



GRASP

APPROFONDIMENTO

Analisi intermedie

Paolo Bruzzi, Genova

Analisi intermedie

(Sinonimo: analisi ad interim)

Analisi dei dati raccolti in uno studio condotte prima della sua conclusione programmata: possono essere classificate in base all'obiettivo (con grosse) implicazioni sul piano statistico



Analisi intermedie

OBIETTIVI POSSIBILI

- organizzativo (andamento del reclutamento, qualità dei dati, ecc.)
- aggiustamenti del piano statistico dello studio (caratteristiche dei pazienti, proporzione di successi/insuccessi, event rate, ecc.)
- analisi di safety
- analisi intermedie classiche, di superiorità
- analisi di futilità



Analisi intermedie

OBIETTIVI POSSIBILI

- ▶ organizzativo (andamento del reclutamento, qualità dei dati, ecc.)
- ▶ aggiustamenti del piano statistico dello studio (caratteristiche dei pazienti, proporzione di successi/insuccessi, event rate, ecc.)
 - *analisi di safety*
 - *analisi intermedie classiche, di superiorità*
 - *analisi di futilità*

Queste analisi dovrebbero essere condotte in cieco a gruppi di trattamento uniti, ma non pongono particolari problemi statistici o metodologici



Analisi intermedie

OBIETTIVI POSSIBILI

- *organizzativo (andamento del reclutamento, qualità dei dati, ecc.)*
- *aggiustamenti del piano statistico dello studio (caratteristiche dei pazienti, proporzione di successi/insuccessi, event rate, ecc.)*
- ▶ analisi di safety
- ▶ analisi intermedie classiche, di superiorità
- ▶ analisi di futilità

Queste analisi implicano un confronto tra i gruppi di trattamento, e richiedono **particolare attenzione sul piano statistico, ma anche metodologico**



Analisi intermedie "classiche"

Analisi comparative sull'**endpoint primario** per valutare se si osserva una differenza "**significativa**" tra i due gruppi, tale da richiedere



L'INTERRUZIONE DELLO STUDIO



Analisi intermedie "classiche"

▶ **ENDPOINT PRIMARIO**

(meno vincoli per endpoints secondari - tossicità)

▶ **ANALISI PRIMA DEL TERMINE PREVISTO**

- a. Arruolamento non completato
- b. Arruolamento completato - Rilevazione endpoint non completata/Non raggiunto numero di eventi previsto



Analisi intermedie "classiche"

SCOPI

a. Arruolamento non completato

(Non è più etico randomizzare pazienti nello studio)



- **Interrompere l'arruolamento**
- Passare i pazienti in trattamento alla terapia più efficace
- Pubblicazione
- Trattare i pazienti futuri con la terapia più efficace



Analisi intermedie "classiche"

SCOPI

b. Arruolamento completato

(Rilevazione endpoint non completata/Non raggiunto numero di eventi previsto)



TERAPIA IN CORSO

(Non è più etico trattare i pazienti con la terapia meno efficace)

-
- **Passare i pazienti in trattamento alla terapia più efficace**
 - Pubblicazione
 - Trattare i pazienti futuri con la terapia più efficace



Analisi intermedie "classiche"

SCOPI

b. Arruolamento completato

(Rilevazione endpoint non completata/Non raggiunto numero di eventi previsto)



TERAPIA COMPLETATA PER TUTTI I PAZIENTI IN STUDIO

- **Pubblicazione**
- **Trattare i pazienti futuri con la terapia più efficace**



Analisi intermedie "classiche"

FATTORI DA CONSIDERARE

- ▶ 1. Arruolamento lungo e/o terapia cronica
 - Problemi etici

- ▶ 2. Arruolamento rapido e terapia breve (es. chirurgica)
 - Pubblicazione
 - Registrazione del farmaco

- ▶ 3. Sempre: minore "sicurezza" del risultato
 - Stabilità statistica?
 - Rischi a lungo termine?



Analisi intermedie "classiche"

PROBLEMI

1. MOLTEPLICITÀ

2. STIME DELL'EFFICACIA DEL TRATTAMENTO



Analisi intermedie "classiche"

PROBLEMI

1. MOLTEPLICITÀ

- Se l'analisi intermedia è incoraggiata dall'osservazione dei risultati si introduce una distorsione

Soluzione: solo analisi pre-pianificate

- Se si analizzano ripetutamente i dati, è più facile osservare, per caso, una differenza significativa

Soluzione: correzione del livello di p per la significatività

2. STIME DELL'EFFICACIA DEL TRATTAMENTO



Analisi intermedie "classiche"

Metodologia corretta (per la molteplicità):

- ▶ si decidono in anticipo (nel protocollo o nel piano di analisi statistica) le occasioni in cui si faranno le analisi
- ▶ si stabilisce il livello di significatività richiesto per interrompere lo studio (<0.05)
- ▶ di solito livelli molto stringenti alle analisi intermedie per salvaguardare l'analisi finale



Analisi intermedie "classiche": metodi

Disegni a gruppi sequenziali

- ▶ Viene stabilito il numero di analisi e quando saranno effettuate, es. 3 analisi, dopo 100, 200 e 300 eventi (finale)
- ▶ Viene stabilito un insieme di regole per interrompere lo studio dopo ogni analisi
- ▶ Queste regole sono dei livelli di significatività (tutti $< \alpha$) calcolati in modo che l'errore α complessivo (p) sia quello desiderato (es. 5%)



In pratica, viene definito un criterio per "**spendere**" progressivamente l'errore α nelle varie analisi

Criterio di spesa "**costante**" (poco usato)

ES. 5 analisi (n)	1	2	3	4	5
Si interrompe lo studio se p					
	$p < 0.0158$	0.0158	0.0158	0.0158	0.0158

Criterio di spesa "**variabile**"

ES. 5 analisi (n)	1	2	3	4	5
Si interrompe lo studio se p					
	$p < 0.0051$	0.0061	0.0073	0.0089	0.0402
Oppure	$p < 0.0005$	0.0005	0.0005	0.0005	0.05



Implicazioni

- ▶ Le "spending functions" vengono calcolate in modo da preservare l'errore α complessivo
- ▶ Comportano una (di solito modesta) perdita di potenza (incremento numero di pazienti?)
- ▶ Con criteri di spesa variabili, lo studio viene interrotto precocemente solo in presenza di fenomeni molto marcati



Analisi intermedie "classiche"

PROBLEMI

1. MOLTEPLICITÀ

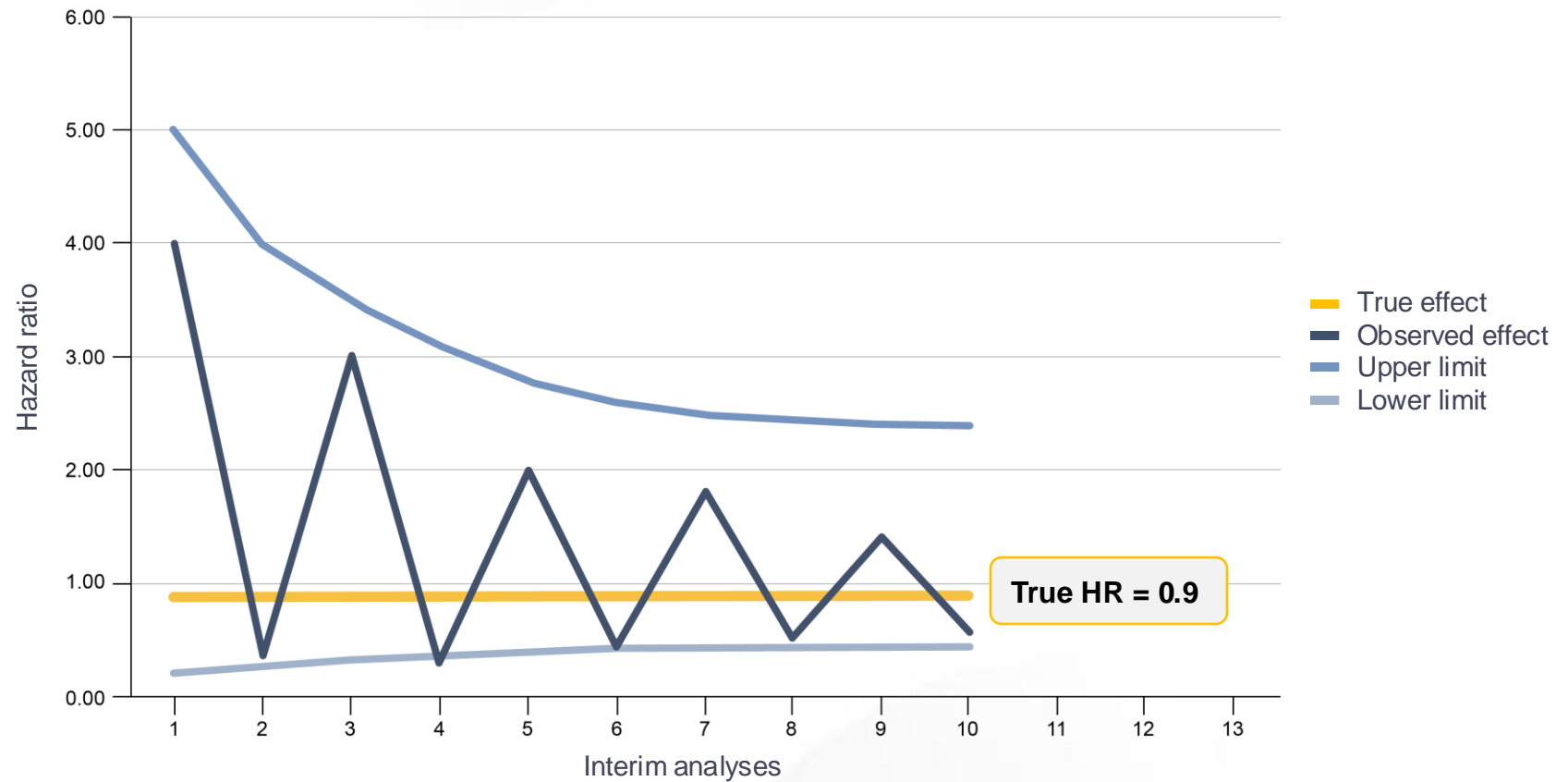
2. STIME DELL'EFFICACIA DEL TRATTAMENTO

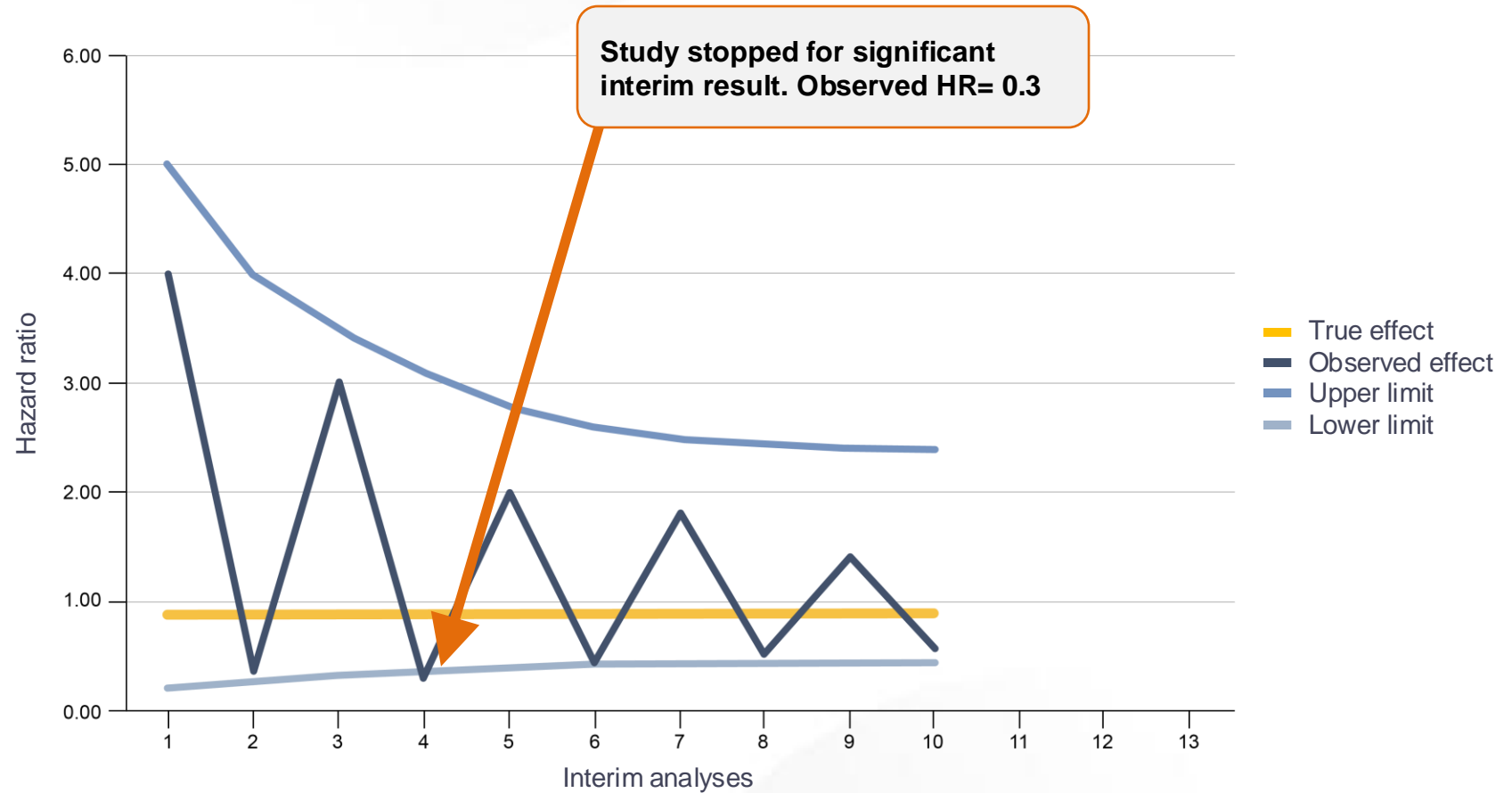


NOTA BENE (1)

- Le "spending functions" vengono calcolate in modo da **preservare l'errore α** complessivo
- Per **motivi statistici**, gli studi interrotti precocemente forniscono sistematicamente una **sovrastima** del beneficio del trattamento più efficace
- Non si dovrebbero quindi utilizzare queste stime per confronti e valutazioni cliniche





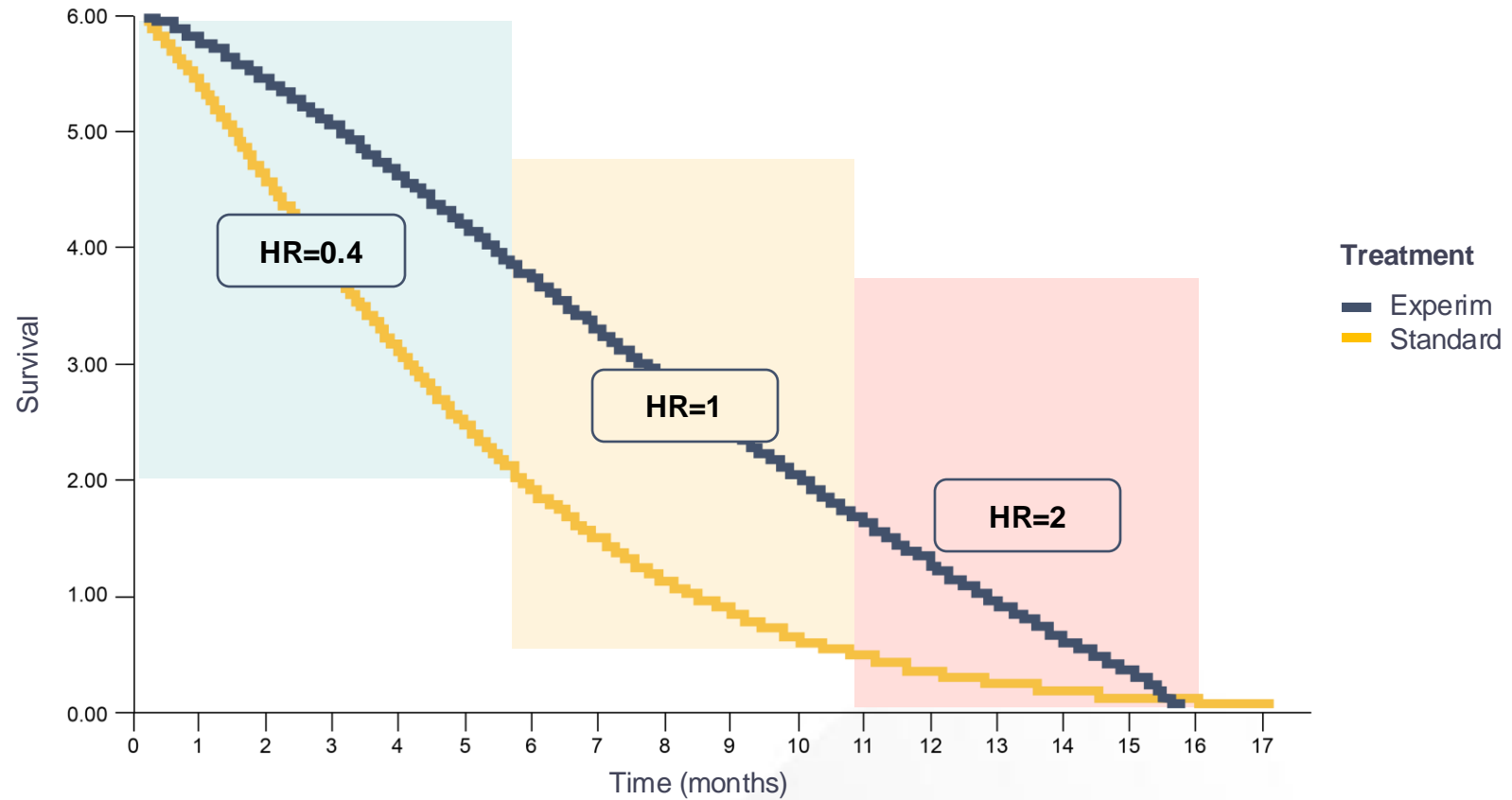


NOTA BENE (2)

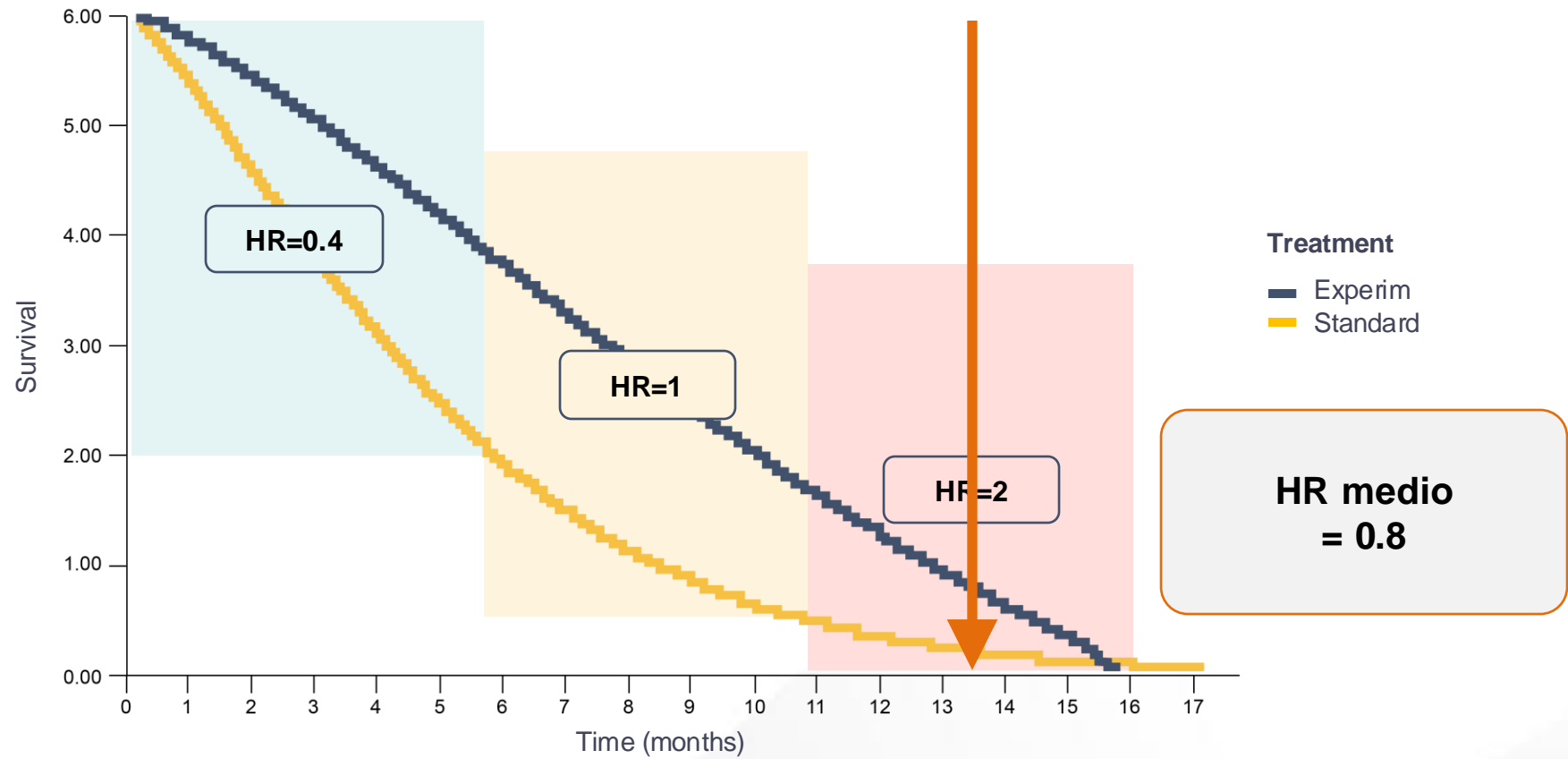
- Le "spending functions" vengono calcolate in modo **da preservare l'errore alfa** complessivo
- Gli studi che hanno come **endpoint primario un tempo a un evento**, se interrotti precocemente forniscono una **stima parziale e potenzialmente ingannevole** del beneficio del trattamento apparentemente più efficace
- Non si dovrebbero quindi utilizzare queste stime per confronti e valutazioni cliniche



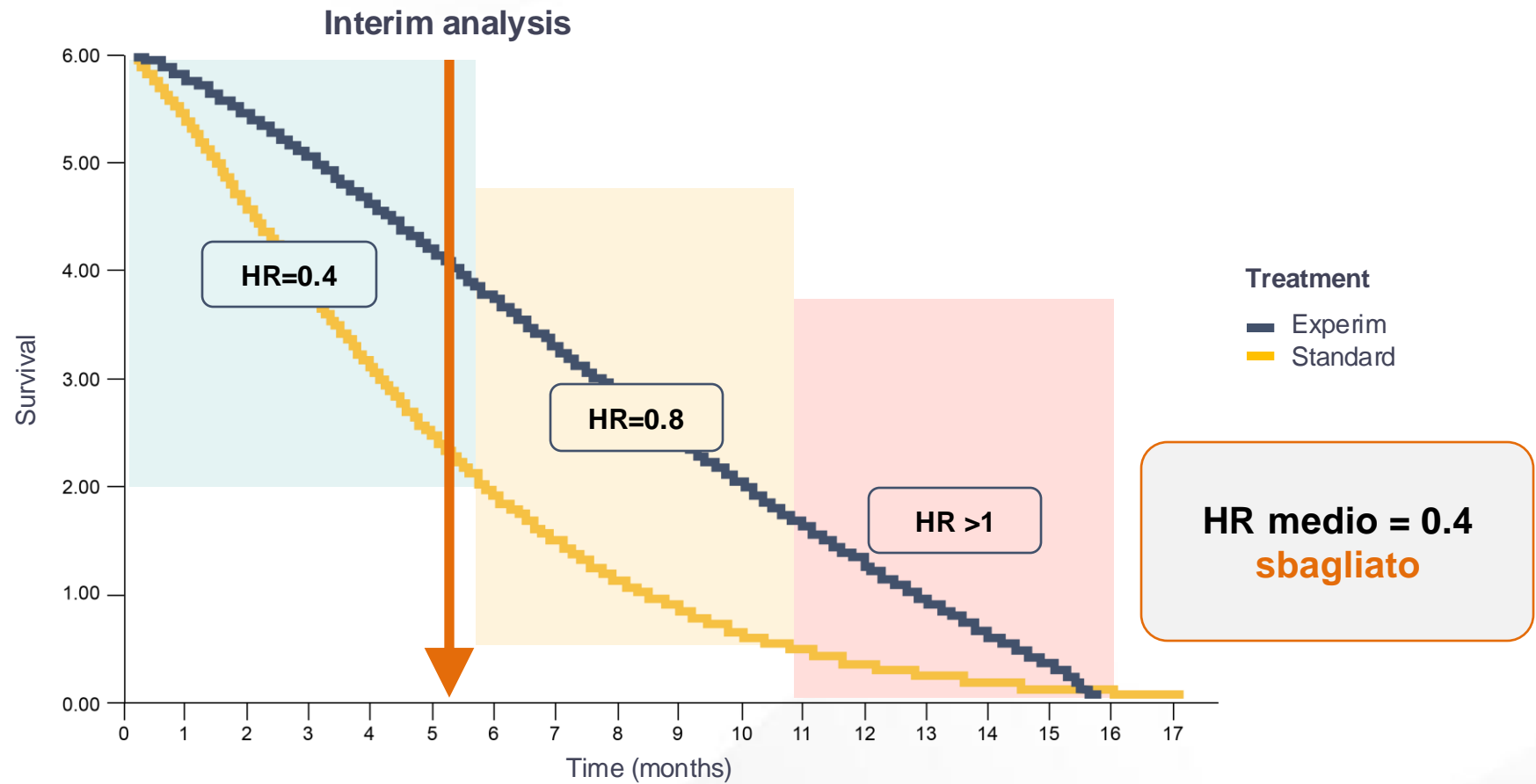
Ipotesi: HR not constant



HR not constant

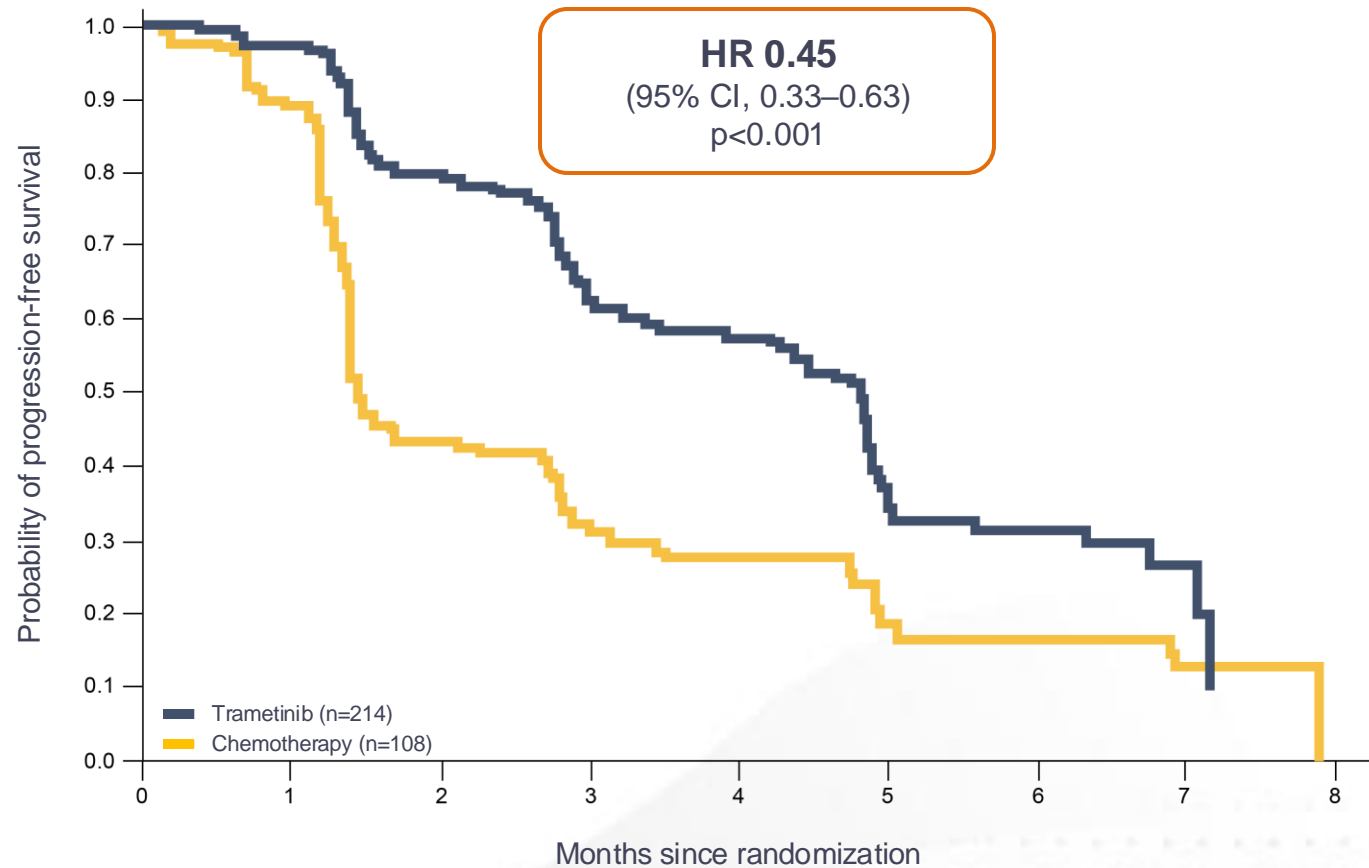


Se interim analysis, HR sovrastimato



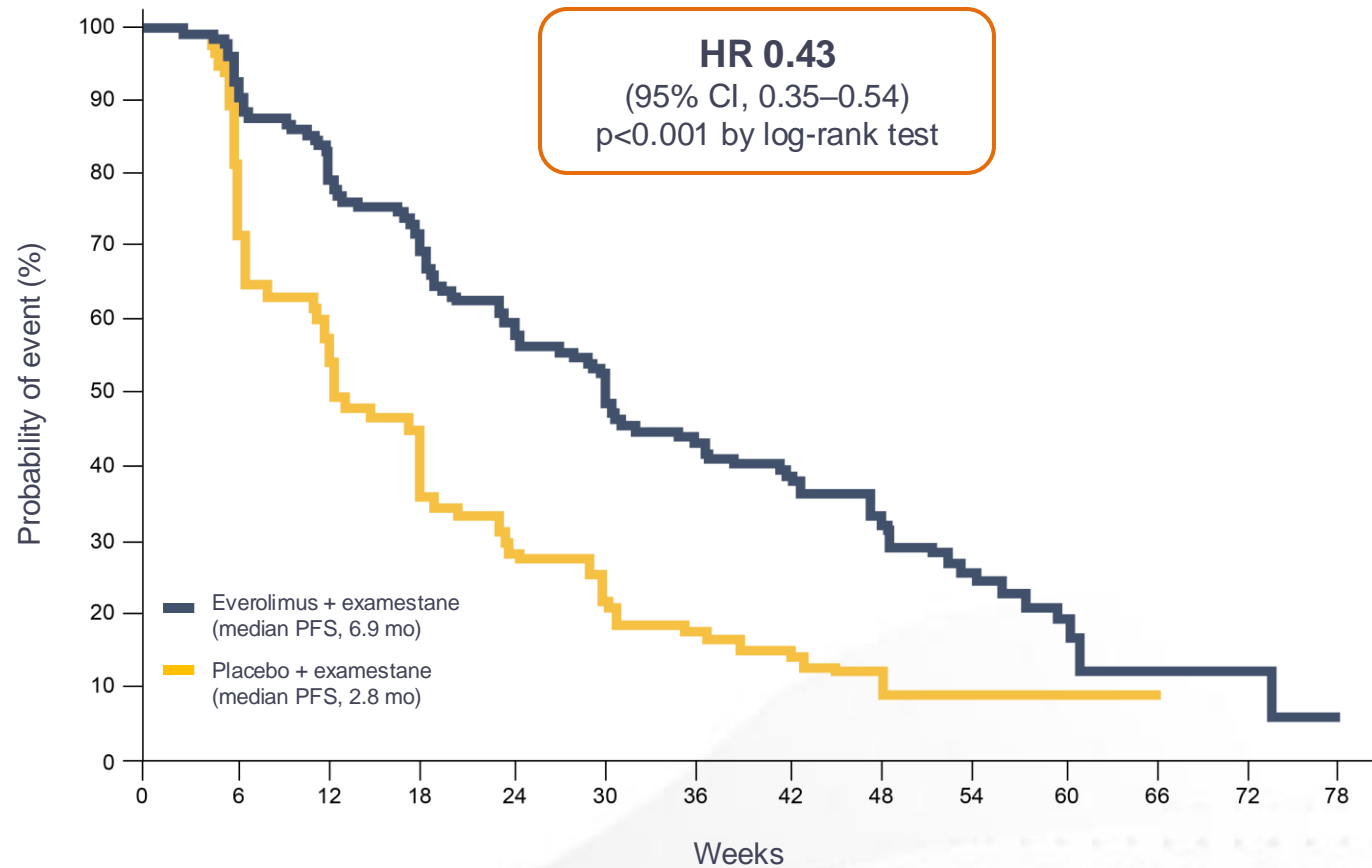
Advanced melanoma

Progression-free survival



Breast cancer

Local assessment



NOTA BENE (3)

Quando l'endpoint è fortemente collegato all'attività del trattamento...

- Es. in **Oncologia**: Risposta → Progression-free survival
- Es. in **Cardiologia**: Attività anti-ipertensiva → Ospedalizzazioni
- Es. in **Neurologia**: Riduzione lesioni MRI → Incidenza recidive cliniche

... i risultati delle analisi intermedie, se troppo precoci, possono essere fortemente fuorvianti



NOTA BENE (4)

Lo stesso problema può insorgere nel caso di studi **sovradimensionati**

Differenza statisticamente significativa a favore del trattamento sperimentale

- Statisticamente corretta: rifiuto dell'ipotesi nulla
- Clinicamente poco rilevante



Analisi intermedie per tossicità ed eventi avversi

- ▶ Necessarie per motivi etici (gli aspetti statistici sono meno rilevanti ma dovrebbero essere considerati)
- ▶ Affidate a un Data Monitoring Committee (DMC) indipendente che, a tempi prefissati, esamina i dati e stabilisce se motivi etici impediscono la prosecuzione dello studio o consiglia modificazioni del protocollo
- ▶ Lo stesso DMC esamina le analisi intermedie sull'endpoint primario e suggerisce provvedimenti



Analisi intermedie per "futility"

SCOPO

Terminare lo studio quando diviene probabile che il trattamento sperimentale non dimostrerà effetti benefici "significativi"

METODI

- a. Conditional Power
- b. Monitoraggio Bayesiano (più efficiente)



Analisi intermedie per "futility"

PERCHÈ?

- Terapia sperimentale tossica
- Terapia sperimentale verosimilmente inefficace (risultati altri studi)
- Terapia sperimentale non più interessante (nuove terapie in sperimentazione)
- L'arruolamento lungo
- Costi



Analisi intermedie per "futility"

METODOLOGIA

- **Conditional Power, Stochastic Curtailment**
Viene ricalcolata la potenza dello studio combinando i risultati già ottenuti con quelli prevedibili/possibili nella parte residua dello studio
- **Monitoraggio Bayesiano**
Viene calcolata la distribuzione di probabilità degli effetti del trattamento sulla base dei risultati raccolti (= si stima la probabilità che il trattamento sia efficace al livello desiderato - più efficiente)



Esempio: Sperimentale vs Controllo

- Differenza minima di interesse clinico: HR=0.75 (riduzione del 25% dell'incidenza)

Per $\alpha=0.05$, $\beta=0.2$ sono necessari 384 eventi

- Dopo 200 eventi di cui 103 nel gruppo sperimentale e 97 nel gruppo di controllo

HR = 1.13 (0.85-1.50)

Conditional Power: qual è la probabilità che lo studio, se proseguito fino alla fine, mi dia un risultato significativo a favore di Sperimentale?

3%!

Interrompo lo studio



Esempio: Sperimentale vs Controllo

- Differenza minima di interesse clinico: HR=0.75 (riduzione del 25% dell'incidenza)

Per $\alpha=0.05$, $\beta=0.2$ sono necessari 384 eventi

- Dopo 200 eventi di cui 103 nel gruppo sperimentale e 97 nel gruppo di controllo

HR = 1.13 (0.85-1.50)

Bayesian Monitoring: qual è la probabilità che l'HR "vero" sia almeno 0.75?

MOLTO BASSA!

Interrompo lo studio



Analisi intermedie per "futility"

NOTA BENE

L'errore α non è influenzato dal numero di analisi per futility

- Non c'è necessità di pianificarle né di correggere il livello di significatività
- Perdita di potenza



Analisi intermedie per "futility"

- ▶ Strumento di grande utilità che si sta diffondendo:
 - per monitorare studi che hanno difficoltà di reclutamento
 - per riconsiderare studi che valutano trattamenti non più interessanti
 - per interrompere studi dove i risultati sono in direzione opposta all'atteso
 - **come strumento per condurre studi a due fasi. Studi adattativi (seamless phase 2-3 trials)**



Conclusioni

- ▶ **1. Nominare DMC** (Data Monitoring Committee) che valuta i risultati delle analisi intermedie ed elabora proposte sulla gestione dello studio

- ▶ **2. Confronti su efficacia per interruzione studio:**
 - preparare un piano statistico adeguato per le analisi intermedie (numero, occasioni, stopping rules)
 - solo eccezionalmente avviare analisi intermedie non previste: compromettono la validità dello studio

- ▶ **3. Altre analisi intermedie (organizzative, safety, futility)**
 - massima libertà
 - riunioni DMC periodiche/su richiesta Steering Committee

