST 5500	2677.77	256	2326	4574	5672
	RPST 6000	2660.30	210	2274	4536	6156
	RPST 7000	2636.48	192	2222	4381	7174
	RPST 10000	2587.72	264	2170	4224	9698
	RPST	2576.50	222	2134	4239	19044
	CLASICO	3962.00	3962	3962	3962	3962
+0 *Tr	Triangular	2401.28	588	2246	3257	5254
	RPST 4000	3146.35	612	3224	3662	4138
	RPST 4500	2643.07	316	2434	3890	4656
	RPST 5000	2517.95	220	2258	3923	5052
	RPST 5100	2511.03	226	2242	3905	5096
	RPST 5200	2486.87	328	2198	3898	5172
	RPST 5300	2486.64	330	2206	3886	5196
	RPST 5305	2474.54	274	2194	3822	5228
	RPST 5400	2472.01	224	2188	3862	5394
	RPST 5500	2488.88	292	2198	3913	5464
	RPST 6000	2425.95	234	2130	3731	5912
	RPST 7000	2377.71	206	2066	3676	6694
	RPST 10000	2309.75	176	2024	3490	9094
	RPST	2307.70	192	2034	3519	12924

# ANEXOS

---

	CLASICO	3962.00	3962	3962	3962	3962
-0.25*Tr	Triangular	2394.75	450	2224	3354	4894
	RPST 4000	3102.46	512	3252	3572	4020
	RPST 4500	2655.34	330	2456	3864	4568
	RPST 5000	2584.76	262	2294	4184	4818
	RPST 5100	2565.60	226	2262	4254	5208
	RPST 5200	2529.94	290	2218	4241	5096
	RPST 5300	2546.25	212	2226	4326	5152
	RPST 5305	2545.97	256	2218	4316	5202
	RPST 5400	2529.31	208	2198	4370	5246
	RPST 5500	2524.61	238	2192	4334	5338
	RPST 6000	2532.59	232	2158	4377	5792
	RPST 7000	2476.41	186	2090	4146	6732
	RPST 10000	2420.11	196	2022	3944	8982
	RPST	2445.88	214	2022	4036	21224
	CLASICO	3962.00	3962	3962	3962	3962
-0.5 *Tr	Triangular	2454.40	342	2320	3508	4850
	RPST 4000	3069.02	374	3240	3480	3896
	RPST 4500	2762.86	192	2762	3830	4350
	RPST 5000	2706.99	202	2492	4198	4748
	RPST 5100	2694.77	250	2432	4264	4954
	RPST 5200	2709.07	208	2436	4342	4996
	RPST 5300	2695.19	186	2384	4417	5070
	RPST 5305	2684.52	182	2366	4416	5016
	RPST 5400	2697.08	222	2368	4484	5096
	RPST 5500	2675.65	216	2328	4550	5162
	RPST 6000	2689.43	208	2272	4887	5722
	RPST 7000	2711.76	168	2216	5350	6556
	RPST 10000	2724.31	186	2142	5103	8908
	RPST	2723.50	176	2114	4970	19066
	CLASICO	3962.00	3962	3962	3962	3962
THETA	TIPO	NMEAN	NMIN	N50	N90	NMAX
-0.75*Tr	Triangular	2366.32	364	2288	3476	4590
	RPST 4000	2929.82	386	3146	3374	3798
	RPST 4500	2660.01	204	2856	3718	4274
	RPST 5000	2612.91	202	2484	4089	4672
	RPST 5100	2624.35	214	2478	4156	4644
	RPST 5200	2613.14	214	2434	4223	4746
	RPST 5300	2613.07	210	2418	4296	4944
	RPST 5305	2629.15	194	2440	4310	4892
	RPST 5400	2634.82	216	2404	4370	4964
	RPST 5500	2600.49	168	2324	4450	4956
	RPST 6000	2633.82	198	2294	4790	5498
	RPST 7000	2661.44	170	2190	5427	6270

# ANEXOS

---

	RPST 10000	2770.25	190	2188	5505	8704	
	RPST	2802.56	178	2164	5479	22732	
	CLASICO	3962.00	3962	3962	3962	3962	
-1	*Tr	Triangular	2094.16	378	2006	3226	4438
	RPST 4000	2605.29	402	2972	3242	3590	
	RPST 4500	2300.19	212	2276	3562	4144	
	RPST 5000	2227.43	202	2022	3900	4374	
	RPST 5100	2244.40	170	2016	3962	4482	
	RPST 5200	2231.64	154	1984	4030	4548	
	RPST 5300	2246.54	172	1992	4102	4674	
	RPST 5305	2221.53	146	1976	4092	4680	
	RPST 5400	2228.95	198	1934	4161	4800	
	RPST 5500	2216.29	168	1930	4216	4858	
	RPST 6000	2226.93	166	1888	4508	5220	
	RPST 7000	2246.87	186	1828	4663	5954	
	RPST 10000	2316.93	178	1818	4591	8492	
	RPST	2297.95	118	1798	4528	14392	
	CLASICO	3962.00	3962	3962	3962	3962	
-1.25	*Tr	Triangular	1711.12	216	1604	2746	4152
	RPST 4000	2158.12	256	2230	3076	3474	
	RPST 4500	1777.51	190	1578	3318	3932	
	RPST 5000	1717.63	170	1462	3510	4180	
	RPST 5100	1700.19	136	1448	3468	4402	
	RPST 5200	1715.63	154	1460	3455	4386	
	RPST 5300	1691.95	136	1444	3404	4420	
	RPST 5305	1683.76	150	1418	3428	4526	
	RPST 5400	1696.42	162	1428	3419	4570	
	RPST 5500	1688.34	170	1438	3330	4594	
	RPST 6000	1668.01	138	1398	3316	4924	
	RPST 7000	1657.76	158	1366	3233	5778	
	RPST 10000	1652.03	136	1332	3222	7916	
	RPST	1677.94	142	1344	3257	12420	
	CLASICO	3962.00	3962	3962	3962	3962	
-1.50	*Tr	Triangular	1357.58	278	1252	2166	3890
	RPST 4000	1713.28	230	1594	2850	3398	
	RPST 4500	1339.04	150	1142	2500	3738	
	RPST 5000	1270.74	172	1068	2360	3992	
	RPST 5100	1247.10	168	1040	2324	4094	
	RPST 5200	1230.15	140	1028	2297	4226	
	RPST 5300	1236.00	116	1036	2315	4272	
	RPST 5305	1229.05	122	1020	2296	4352	
	RPST 5400	1245.20	132	1032	2324	4258	
	RPST 5500	1229.59	114	1022	2292	4450	
	RPST 6000	1215.99	168	1014	2256	4742	
	RPST 7000	1198.58	126	990	2251	5466	
	RPST 10000	1193.44	132	982	2233	7034	

## ANEXOS

---

RPST	1181.46	116	964	2207	7538
CLASICO	3962.00	3962	3962	3962	3962

# ANEXOS

## AHORRO

THETA	TIPO	AHORRO	DMEAN	DMIN	D50	D90	DMAX
+1.5 *Tr	Triangular	97.2	1969.69	-1658	2130	2902	3562
	RPST 4000	85.1	1470.04	-940	1660	2782	3584
	RPST 4500	92.5	2022.07	-1284	2314	3174	3736
	RPST 5000	93.7	2105.94	-1944	2384	3250	3734
	RPST 5100	93.9	2130.06	-1948	2432	3256	3786
	RPST 5200	94.2	2140.15	-2226	2434	3264	3770
	RPST 5300	94.1	2150.31	-2254	2440	3272	3730
	RPST 5305	94.1	2157.43	-2230	2454	3279	3730
	RPST 5400	94.0	2153.44	-2408	2450	3280	3720
	RPST 5500	93.9	2143.62	-2330	2436	3290	3760
	RPST 6000	94.6	2197.94	-2842	2496	3310	3756
	RPST 7000	94.5	2207.51	-3836	2518	3314	3768
	RPST 10000	94.8	2232.43	-6664	2526	3324	3790
	RPST	94.5	2218.54	-6448	2542	3326	3782
	CLASICO	100.0	.00	0	0	0	0
+1.25*Tr	Triangular	92.6	1624.28	-1900	1774	2778	3484
	RPST 4000	71.8	990.90	-896	872	2560	3496
	RPST 4500	80.6	1516.69	-1502	1792	3070	3722
	RPST 5000	84.0	1620.36	-1884	1950	3142	3774
	RPST 5100	84.5	1628.77	-1990	1966	3140	3812
	RPST 5200	84.7	1643.76	-2188	1978	3173	3786
	RPST 5300	84.3	1638.68	-2262	2008	3174	3812
	RPST 5305	84.9	1644.61	-2364	1990	3174	3746
	RPST 5400	85.0	1645.74	-2236	1982	3178	3770
	RPST 5500	85.6	1674.72	-2364	2048	3192	3810
	RPST 6000	86.1	1688.24	-2912	2072	3200	3754
	RPST 7000	86.2	1681.12	-3980	2086	3233	3786
	RPST 10000	86.5	1710.04	-7058	2152	3248	3786
	RPST	87.0	1688.43	-9484	2140	3250	3772
	CLASICO	100.0	.00	0	0	0	0
+1 *Tr	Triangular	87.5	1290.76	-1930	1408	2578	3496
	RPST 4000	66.0	606.64	-756	144	2206	3578
	RPST 4500	66.2	1031.75	-1178	1094	2364	3700
	RPST 5000	71.2	1086.89	-1758	1332	2362	3750
	RPST 5100	71.4	1097.86	-1842	1400	2382	3688
	RPST 5200	72.7	1117.16	-1928	1424	2316	3706
	RPST 5300	72.4	1101.70	-2242	1426	2320	3776
	RPST 5305	72.5	1094.28	-2034	1418	2385	3786
	RPST 5400	72.7	1108.69	-2364	1464	2332	3800
	RPST 5500	73.5	1115.02	-2282	1476	2336	3726

# ANEXOS

---

	RPST 6000	74.7	1105.61	-2930	1532	3036	3744
	RPST 7000	76.1	1095.86	-3704	1606	3062	3748
	RPST 10000	76.5	1030.92	-7026	1672	3100	3792
	RPST	76.6	1000.34	-14084	1646	3112	3748
	CLASICO	100.0	.00	0	0	0	0
+0.75*Tr	Triangular	85.1	1099.46	-1684	1210	2258	3432
	RPST 4000	76.3	433.45	-666	166	1446	3454
	RPST 4500	58.6	775.87	-1068	598	2499	3714
	RPST 5000	62.9	774.01	-1588	904	2665	3700
	RPST 5100	64.4	809.21	-1710	984	2709	3816
	RPST 5200	65.2	797.53	-1856	1026	2667	3756
	RPST 5300	65.2	814.34	-2062	1088	2746	3756
	RPST 5305	64.7	788.27	-1892	1054	2727	3756
	RPST 5400	66.2	783.95	-2270	1084	2682	3762
	RPST 5500	67.3	822.00	-2096	1134	2772	3750
	RPST 6000	68.1	777.01	-2642	1214	2779	3738
	RPST 7000	69.7	744.05	-3600	1312	2842	3742
	RPST 10000	71.6	665.17	-6600	1354	2851	3756
	RPST	71.7	575.64	-24328	1362	2844	3740
	CLASICO	100.0	.00	0	0	0	0

# ANEXOS

THETA	TIPO	AHORRO	DMEAN	DMIN	D50	D90	DMAX
+0.5 *Tr	Triangular	89.1	1175.13	-1626	1326	2058	3448
	RPST 4000	90.6	480.34	-410	294	1228	3364
	RPST 4500	65.4	834.35	-966	856	2034	3700
	RPST 5000	68.7	878.26	-1452	1142	2189	3726
	RPST 5100	70.1	903.33	-1454	1188	2221	3698
	RPST 5200	71.4	935.11	-1694	1266	2252	3722
	RPST 5300	71.9	923.93	-1748	1274	2238	3756
	RPST 5305	71.1	911.36	-1744	1262	2261	3736
	RPST 5400	71.1	897.27	-1798	1266	2250	3724
	RPST 5500	72.2	924.04	-1904	1328	2282	3678
	RPST 6000	72.9	899.10	-2374	1360	2314	3734
	RPST 7000	75.0	892.72	-3442	1460	2373	3728
	RPST 10000	77.2	891.20	-6158	1548	2394	3720
	RPST	76.6	844.91	-14042	1536	2413	3776
	CLASICO	100.0	.00	0	0	0	0
+0.25*Tr	Triangular	95.7	1424.47	-1094	1610	2098	3488
	RPST 4000	97.8	654.48	-428	500	1312	3268
	RPST 4500	82.0	1115.57	-920	1342	2030	3676
	RPST 5000	83.2	1232.93	-1240	1560	2164	3622
	RPST 5100	83.8	1254.00	-1454	1576	2196	3750
	RPST 5200	83.7	1248.81	-1450	1584	2212	3700
	RPST 5300	84.7	1274.40	-1562	1612	2220	3680
	RPST 5305	84.1	1273.25	-1468	1620	2228	3718
	RPST 5400	84.2	1257.26	-1718	1596	2234	3678
	RPST 5500	84.9	1284.23	-1710	1636	2254	3706
	RPST 6000	85.6	1301.70	-2194	1688	2288	3752
	RPST 7000	86.5	1325.53	-3212	1740	2318	3770
	RPST 10000	88.1	1374.28	-5736	1792	2374	3698
	RPST	88.0	1385.50	-15082	1828	2382	3740
	CLASICO	100.0	.00	0	0	0	0
+0 *Tr	Triangular	98.1	1560.72	-1292	1716	2152	3374
	RPST 4000	99.6	815.65	-176	738	1408	3350
	RPST 4500	92.7	1318.93	-694	1530	2088	3646
	RPST 5000	90.4	1444.05	-1090	1704	2228	3742
	RPST 5100	90.5	1450.97	-1134	1720	2238	3736
	RPST 5200	90.4	1475.13	-1210	1764	2260	3634
	RPST 5300	90.7	1475.36	-1234	1756	2268	3632
	RPST 5305	91.2	1487.46	-1266	1768	2268	3688
	RPST 5400	90.8	1489.99	-1432	1774	2280	3738
	RPST 5500	90.4	1473.12	-1502	1764	2286	3670
	RPST 6000	91.9	1536.05	-1950	1832	2324	3728
	RPST 7000	92.4	1584.29	-2732	1896	2376	3756
	RPST 10000	93.6	1652.25	-5132	1938	2412	3786
	RPST	94.1	1654.30	-8962	1928	2418	3770

# ANEXOS

	CLASICO	100.0	.00	0	0	0	0
-0.25*Tr	Triangular	97.6	1567.25	-932	1738	2208	3512
	RPST 4000	100.0	859.54	-58	710	1472	3450
	RPST 4500	94.4	1306.66	-606	1506	2150	3632
	RPST 5000	85.2	1377.24	-856	1668	2284	3700
	RPST 5100	85.7	1396.40	-1246	1700	2298	3736
	RPST 5200	87.0	1432.06	-1134	1746	2312	3672
	RPST 5300	86.4	1415.75	-1190	1738	2330	3750
	RPST 5305	86.2	1416.03	-1240	1744	2334	3706
	RPST 5400	86.5	1432.69	-1284	1764	2342	3754
	RPST 5500	87.0	1437.39	-1376	1770	2350	3724
	RPST 6000	86.5	1429.41	-1830	1804	2386	3730
	RPST 7000	88.6	1485.59	-2770	1872	2432	3776
	RPST 10000	90.1	1541.89	-5020	1940	2472	3766
	RPST	89.4	1516.12	-17262	1940	2470	3748
	CLASICO	100.0	.00	0	0	0	0
-0.5 *Tr	Triangular	96.8	1507.60	-888	1642	2287	3620
	RPST 4000	100.0	892.98	66	722	1545	3588
	RPST 4500	97.3	1199.14	-388	1200	2250	3770
	RPST 5000	77.4	1255.01	-786	1470	2400	3760
	RPST 5100	76.9	1267.23	-992	1530	2418	3712
	RPST 5200	76.2	1252.93	-1034	1528	2428	3754
	RPST 5300	76.8	1266.81	-1108	1578	2456	3776
	RPST 5305	77.3	1277.48	-1054	1596	2446	3780
	RPST 5400	77.0	1264.92	-1134	1594	2452	3740
	RPST 5500	77.9	1286.35	-1200	1634	2477	3746
	RPST 6000	79.0	1272.57	-1760	1690	2504	3754
	RPST 7000	80.3	1250.24	-2594	1746	2556	3794
	RPST 10000	81.6	1237.69	-4946	1822	2594	3776
	RPST	82.8	1238.50	-15104	1848	2596	3786
	CLASICO	100.0	.00	0	0	0	0
THETA	TIPO	AHORRO	DMEAN	DMIN	D50	D90	DMAX
-0.75*Tr	Triangular	98.1	1595.68	-628	1674	2541	3598
	RPST 4000	100.0	1032.18	164	816	1875	3576
	RPST 4500	99.5	1301.99	-312	1108	2752	3758
	RPST 5000	81.4	1349.09	-710	1478	2893	3760
	RPST 5100	77.2	1337.65	-682	1484	2930	3748
	RPST 5200	75.9	1348.86	-784	1528	2920	3748
	RPST 5300	75.0	1348.93	-982	1546	2954	3752
	RPST 5305	74.3	1332.85	-930	1522	2943	3768
	RPST 5400	74.2	1327.18	-1002	1558	2922	3746
	RPST 5500	75.4	1361.51	-994	1640	2967	3794
	RPST 6000	76.4	1328.18	-1536	1668	2976	3764
	RPST 7000	78.2	1300.56	-2308	1772	3002	3792

# ANEXOS

---

	RPST 10000	78.9	1191.75	-4742	1774	3015	3772	
	RPST	79.3	1159.44	-18770	1798	3045	3784	
	CLASICO	100.0	.00	0	0	0	0	
-1	*Tr	Triangular	99.1	1867.84	-476	1956	2876	3584
	RPST 4000	100.0	1356.71	372	990	2594	3560	
	RPST 4500	100.0	1661.81	-182	1686	3102	3750	
	RPST 5000	93.4	1734.57	-412	1940	3201	3760	
	RPST 5100	90.1	1717.60	-520	1946	3188	3792	
	RPST 5200	86.7	1730.36	-586	1978	3207	3808	
	RPST 5300	84.7	1715.46	-712	1970	3206	3790	
	RPST 5305	84.8	1740.47	-718	1986	3214	3816	
	RPST 5400	83.2	1733.05	-838	2028	3214	3764	
	RPST 5500	83.6	1745.71	-896	2032	3229	3794	
	RPST 6000	84.3	1735.07	-1258	2074	3260	3796	
	RPST 7000	85.1	1715.13	-1992	2134	3260	3776	
	RPST 10000	85.4	1645.07	-4530	2146	3270	3784	
	RPST	85.9	1664.05	-10430	2164	3282	3844	
	CLASICO	100.0	.00	0	0	0	0	
-1.25	*Tr	Triangular	99.9	2250.88	-190	2358	3094	3746
	RPST 4000	100.0	1803.88	488	1732	2934	3706	
	RPST 4500	100.0	2184.49	30	2386	3304	3772	
	RPST 5000	99.0	2244.37	-218	2500	3368	3792	
	RPST 5100	98.3	2261.81	-440	2514	3376	3826	
	RPST 5200	96.9	2246.37	-424	2502	3370	3808	
	RPST 5300	96.3	2270.05	-458	2518	3384	3826	
	RPST 5305	96.0	2278.24	-564	2544	3392	3812	
	RPST 5400	95.3	2265.58	-608	2534	3386	3800	
	RPST 5500	94.7	2273.66	-632	2524	3394	3792	
	RPST 6000	94.2	2293.99	-962	2564	3408	3824	
	RPST 7000	94.7	2304.24	-1816	2596	3442	3804	
	RPST 10000	94.7	2309.97	-3954	2630	3446	3826	
	RPST	94.7	2284.06	-8458	2618	3437	3820	
	CLASICO	100.0	.00	0	0	0	0	
-1.50	*Tr	Triangular	100.0	2604.42	72	2710	3244	3684
	RPST 4000	100.0	2248.72	564	2370	3146	3732	
	RPST 4500	100.0	2622.96	224	2820	3424	3812	
	RPST 5000	100.0	2691.26	-30	2894	3478	3790	
	RPST 5100	99.9	2714.90	-132	2922	3484	3794	
	RPST 5200	99.9	2731.85	-264	2934	3490	3822	
	RPST 5300	99.7	2726.00	-310	2926	3498	3846	
	RPST 5305	99.6	2732.95	-390	2942	3500	3840	
	RPST 5400	99.4	2716.80	-296	2930	3500	3830	
	RPST 5500	99.4	2732.41	-488	2940	3492	3848	
	RPST 6000	99.0	2746.01	-780	2948	3506	3794	
	RPST 7000	99.0	2763.42	-1504	2972	3526	3836	
	RPST 10000	98.9	2768.56	-3072	2980	3530	3830	

## ANEXOS

---

RPST	99.0	2780.54	-3576	2998	3524	3846
CLASICO	100.0	.00	0	0	0	0

**CARACTERÍSTICA OPERANTE**

THETA	TIPO	SIGNI
+1.5 *Tr	Triangular	98.6
	RPST 4000	98.7
	RPST 4500	99.1
	RPST 5000	99.1
	RPST 5100	99.0
	RPST 5200	99.0
	RPST 5300	99.0
	RPST 5305	98.8
	RPST 5400	98.9
	RPST 5500	98.9
	RPST 6000	99.1
	RPST 7000	98.9
	RPST 10000	98.7
	RPST	98.9
	CLASICO	98.2
	+1.25*Tr	Triangular
RPST 4000		93.5
RPST 4500		94.7
RPST 5000		94.5
RPST 5100		94.9
RPST 5200		94.7
RPST 5300		94.5
RPST 5305		94.9
RPST 5400		94.4
RPST 5500		94.9
RPST 6000		94.7
RPST 7000		94.8
RPST 10000		94.9
RPST		95.0
CLASICO		92.8
+1 *Tr		Triangular
	RPST 4000	79.2
	RPST 4500	81.9
	RPST 5000	81.8
	RPST 5100	82.0
	RPST 5200	81.4
	RPST 5300	81.9
	RPST 5305	80.7
	RPST 5400	81.0
	RPST 5500	81.2

## ANEXOS

---

	RPST 6000	81.7
	RPST 7000	82.0
	RPST 10000	82.0
	RPST	81.1
	CLASICO	79.4
+0.75*Tr	Triangular	56.0
	RPST 4000	55.8
	RPST 4500	56.8
	RPST 5000	56.4
	RPST 5100	57.0
	RPST 5200	55.8
	RPST 5300	56.8
	RPST 5305	55.7
	RPST 5400	55.4
	RPST 5500	56.2
	RPST 6000	56.1
	RPST 7000	54.8
	RPST 10000	54.6
	RPST	53.8
	CLASICO	57.0

# ANEXOS

---

THETA	TIPO	SIGNI
+0.5 *Tr	Triangular	27.9
	RPST 4000	29.1
	RPST 4500	29.7
	RPST 5000	29.3
	RPST 5100	28.3
	RPST 5200	27.9
	RPST 5300	28.1
	RPST 5305	27.8
	RPST 5400	27.0
	RPST 5500	28.2
	RPST 6000	27.6
	RPST 7000	27.0
	RPST 10000	25.6
	RPST	24.5
	CLASICO	29.5
	+0.25*Tr	Triangular
RPST 4000		10.6
RPST 4500		10.9
RPST 5000		9.8
RPST 5100		10.3
RPST 5200		10.3
RPST 5300		9.9
RPST 5305		10.3
RPST 5400		10.1
RPST 5500		10.5
RPST 6000		10.1
RPST 7000		9.7
RPST 10000		10.0
RPST		9.6
CLASICO		11.4
+0 *Tr		Triangular
	RPST 4000	4.8
	RPST 4500	5.1
	RPST 5000	5.1
	RPST 5100	5.3
	RPST 5200	5.4
	RPST 5300	5.1
	RPST 5305	5.1
	RPST 5400	5.3
	RPST 5500	5.0
	RPST 6000	4.9
	RPST 7000	5.4
	RPST 10000	5.0
	RPST	5.5

# ANEXOS

---

	CLASICO	5.0
-0.25*Tr	Triangular	10.9
	RPST 4000	9.7
	RPST 4500	11.3
	RPST 5000	10.5
	RPST 5100	10.5
	RPST 5200	10.6
	RPST 5300	10.2
	RPST 5305	10.4
	RPST 5400	10.6
	RPST 5500	9.9
	RPST 6000	10.6
	RPST 7000	9.7
	RPST 10000	9.3
	RPST	8.8
	CLASICO	11.2
-0.5 *Tr	Triangular	28.3
	RPST 4000	28.8
	RPST 4500	28.8
	RPST 5000	28.3
	RPST 5100	28.1
	RPST 5200	28.1
	RPST 5300	27.7
	RPST 5305	27.8
	RPST 5400	28.3
	RPST 5500	27.9
	RPST 6000	28.2
	RPST 7000	26.9
	RPST 10000	25.6
	RPST	25.1
	CLASICO	34.1

THETA	TIPO	SIGNI
-0.75*Tr	Triangular	56.0
	RPST 4000	54.9
	RPST 4500	57.5
	RPST 5000	56.4
	RPST 5100	55.9
	RPST 5200	56.8
	RPST 5300	56.1
	RPST 5305	55.9
	RPST 5400	55.8
	RPST 5500	56.1
	RPST 6000	56.1
	RPST 7000	54.8

# ANEXOS

---

	RPST 10000	55.3
	RPST	54.9
	CLASICO	65.5
-1	*Tr Triangular	80.5
	RPST 4000	79.8
	RPST 4500	81.4
	RPST 5000	81.7
	RPST 5100	81.5
	RPST 5200	80.8
	RPST 5300	81.2
	RPST 5305	81.9
	RPST 5400	80.9
	RPST 5500	81.0
	RPST 6000	81.9
	RPST 7000	81.1
	RPST 10000	81.3
	RPST	81.6
	CLASICO	88.0
-1.25	*Tr Triangular	93.8
	RPST 4000	93.9
	RPST 4500	94.7
	RPST 5000	94.6
	RPST 5100	94.6
	RPST 5200	94.4
	RPST 5300	94.7
	RPST 5305	94.6
	RPST 5400	94.9
	RPST 5500	94.7
	RPST 6000	94.7
	RPST 7000	94.9
	RPST 10000	94.9
	RPST	95.0
	CLASICO	97.9
-1.50	*Tr Triangular	98.8
	RPST 4000	98.7
	RPST 4500	98.8
	RPST 5000	99.0
	RPST 5100	98.9
	RPST 5200	99.1
	RPST 5300	98.8
	RPST 5305	98.8
	RPST 5400	98.9
	RPST 5500	99.1
	RPST 6000	99.1
	RPST 7000	98.9
	RPST 10000	99.1

## ANEXOS

---

RPST	98.8
CLASICO	99.9



## ANEXOS

---

```
data (tp2(i),i=1,nThetas) & ! Trat B
  / 0.1513, 0.1414, 0.1321, 0.1234, 0.1151, 0.1073, &
    0.1 , 0.09313,0.08669,0.08065,0.075,0.06971, 0.06478 /

DO i=1,nThetas ! Trat B
  tp1(i) =.1
END DO

nrep=10000
nFix=3962

valor = 1
call set_seed(valor)

! Inicio estudio
!   Trat A   Trat B
!   -----
!   A   |   C   AC   exito   A = Exitos   en A
!   -----
!   B   |   D   BD   fracaso  C= Exitos   en B
!   -----
!   AB           CD   T
!   -----

open(10,file='todos.dat',STATUS='new')
```

```

do iTipo = 1, nTipos + 1
  Vertice=Ordo(iTipo)/PendI(iTipo)
  do icomp = 1, nThetas
    pA=Tp1(icom)
    pB=Tp2(icom)
    DO irepe = 1,nrep
      A=0.; B=0.; C=0.; D=0.
      AB=0.; AC=0.; BD=0.; CD=0.; T=0.
      CONT=0.; VLAST=0.
      FIN=0;
      DO while (fin.eq.0)
        IF (random() .gt. pA ) then
          A=A+1. ! exito A
        else
          B=B+1. ! fracaso A
        END IF
        IF (random() .gt. pB ) then
          C=C+1. ! exito B
        else
          D=D+1. ! fracaso B
        END IF
        AB = A + B; AC = A + C; BD = B + D; CD = C + D; T = AB + CD

        Z = -1.*(B*CD - D*AB)/T; V = AB/T*AC/T*BD/T*CD
        CONT = 0.583*SQRT(V - VLAST); VLAST = V

        IF (iTipo .eq. nTipos+1) then ! diseño N fija
          IF (T.ge.Nfix) then
            sqrtV = sqrt(V) * 1.96
            IF (Z.LE.0) then
              IF (Z.LE.-1.*sqrtV) FIN=8
              IF (Z.GT.-1.*sqrtV) FIN=5
            ELSE
              IF (Z.GE.sqrtV) FIN=1
              IF (Z.LT.sqrtV) FIN=4
            END IF
          END IF
        ELSE ! diseños secuenciales
          BExterna =PendS(iTipo)*V+Ordo(iTipo)-CONT
          BInterna =PendI(iTipo)*V-Ordo(iTipo)+CONT
          IF ((V.LE.Vmax(iTipo)) .or. (iTipo.eq.1) .or. (iTipo.eq.14) ) THEN
            IF (Z .GT. BExterna ) then
              FIN = 1
            ELSE IF (Z .LT. -1.* BExterna ) then
              FIN = 8
            ELSE IF (V .GT. Vertice) THEN
              IF (Z.GT.0) THEN
                IF (Z.LT.BInterna) FIN = 4
              ELSE
                IF (Z.GT.-1*BInterna) FIN = 5
              END IF
            END IF
          END IF
        END IF
      END DO
    END DO
  END DO
END DO

```

## ANEXOS

---

```

        END IF
    ELSE
        ! Diseños truncados que
llegan a Vmax
        IF (Z .GE. PendS(iTipo)*V ) then
            FIN = 2
        ELSE IF ( (Z .LT. PendS(iTipo)*V ) .AND. (Z .GT. (PendS(iTipo)*V - Ordo(iTipo))) )
then
        ELSE IF ( (Z .LT. PendS(iTipo)*V ) .AND. (Z .GT. 0) ) then
            FIN = 3
        ELSE IF ( (Z .LT. (-1.*PendI(iTipo)*V + Ordo(iTipo))) .AND. (Z .GT. -
1.*PendI(iTipo)*V )) THEN
            FIN = 6
        ELSE IF (Z .LT. -1.*PendI(iTipo)*V ) then
            FIN = 7
        END IF
    END IF
END IF
END IF
if (t.ge.50000) then
    stop
!
!2
!
!
    write (6,2)int(t),z,v,Vertice,BInterna,Vmax(iTipo),ibucle1,ibucle2,ibucle3,ibucle4
    format(i8,2f10.3,3f9.3,10i3)
    print *,PendS(iTipo)*V, (PendS(iTipo)*V - Ordo(iTipo))
    print *, ''
END IF

END DO

write (10,'(i3,i3,i8,2f10.3,i3)')itipo,icomp,int(t),z,v,fin
END DO
END DO
END DO
END
```

**PROGRAMAS EN SAS UTILIZADOS PARA LEER Y ANALIZAR  
LOS DATOS SIMULADOS****LECTURA DE DATOS SIMULADOS**

```
libname simul '.';

proc format;
  value fin_fmt 1 = 'A *'
                2 = 'A '
                3 = 'B '
                4 = 'B *';
  value finfmt 1 = '++++A *'
                2 = ' +++A *'
                3 = '  ++A '
                4 = '   +A '
                5 = '    +B '
                6 = '   ++B '
                7 = '  +++B *'
                8 = '++++B *' ;
  value icompft 1 = '-1.50*Tr'
                2 = '-1.25*Tr'
                3 = '-1   *Tr'
                4 = '-0.75*Tr'
                5 = '-0.5  *Tr'
                6 = '-0.25*Tr'
                7 = '+0   *Tr'
                8 = '+0.25*Tr'
                9 = '+0.5  *Tr'
                10 = '+0.75*Tr'
                11 = '+1   *Tr'
                12 = '+1.25*Tr'
                13 = '+1.5  *Tr';
  value iTipoft 1 = 'Triangular'
                2 = 'RPST 4000'
                3 = 'RPST 4500'
                4 = 'RPST 5000'
                5 = 'RPST 5100'
                6 = 'RPST 5200'
                7 = 'RPST 5300'
                8 = 'RPST 5305'
                9 = 'RPST 5400'
```

```
10= 'RPST 5500'  
11= 'RPST 6000'  
12= 'RPST 7000'  
13= 'RPST 10000'  
14= 'RPST '  
15= 'CLASICO';
```

```
data simul.todos;  
  infile 't:\tim\work\sim\sas\todos.dat' missover lrecl=80;  
  input itipo  
        icoomp  
        t  
        z  
        v  
        fin;  
  if      1 le fin le 2 then fin_ =1;  
  else if 3 le fin le 4 then fin_ =2;  
  else if 5 le fin le 6 then fin_ =3;  
  else if 7 le fin le 8 then fin_ =4;  
  if (fin_ =1) or (fin_ =4) then signi =1;  
  else signi =0;  
  
  if      (t gt 3962) then gtNfix = 1;  
  else gtNfix = 0;  
  nrep = 1;  
  DifNfix = 3962 - T;  
  label iTipo      = 'Tipo Estudio'  
        iComp      = 'Comparacion'  
        Fin        = 'Lugar de Corte'  
        Fin_       = 'Lugar de Corte (cod)'  
        DifNfix    = 'Nfija - T'  
        gtNfix     = 'N gt Nfija'  
        SIGNI      = 'p<= 0.05';  
  
  format fin finfmt. fin_ fin_fmt. icoomp icoompft. itipo itipoft.;  
run;  
  
run;
```

*ANÁLISIS DE LOS DATOS SIMULADOS*

```
libname simul '.';

proc format;
  value fin_fmt 1 = 'A *'
                2 = 'A '
                3 = 'B '
                4 = 'B *';
  value finfmt 1 = '++++A *'
                2 = ' +++A *'
                3 = '  ++A '
                4 = '   +A '
                5 = '    +B '
                6 = '   ++B '
                7 = '  +++B *'
                8 = '++++B *';
  value icompft 1 = '-1.50*Tr'
                2 = '-1.25*Tr'
                3 = '-1   *Tr'
                4 = '-0.75*Tr'
                5 = '-0.5  *Tr'
                6 = '-0.25*Tr'
                7 = '+0   *Tr'
                8 = '+0.25*Tr'
                9 = '+0.5  *Tr'
                10 = '+0.75*Tr'
                11 = '+1   *Tr'
                12 = '+1.25*Tr'
                13 = '+1.5  *Tr';
  value iTipoft 1 = 'Triangular'
                2 = 'RPST 4000'
                3 = 'RPST 4500'
                4 = 'RPST 5000'
                5 = 'RPST 5100'
                6 = 'RPST 5200'
                7 = 'RPST 5300'
                8 = 'RPST 5305'
                9 = 'RPST 5400'
                10 = 'RPST 5500'
                11 = 'RPST 6000'
                12 = 'RPST 7000'
                13 = 'RPST 10000'
                14 = 'RPST ';
```

## ANEXOS

---

```
15= 'CLASICO';

format fin finfmt. fin_ fin_fmt. icomp icompft. iTipo tipoft.;
/*
proc print data=simul.todos; run;
*/

proc univariate data=simul.todos noprint;
  by iTipo iComp;
  var t;
  output out=simul.SIZE
          mean      =Nmean
          min       =Nmin
          max       =Nmax
          Median    =P50
          p90      =P90;
proc univariate data=simul.todos noprint;
  by iTipo iComp;
  var nrep gtNfix;
  output out=simul.AHORRO sum = nrep Ahorro;
proc univariate data=simul.todos noprint;
  by iTipo iComp;
  var nrep signi;
  output out=simul.CO sum = nrep signi;

proc univariate data=simul.todos noprint;
  by iTipo iComp;
  var DifNfix;
  output out=simul.DNfix
          mean      =Nmean
          min       =Nmin
          max       =Nmax
          Median    =P50
          p90      =P90;

data simul.ahorro; set simul.ahorro; ahorro= 100 - (ahorro/nrep *100); drop nrep;
data simul.co; set simul.co; signi= (signi/nrep *100); drop nrep;

proc print data=simul.size;
proc print data=simul.co; format signi 5.1 ;
proc print data=simul.ahorro; format ahorro 5.1 ;
proc print data=simul.DNfix;

run;
```

## Bibliografía

1. Arós F, Marrugat J, Bayón J. Datos epidemiológicos y fisiopatología del infarto agudo de miocardio. *Rev Esp Cardiol* 1994; 47 (supl. 1):S13-16.
2. Hennekens C.H., Albert C.M., Godfried SL, et al. Adjunctive drug therapy of acute myocardial infarction- Evidence from clinical trials. *N Engl J Med* 1996; 335:1660-1667.
3. Yusuf S, Anand S, Avezum A, Jr., Flather M, Coutinho M. Treatment for acute myocardial infarction. Overview of randomized clinical trials. *Eur.Heart J.* 1996; 17 Suppl F:16-29.
4. Glogar D, Yang P, Steurer G. Management of acute myocardial infarction: evaluating the past, practicing in the present, elaborating the future. *Am Heart J* 1996; 132 (Suppl2):465-470.
5. Rappaport E, Gheorghide M. Pharmacologic therapies after myocardial infarction. *Am J Med* 1996; 101 (Suppl 4A):61s-69s.
6. Fuster V., Badimon L, Badimon JJ, Chesebri JH. The pathogenesis of coronary artery disease and the acute coronary syndromes. (part I). *N Engl J Med* 1992; 326:242-250.
7. Fuster V., Badimon L, Badimon JJ. The pathogenesis of coronary artery disease and the acute coronary syndromes. (part II). *N Engl J Med* 1992; 326:310-318.
8. Almony GT, Lefkovits J, Topol EJ. Antiplatelet and anticoagulant use after myocardial infarction. *Clin Cardiol* 1996; 19:357-365.

9. Badimon L., Martínez J, Badimon JJ. Hemostasia primaria. In: Castillo J., Martínez-Vila E, editors. Trombosis, fármacos antitrombóticos y enfermedad cerebrovascular. Barcelona: J Uriach & Cia, 1995:
10. Antiplatelet Trialists' Collaboration. Collaborative overview of randomised trials of antiplatelet therapy. *Br Med J* 1994; 308:159-168.
11. Schör K. Antiplatelet drugs. A comparative review. *Drugs* 1995; 50:7-28.
12. Rimbau V, López R, Fernandez A, Forn J. Estudio farmacocinético del triflusal (UR-1501). *Arch Farm Tox* 1981; 7:11-16.
13. García-Rafanell J, Manresa G, Miserach N, Ventura A. Estudio del efecto del triflusal (UR-1501) sobre la agregación plaquetaria y la coagulación plasmática en voluntarios sanos. *Med Clin.(Barc)*. 1978; 71:397-401.
14. Albors M, Castellarnau C, Vila L, Sola J, Rutllan M. Inhibition of thromboxane production and platelet function by triflusal in healthy volunteers. *Rev Farmacol.Clin.Exp.* 1987; 4:11-16.
15. Sánchez de Miguel L, Millas I, García-Durán M, et al. New mechanisms for the antiaggregating platelet effects of triflusal: role of nitric oxide and neutrophils. *Eur J Cancer* 1997; 18 (suppl):19
16. Boyssen G, Boss AH, Odum N, Olsen JS. Prolongation of bleeding time and inhibition of platelet aggregation by low dose acetylsalicylic acid in patients with cerebrovascular disease. *Stroke* 1984; 15:241-243.
17. Putz P, Buyse H, Delvaux D, Kutnowskui M, Demulder A, Dumont N, et al. Triflusal versus acetylsalicylic acid: a double blind study for the prophylaxis of deep vein thrombosis after hip surgery. *Acta Chir Belg* 1991; 6:269-276.

18. Bakke OM, Carné X, García Alonso F. Ensayos clínicos con medicamentos: fundamentos básicos, metodología y práctica. Barcelona: Ediciones Doyma, S.A., 1994.
19. Pocock SJ. Clinical trials: a practical approach. New York: Wiley, 1982.
20. Spilker B. Guide to clinical trials. New York: Raven Press Ltd, 1991.
21. Armitage P, Berry G. Estadística para la investigación biomédica . 2 ed. Barcelona: Ediciones Doyma, 1992.
22. Carné X, Moreno V, Porta-Serra M, Velilla E. El cálculo del número de pacientes necesarios en la planificación de un estudio clínico. Med Clin.(Barc). 1989; 92:72-77.
23. Freiman JA, Chalmers TC, Smith H, Kuebler RR. The importance of beta, the type II error and sample size in the design and interpretation of the randomized controlled trials. N Engl J Med 1978; 299:690-694.
24. Gail M, Gart JJ. The determination of sample sizes for use with the exact conditional test in 2x2 comparative trials. Biometrics. 1973; 29:441-448.
25. Haseman JK. Exact sample sizes for use with the Fisher-Irwin test for 2x2 tables. Biometrics. 1978; 34:106-109.
26. Schlesselman JJ. Case-control studies. New York: Oxford University Press, 1982.
27. Gail MH. Applicability of sample size calculations based on a comparison of proportions for use with the logrank test. Controlled Clinical.Trials. 1985; 6:112-119.
28. Lemeshow S, Hosmer DW, Jr., Klar J. Sample size requirements for studies estimating odds ratios or relative risks. Stat.Med 1988; 7:759-764.

29. Lubin JH, Gail MH, Ershow AG. Sample size and power for case-control studies when exposures are continuous. *Stat.Med* 1988; 7:363-376.
30. Muhm JM, Olshan AF. A program to calculate sample size, power and least detectable relative risk using a programmable calculator. *Am.J Epidemiol.* 1989; 129:205-211.
31. O'Neil RT. On sample sizes to estimate the protective efficacy of a vaccine. *Stat.Med* 1988; 7:1279-1288.
32. Schlesselman JJ. Sample size requirements in cohort and case-control studies of disease. *Am.J Epidemiol.* 1974; 99:381-384.
33. Walter SD, Morgenstern H. A note on optimal sampling for the comparison of proportions or rates. *Stat.Med* 1985; 4:541-542.
34. Woolson RF, Bean JA, Rojas PB. Sample size for case-control studies using Cochran's statistic. *Biometrics.* 1986; 42:927-932.
35. Dupont WD, Plummer WD. Power and sample size calculations: a review and computer program. *Control.Clin.Trials.* 1990; 11:116-128.
36. Machin D, Campbell MJ. *Statistical tables for the design of clinical trials.* Oxford: Blackwell Scientific Publications, 1987.
37. Garbe E, Röhmel J, Gundert-Remy U. Clinical and statistical issues in therapeutic equivalence trials. *Eur J Clin.Pharmacol.* 1993; 45:1-7.
38. Makuch R, Johnson M. Issues in planning and interpreting active control equivalence studies. *J Clin.Epidemiol.* 1989; 42:503-511.
39. Ng T-H. A specification of treatment difference in the design of clinical trials with active controls. *Drug Information Journal* 1993; 27:705-719.

## BIBLIOGRAFIA

---

40. Blackwelder WC. Similarity/Equivalence Trials for combination vaccines. *Annals New York Academy of Sciences* 9999; 321-328.
41. Victor N. On clinically relevant differences and shifted nullhypotheses. *Meth.Inform.Med.* 1987; 26:109-116.
42. Jones B, Jarvis P, Lewis JA, Ebbutt AF. Trials to assess equivalence: the importance of rigorous methods [see comments]. *BMJ.* 1996; 313:36-39.
43. Kirshner B. Methodological standards for assessing therapeutic equivalence. *J Clin.Epidemiol.* 1991; 44:839-849.
44. Roebruck P, Kuhn A. Comparison of tests and sample size formulae for proving therapeutic equivalencebased on the difference of binomial probabilities. *Stat.Med.* 1995; 14:1583-1594.
45. Kingman A. Statistical vs clinical significance in product testing: can they be designed to satisfy equivalence? *KIngman, A* 1992; 52:353-360.
46. Ware JH, Antman EM. Equivalence trials. *N.Engl.J.Med.* 1997; 337:1159-1161.
47. Zanen Z, Lammers J-WJ. Sample sizes for comparative inhaled corticosteroid trials with emphasis on showing therapeutic equivalence. *Eur J Clin.Pharmacol.* 1995; 48:179-184.
48. Schuirmann DJ. A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *J Pharmacokin.Biopharm.* 1987; 15:657-680.
49. Langman MJS. Towards estimation and confidence intervals. *Br Med J* 1986; 292:716-,estimacio.

## BIBLIOGRAFIA

---

50. McHugh RB, Le CT. Confidence estimation and the size of a clinical trial. *Controlled Clinical.Trials*. 1984; 5:157-163.
51. Daly LE. Confidence intervals and sample sizes: don't throw out all your old sample sizetables [see comments]. *BMJ*. 1991; 302:333-336.
52. Braitman LE. Confidence intervals extract clinically useful information from data. *Ann Int Med* 1988; 108:296-298.
53. Braitman LE. Confidence intervals assess both clinical significance and statistical significance. *Ann Int Med* 1991; 114:515-517.
54. Westlake WJ. Symmetrical confidence intervals for bioequivalence trials. *Biometrics*. 1976; 32:741-744.
55. Anderson S, Hauck WW. A new procedure for testing equivalence in comparative trials. *Commun Statist-Theor Meth* 1983; 12:2663-2692.
56. Hauck WW, Anderson S. A new statistical procedure for testing equivalence in two-group comparative bioavailability trials. *J Pharmacokin.Biopharm*. 1984; 12:83-91.
57. Hauck WW, Anderson S. A proposal for interpreting and reporting negative studies. *Stat.Med* 1986; 5:203-209.
58. Owen DB. A especial case of the bivariate non-central t-distribution. *Biometrika*. 1965; 52:437-446.
59. Phillips KF. Power of the two one-sided tests procedure in bioequivalence. *J Pharmacokin.Biopharm*. 1990; 18:37-143.
60. Diletti E, Hauschke D, Steinijans VW. Sample size determination for bioequivalence assessment by means of confidence intervals. *Int.J.Clin.Pharmacol.Ther.Toxicol*. 1991; 29:1-8.

61. Diletti E, Hauschke D, Steinijans VW. Sample size determination: extended tables for the multiplicative model and bioequivalence ranges of 0.9 to 1.11 and 0.7 to 1.43. *Int.J.Clin.Pharmacol.Ther.Toxicol.* 1992; 30 (Suppl. N° 1):S59-62.
62. Liu JP, Chow SC. Sample size determination for the two one-sided test procedure in bioequivalence. *J Pharmacokin.Biopharm.* 1992; 20:101-104.
63. Hauschke D, Steinijans VW, Diletti E, Burke M. Sample size determination for bioequivalence assessment using a multiplicative model. *J Pharmacokin.Biopharm.* 1992; 20:557-561.
64. Peace KE. Some thoughts on one-tailed tests. *Biometrics.* 1988; 44:911-912.
65. Steinijans VW, Diletti E. Statistical analysis of bioavailability studies: parametric and nonparametric confidence intervals. *Eur J Clin.Pharmacol.* 1983; 24:127-136.
66. Hauschke D, Steinijans VW, Diletti E. A distribution-free procedure for the statistical analysis of bioequivalence studies. *Int.J.Clin.Pharmacol.Ther.Toxicol.* 1990; 28:72-78.
67. Dunnett CW, Gent M. Significance testing equivalence between treatments, with special reference to data in the form of 2 x 2 tables. *Biometrics.* 1977; 33:593-602.
68. Blackwelder WC. "Proving the null hypothesis" in clinical trials. *Controlled Clinical.Trials.* 1982; 3:345-353.
69. Makuch R, Simon R. Sample size for evaluating a conservative therapy. *Cancer Treatment.Reports.* 1978; 62:1037-1040.

## BIBLIOGRAFIA

---

70. Blackwelder WC, Chang MA. Sample size graph for "proving the null hypothesis". *Controlled Clinical.Trials.* 1984; 5:97-105.
71. Heiselbetz C, Edler L. A sample size program for "proving the null hypothesis". *Controlled Clinical.Trials.* 1987; 8:45-48.
72. Lee ML, Lusher JM. The problem of therapeutic equivalence with paired qualitative data: an example from a clinical trial using haemophiliacs with an inhibitor to factor VIII. *Stat.Med.* 1991; 10:433-441.
73. Kieser M. A confirmatory strategy for therapeutic equivalence trials. *Int.J.Clin.Pharmacol.Ther.* 1995; 33:388-390.
74. Rodda BE, Davis RL. Determining the probability of an important difference in bioavailability. *Clin.Pharmacol.Ther.* 1980; 28:247-252.
75. Mandallaz D, Mau J. Comparison of different methods for decision-making in bioequivalence assessment. *Biometrics.* 1981; 37:213-222.
76. Selwyn MR, Dempster AP, Hall NR. A bayesian approach to bioequivalence for the 2x2 changeover design. *Biometrics.* 1981; 37:11-21.
77. Selwyn MR, Hall NR. On bayesian methods for bioequivalence. *Biometrics.* 1984; 40:1103-1108.
78. Mau J. A statistical assessment of clinical equivalence. *Stat.Med* 1988; 7:1267-1277.
79. Anderson S, Hauck WW. Consideration of individual bioequivalence. *J Pharm Pharmacokinet* 1990; 18:259-273.
80. Ashby D, Machin D. Stopping rules, interim analyses and data monitoring committees [editorial]. *Br.J.Cancer* 1993; 68:1047-1050.

81. Task force of the working group on arrhythmias of the European Society of Cardiology. The early termination of clinical trials: causes, consequences, and control. With special reference to trials in the field of arrhythmias and sudden death. Task Force of the Working Group on Arrhythmias of the European Society of Cardiology. *Eur.Heart J.* 1994; 15:721-738.
82. Armitage P, McPherson CK, Rowe BC. Repeated significance tests on accumulating data. *J.Roy.Statist.Soc.* 1969; 132:235-244.
83. Paykel ES. Interim analyses and publication of clinical trials [letter]. *Arch Gen Psychiatry* 1986; 43:613
84. Wald A. *Sequential Analysis*. New York: Wiley, 1947.
85. Armitage P. *Sequential Medical Trials*. 2 ed. Oxford: Blackwell, 1975.
86. Kilpartick GS, Oldham PD. Calcium chloride and adrenaline as bronchial dilators compared by sequential analysis. *Br Med J* 1954; 2:1388-1391.
87. Freireich EJ, Gehan E, Frei E. The effect of 6-Mercaptopurine on the duration of steroid-induced remissions in acute leukemia: A model for the evaluation of other potential useful therapy. *Blood* 1966; 21:699-716.
88. Bellissant E, Duhamel JF, Guillot M, Pariente-Khayat A, Olive G, Pons G. The triangular test to assess the efficacy of metoclopramide in gastroesophageal reflux. *Clin.Pharmacol.Ther.* 1997; 61:377-384.
89. CONSENSUS Trial Study Group. Effects of enalapril on mortality in severe congestive heart failure: results of the cooperative North Scandinavian enalapril survival study (CONSENSUS). *N.Engl.J.Med.* 1987; 316:1429-1435.
90. Storb R, Deeg HJ, Whitehead J. Methotrexate and cyclosporine compared with cyclosporine alone for prophylaxis of acute graft versus host disease

- after marrow transplantation for leukemia. *N Engl J Med* 1986; 314:729-735.
91. Jones DR, Newman CE, Whitehead J. The design of a sequential clinical trial for the comparison of two lung cancer treatments. *Stat.Med* 1982; 1:73-82.
92. Abramowitz MD, OH TH, Epstein BS, Ruttimann U, Friendly D. The antiemetic effect of droperidol following outpatient strabismus surgery in children. *Anesthesiology* 1983; 59:579-583.
93. Cardiac arrhythmia suppression trial investigators. Effect of encainide and flecainide on mortality in a randomised trial of arrhythmia suppression after myocardial infarction. *N Engl J Med* 1989; 321:406-412.
94. Coronary drug project research group. Practical aspects of decision making in clinical trials: the coronary drug project as a case study. *Control.Clin.Trials*. 1981; 1:363-376.
95. Dillman RO, Seagren SL, Propert KJ, Guerra J, Eaton WL, Perry MC, et al. A randomised trial of induction chemotherapy plus high-dose radiation alone in stage III non-small-cell lung cancer. *N Engl J Med* 1990; 323:940-945.
96. Facey KM. A sequential procedure for a phase II efficacy trial in hypercholesteremia. *Control.Clin.Trials*. 1992; 13:122-133.
97. Fischl MA, Richman DD, Grieco MH, et al. The efficacy of azidothymidine (AZT) in the treatment of patient with AIDS and AIDS-related complex: double-blind placebo-controlled trial. *N Engl J Med* 1987; 317:185-191.
98. Gortner L, Bernsa U. Does prophylactic use of bovine surfactant change drug utilisation in very premature infants during neonatal period? *Developmental Pharmacology and Therapeutics* 1991; 16:1-6.

99. Hackett GH, Harris MNE, Plantevin M, Pringle HM, Garrioch DB, Avery A. Anaesthesia for out-patient termination of pregnancy, a comparison of two anaesthetic techniques. *British Journal of Anaesthesia* 1982; 54:865-870.
100. Hardy RJ, Davis BR. Statistical considerations in terminating randomisations in the multicenter trial of cryotherapy for retinopathy of prematurity. *Control.Clin.Trials*. 1991; 12:293-303.
101. Kirwan JR, Haskard DO, Higgins CS. The use of sequential analysis to assess patient preference for local skin anaesthesia during knee aspiration. *British Journal of Rheumatology* 1984; 23:210-213.
102. Lewis HE, et al. Protective effects of aspirin against acute myocardial infarction and death in men with unstable angina. *N Engl J Med* 1983; 309:396-403.
103. Montaner JSG, et al. Corticosteroids prevent early deterioration in patients with moderately severe pneumocystis carinii pneumonia and the Acquired Immunodeficiency Syndrome (AIDS). *Ann Int Med* 1990; 113:14-20.
104. Pawitan Y, Hallstrom A. Statistical interim monitoring of the Cardiac Arrhythmia Suppression Trial. *Stat.Med.* 1990; 9:1081-1090.
105. Rubens RD, Bartelink H, Engelsman E, Hayward JL, Rotmensz N, Sylvester R, et al. Locally advanced breast cancer: the contribution of cytotoxic and endocrine treatment to radiotherapy. *European Journal of Cancer and Clinical Oncology* 1989; 25:667-668.
106. Snell ES, Armitage P. Clinical comparison of diamorphine and pholcodine as cough suppressants, by a new method of sequential analysis. *Lancet* 1957; 1:860-862.

## BIBLIOGRAFIA

---

107. Sylvester R, Bartelink H, Rubens R. A reversal of fortune: practical problems in the monitoring and interpretation of an EORTC breast cancer trial. *Stat.Med.* 1994; 13:1329-1355.
108. van Es GA, Tijssen JG, Lubsen J, van Strik R. Early termination of clinical trials with prolonged observation of individual patients: a case study. *Stat.Med* 1987; 6:927-937.
109. Whitehead J. Sequential designs for pharmaceutical clinical trials. *Pharmaceutical Medicine* 1992; 6:179-191.
110. Whitehead J. Application of sequential methods to a phase III clinical trial in stroke. *Drug Information Journal* 1993; 27:733-740.
111. Whitehead J, Jones DR, Ellis SH. The analysis of a sequential clinical trial for the comparison of two lung cancer treatments. *Stat.Med.* 1983; 2:183-190.
112. Whitehead J. The design and analysis of sequential clinical trials. 2 ed. Chichester: Ellis Horwood, 1992.
113. Bellissant E, Benichou J, Chastang C. The group sequential triangular test for phase II cancer clinical trials. *Am.J.Clin.Oncol.* 1996; 19:422-430.
114. Brunier H, Whitehead J. PEST3.0 Operating Manual. Reading: Reading University, 1993.
115. DeGroot MH. Probability and statistics. 2 ed. Reading: Addison-Wesley, 1986.
116. Stein C. A two-sample test for a linear hypothesis whose power is independent of the variance. *Ann Math Stat* 1945; 16:243-258.

117. Gould AL. Interim analyses for monitoring clinical trials that do not materially affect the type I error rate. *Stat.Med.* 1992; 11:55-66.
118. Gould AL, Shih WJ. Sample size re-estimation without unblinding for normally distributed data with unknown variance. *Commun Statist-Theor Meth* 1992; 21:2833-2853.
119. Peace KE. *Biopharmaceutical sequential statistical applications*. New York: Marcel Dekker, Inc., 1992.
120. Armitage P. Sequential methods in clinical trials. *Amer.J.Pub.Health* 1958; 48:1395-1402.
121. Pocock SJ. Group sequential methods in the design and analysis of clinical trials. *Biometrika.* 1977; 64:191-199.
122. O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. *Biometrics.* 1979; 35:549-556.
123. Pocock SJ. Interim analyses for randomized clinical trials: the group sequential approach. *Biometrics.* 1982; 38:153-162.
124. Machin D. Interim analysis and ethical issues in the conduct of trials. In: Williams CJ, editor. *Introducing new treatments for cancer: practical, ethical and legal problems*. Wiley, 1992:
125. Peto R, Pike MC, Armitage P. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. *Br J Cancer* 1976; 34:585-612.
126. DeMets DL, Ware JH. Group sequential methods for clinical trials with a one-sided hypothesis. *Biometrika.* 1980; 67:651-660.

127. Emerson SS, Fleming TR. Symmetric group sequential test designs. *Biometrics*. 1989; 45:905-923.
128. Jennison C. Efficient group sequential tests with unpredictable group sizes. *Biometrika*. 1987; 74:155-165.
129. Kim K. SEQPWR and SEQOPR: computer programs for design of maximum information trials based on group sequential logrank tests. *Comput.Methods Programs Biomed*. 1995; 46:143-153.
130. Cytel Software Corporation. EaSt: A software package for the design and interim monitoring of group sequential trials. Cambridge MA, 1992.
131. Lewis RJ. An introduction to the use of interim data analyses in clinical trials. *Ann.Emerg.Med*. 1993; 22:1463-1469.
132. Pocock SJ, Geller NL. Interim analysis in randomized clinical trials. *Drug Information Journal* 1986; 20:263-270.
133. Lan KKG, DeMets DL. Discrete sequential boundaries for clinical trials. *Biometrika*. 1983; 70:659-663.
134. Lan KK, DeMets DL. Changing frequency of interim analysis in sequential monitoring. *Biometrics*. 1989; 45:1017-1020.
135. Kim K, DeMets DL. Design and analysis of group sequential tests based on the type I error spending rate function. *Biometrika*. 1987; 74:149-154.
136. Lan KKG, DeMets DL. Group sequential procedures: calendar versus information time. *Stat.Med*. 1989; 8:1191-????
137. Jennison C, Turnbull BW. Interim analysis: the repeated confidence interval approach. *J.Roy.Statist.Soc*. 1989; B51:305-361.

138. Slud EV, Wei LJ. Two-sample repeated significance test based on the modified Wilcoxon statistic. *J Am.Statistical.Association.* 1982; 77:862-868.
139. Jennison C, Turnbull BW. Sequential equivalence testing and repeated confidence intervals, with applications to normal and binary responses. *Biometrics.* 1993; 49:31-43.
140. Wang SK, Tsiatis AA. Approximately optimal one-parameter boundaries for group sequential trials. *Biometrics.* 1987; 43:193-199.
141. Pampallona S, Tsiatis AA. Group sequential designs for one-sided and two-sided hypothesis testing with provision for early stopping in favor of the null hypothesis. *Journal of Statistical Planning and Inference* 1994; 42:19-35.
142. Barnard GA. Sequential test in industrial statistics. *J.Roy.Statist.Soc.* 1946; Suppl.8:S1-26.
143. Anscombe FJ. Sequential medical trials. *J.Amer.Statist.Assoc.* 1963; 58:365-383.
144. Whitehead J, Jones D. The analysis of sequential clinical trials. *Biometrika.* 1979; 66:443-452.
145. Whitehead J. Interim analyses and stopping rules in cancer clinical trials. *Br.J.Cancer* 1993; 68:1179-1185.
146. Wald A, Wolfowitz J. Optimal character of the sequential probability ratio test. *Ann.Math.Stat.* 1948; 19:326-339.
147. Armitage P. Restricted sequential procedures. *Biometrika.* 1957; 44:9-26.
148. Anderson TW. A modification of the sequential probability ratio test to reduce sample size. *Ann.Math.Stat.* 1960; 31:165-197.

149. Whitehead J, Stratton I. Group sequential clinical trials with triangular continuation regions. *Biometrics*. 1983; 39:227-236.
150. Proschan MA, Follmann DA, Waclawiw MA. Effects of assumption violations on type I error rate in group sequential monitoring. *Biometrics*. 1992; 48:1131-1143.
151. Sellke T, Siegmund D. Sequential analysis of the proportional hazards model. *Biometrika*. 1983; 70:315-326.
152. Lan KKG, Zucker DM. Sequential monitoring of clinical trials: the role of information and Brownian motion. *Stat.Med*. 1993; 12:753-765.
153. Fairbanks K, Madsen R. P values for tests using a repeated significance test design. *Biometrika*. 1982; 69:69-74.
154. Rosner GL, Tsiatis AA. Exact confidence limits following group sequential tests. *Biometrika*. 1988; 75:723-729.
155. Chang MN. Confidence intervals for a normal mean following a group sequential test. *Biometrics*. 1989; 45:247-254.
156. Emerson SS, Fleming TR. Parameter estimation following group sequential hypothesis testing. *Biometrika*. 1990; 77:875-892.
157. Whitehead J. On the bias of maximum likelihood estimation following a sequential test. *Biometrika*. 1986; 73:573-581.
158. Anderson TW. Sequential analysis with delayed observations. *J Am.Statistical.Association*. 1964; 59:1006-1015.
159. Choi SC, Clark VA. Sequential decision for a binomial parameter with delayed observations. *Biometrics*. 1970; 26:411-420.

160. Whitehead J. Overrunning and underrunning in sequential clinical trials. *Control.Clin.Trials*. 1992; 13:106-121.
161. Whitehead J. Supplementary analysis at the conclusion of a sequential clinical trial. *Biometrics*. 1986; 42:461-471.
162. Jennison C, Turnbull BW. Repeated confidence intervals for group sequential clinical trials. *Controlled Clin.Trials*. 1984; 5:33-45.
163. Jennison C, Turnbull BW. Statistical approaches to interim monitoring of medical trials: A review and a commentary. *Statistical Science* 1990; 5:299-317.
164. Robbins H. Statistical methods related to the law of the iterated logarithm. *Ann.Math.Stat.* 1970; 41:1397-1409.
165. Lan KKG, Simon R, Halperin M. Stochastically curtailed tests in long-term clinical trials. *Sequential Analysis* 1982; 1:207-219.
166. Lindley DV. Part 2, Inference. In: Anonymous Introduction to probability and statistics from a bayesian viewpoint. Cambridge: Cambridge University Press, 1965:
167. Barnett V. Comparative statistical inference. 2 ed. Chichester: Wiley, 1982.
168. Spiegelhalter DJ. Statistical aids in clinical decision making. *Statistician* 1982; 31:19-36.
169. Racine A, Grieve P, Flühler H, Smith AFM. Bayesian methods in practice: experiences in the pharmaceutical industry. *Appl.Statist.* 1986; 35:93-150.
170. Freedman LS, Spiegelhalter DJ. Response to letter to the editor. *Controlled Clin.Trials*. 1991; 12:346-350.

## BIBLIOGRAFIA

---

171. Kober L, Torp-Pedersen C, Cole D, Hampton JR, Camm AJ. Bayesian interim statistical analysis of randomised trials: the case against. *Lancet* 1997; 349:1168-1169.
172. Brophy JM, Joseph L. Bayesian interim statistical analysis of randomised trials. *Lancet* 1997; 349:1166-1168.
173. Armitage P, Berry G. *Statistical methods in medical research*. Oxford: Blackwell, 1992.
174. Etzioni RD, Kadane JB. Bayesian statistical methods in public health and medicine. *Annu.Rev.Public Health* 1995; 16:23-41.
175. Spiegelhalter DJ, Freedman LS, Parmar MKB. Bayesian approaches to randomized trials. *J.Roy.Statist.Soc.* 1994; 157:357-416.
176. Grossman J, Parmar MK, Spiegelhalter DJ, Freedman LS. A unified method for monitoring and analysing controlled trials. *Stat.Med.* 1994; 13:1815-1826.
177. Whitehead J. Sequential designs for equivalence studies. *Stat.Med.* 1996; 15:2703-2715.
178. Castle W, Fuller R, Hall J, Palmer J. Serevant nationwide surveillance study: comparison of salmeterol with salbutamol in asthmatic patientes who require regular bronchodilator treatment. *Br Med J* 1993; 306:1034-1037.
179. ISIS-2 Collaborative Group. Randomized trial of intavenous streptokinase, oral aspirin, both or neither among 17187 cases of suspected acute myocardial infarction: ISIS-2. *Lancet* 1988; 2:349-360.
180. Antiplatelet Trialist's Collaboration. Secondary prevention of vascular disease by prolonged antiplatelet treatment. *BMJ.* 1988; 296:320-331.

UAB

Biblioteca General  
Edifici A

Universitat de Barcelona (Barcelona) Espanya

## BIBLIOGRAFIA

---

181. Metcalf M, Reid J. FTN90 user's guide. Oxford: Oxford University Press, 1995.
182. Anonymous SAS/STAT user's guide, release 6.03 edition. Cary NC: SAS Institute Inc, 1988.
183. Anonymous SPSS-x user's guide. 3 ed. Chicago: SPSS Inc, 1988.