

**Per - corso di Statistica di Base
Modulo 6 – Elementi di epidemiologia**

Roma, 1 Giugno 2018

Analisi Statistica dei Trial Clinici

Marco Massari – Roberto Raschetti

Reparto di Farmacoepidemiologia
Centro Nazionale Ricerca e Valutazione Preclinica e Clinica dei Farmaci
Istituto Superiore di Sanità - Roma

Analisi Statistica dei Trial Clinici

✓ Protocollo, Statistical Analysis Plan (SAP) \Rightarrow Pre-Specificazione

✓ Metodi di Randomizzazione

Definizione del Set di Analisi \Rightarrow Intention To Treat, Per Protocol

Metodi Statistici (Endpoint: Categorici, Continui, Time To Event)

Sample Size Calculation \Rightarrow Effetto Clinicamente Rilevante

Analisi «Intention To Treat» e «Per Protocol»

Nel corso di uno studio può avvenire che i pazienti, per ragioni cliniche, possano essere trattati diversamente da quanto inizialmente previsto oppure possono, per diversi motivi, abbandonare la sperimentazione.

Come analizzare i dati?

- 1) per “**Intention to Treat**” (ITT), che significa analizzare i dati di tutti i pazienti sulla base della loro assegnazione iniziale ai gruppi, indipendentemente da come si sia poi evoluta la loro partecipazione alla sperimentazione.
- 2) per “**Protocollo**”, che significa analizzare i dati dei soli pazienti che hanno rispettato per tutto il tempo il protocollo originariamente previsto;

Stima dell'effetto del trattamento

Tipo di Endpoint	Misure di effetto - Analisi Statistica	
Qualitativo (categorico)	Rapporti di incidenza	Rischio Relativo (RR) Odds Ratio (OR)
Dicotomico Es. Responders	Differenze di incidenza	Riduzione assoluta di rischio (ARR) Numero di pazienti da trattare (NNT)
Quantitativo	Differenze tra medie	T-Test, ANOVA, ANCOVA
Time To Event	Confronto dei Tempi di Sopravvivenza	Analisi della sopravvivenza, Curve K-M, Hazard Ratio (HR), Modello di Cox

Misure di efficacia: Endpoint Qualitativo

Tipo di Endpoint	Misure di effetto - Analisi Statistica	
Qualitativo (categorico) Dicotomico / Responder	Differenze di incidenza	Riduzione assoluta di rischio (ARR)
		Numero di pazienti da trattare (NNT)

L'effetto assoluto è la differenza nelle incidenze tra una popolazione trattata e una non trattata

Fornisce una misura dell'impatto del trattamento

Riduzione assoluta di rischio (RAR)

Trattamento	Evento		Tot
	+	-	
A	a	b	a+b
B	c	d	c+d

$$I_A = \frac{a}{a+b} \quad I_B = \frac{c}{c+d}$$

$$RAR = I_B - I_A = \frac{c}{c+d} - \frac{a}{a+b}$$

Esempio RCT (II): Riduzione Assoluta di Rischio

PREVENTING MORTALITY FROM CORONARY HEART DISEASE WITH PRAVASTATIN

PREVENTION OF CARDIOVASCULAR EVENTS AND DEATH WITH PRAVASTATIN
IN PATIENTS WITH CORONARY HEART DISEASE AND A BROAD RANGE
OF INITIAL CHOLESTEROL LEVELS

THE LONG-TERM INTERVENTION WITH PRAVASTATIN IN ISCHAEMIC DISEASE (LIPID) STUDY GROUP*

NEJM, 1998;339:1349-1357

Trattamento	Numero di decessi cardiovascolari o di infarto del miocardio non fatale		Tot	
	+	-		
Pravastatina	557	3.955	4.512	$I_{Pr} = \frac{557}{4.512} \cdot 100 = 12,3\%$
Placebo	715	3.787	4.502	$I_{Pl} = \frac{715}{4.502} \cdot 100 = 15,9\%$

Riduzione Assoluta di Rischio (RAR)

$$I_{Pl} - I_{Pr} = 15,9\% - 12,3\% = 3,6\%$$

➔ Trattando 100 soggetti con Pravastatina si prevengono 4 casi di decessi per CHD o MI non fatali

Numero di pazienti da trattare (NNT)

Approccio per valutare i risultati di un trial: sintetizzare gli effetti del trattamento attraverso il **numero di pazienti** che un clinico deve trattare per prevenire un evento.

$$\text{NNT} = \frac{1}{\text{RAR}} = \frac{1}{I_B - I_A}$$

Esempio RCT (II): Numero di pazienti da trattare

PREVENTING MORTALITY FROM CORONARY HEART DISEASE WITH PRAVASTATIN

PREVENTION OF CARDIOVASCULAR EVENTS AND DEATH WITH PRAVASTATIN
IN PATIENTS WITH CORONARY HEART DISEASE AND A BROAD RANGE
OF INITIAL CHOLESTEROL LEVELS

THE LONG-TERM INTERVENTION WITH PRAVASTATIN IN ISCHAEMIC DISEASE (LIPID) STUDY GROUP*

NEJM, 1998;339:1349-1357

Trattamento	Numero di decessi cardiovascolari o di infarto del miocardio non fatale		Tot
	+	-	
Pravastatina	557	3.955	4.512
Placebo	715	3.787	4.502

$$I_{Pr} = \frac{557}{4.512} \cdot 100 = 12,3\%$$

$$I_{Pl} = \frac{715}{4.502} \cdot 100 = 15,9\%$$

Numero di pazienti da trattare (NNT)

$$NNT = \frac{1}{ARR} = \frac{1}{0,036} = 28$$

➔ Per prevenire un evento in studio (decessi per CHD o MI non fatali) bisogna trattare con pravastatina 28 soggetti

Misure di efficacia: Endpoint Continuo

Tipo di Endpoint	Misure di effetto - Analisi Statistica	
Quantitativo	Differenze tra medie	t-Test, ANOVA, ANCOVA (°)

(°) e corrispondenti test non parametrici

L'effetto del trattamento al tempo t si misura come delta (Δ) delle medie delle differenze rispetto al baseline (tempo 0)

«Mean Change From Baseline»

Unpaired t test (2 bracci)^(*), ANOVA (2+ bracci)^(), ANOVA per misure ripetute, ANCOVA**

(*) Corrispondente non parametrico Mann-Whitney U test

(**) Corrispondente non parametrico Kruskal Wallis test

Dal test t di Student ...

Trattamento	N.	Evento	
		Media	Ds
A	n_A	\bar{X}_A	ds_A
B	n_B	\bar{X}_B	ds_B

$$H_0 : \bar{X}_A = \bar{X}_B \quad H_1 : \bar{X}_A \neq \bar{X}_B$$

Test statistico: t di Student (t)

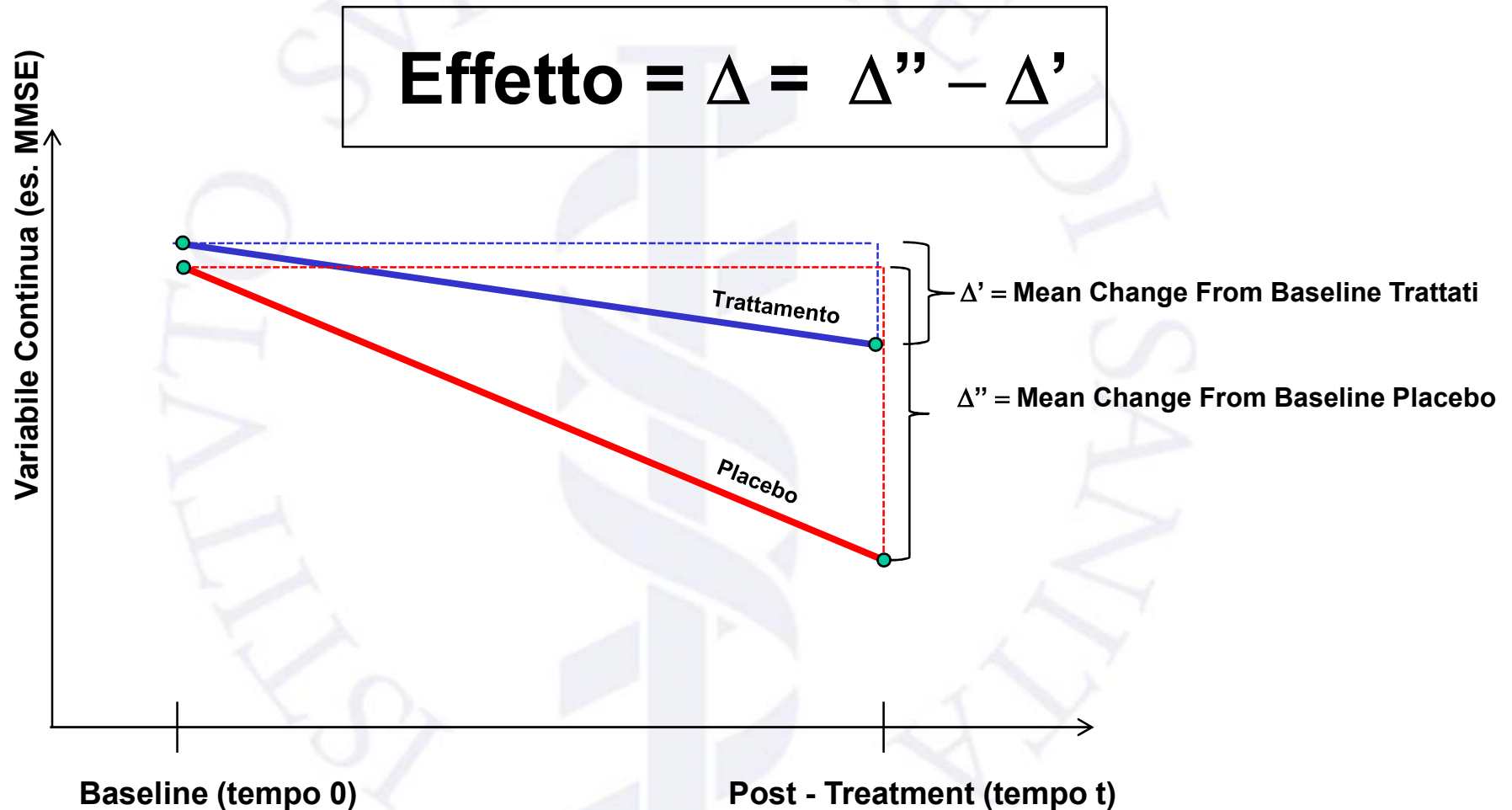
$$t = \frac{\bar{X}_B - \bar{X}_A}{\sqrt{\frac{ds_B^2}{n_B} + \frac{ds_A^2}{n_A}}}$$

Nota bene:

$$ES = \frac{ds}{\sqrt{n}}$$

$$ds = ES \times \sqrt{n}$$

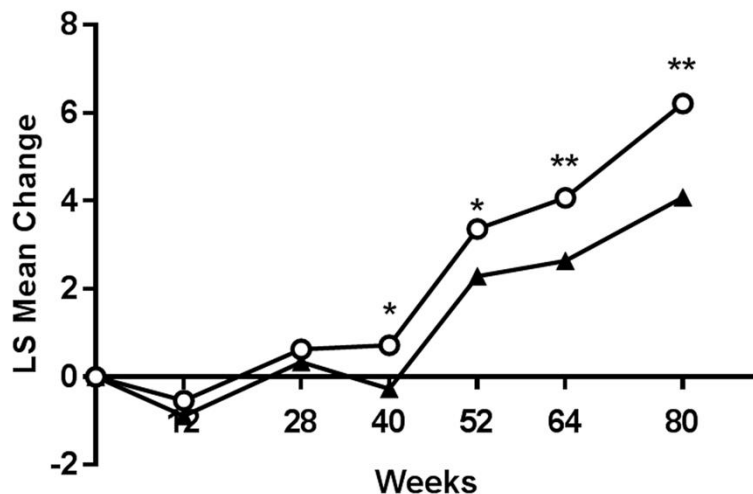
... al test t di Student (Δ Pre-Post) ...



Il Mini-Mental State Examination (MMSE) (Folstein et al., 1975), è un test neuropsicologico per la valutazione dei disturbi dell'efficienza intellettiva e della presenza di deterioramento cognitivo.

... all'ANOVA (misura ripetute) ...

ADAS-Cog₁₄

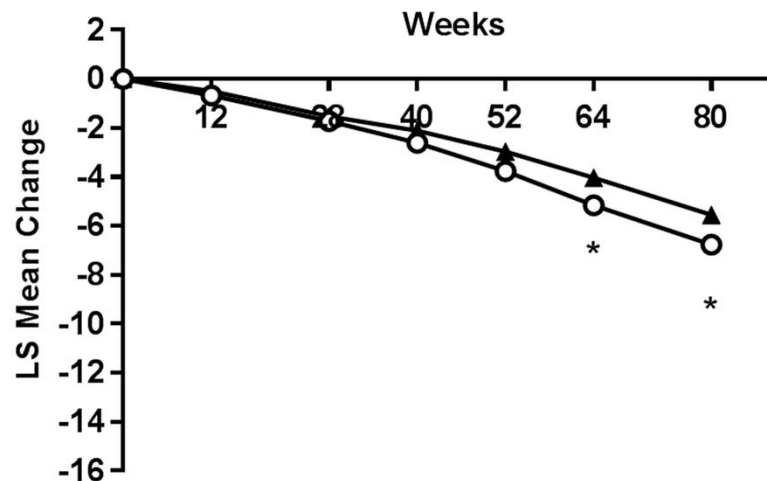


Placebo n=660
Solanezumab n=654

n=520
n=498

34% slowing in cognitive decline at 80 weeks

ADCS-iADL



Placebo n=660
Solanezumab n=652

n=521
n=500

18% slowing in functional decline at 80 weeks

○ Placebo
▲ Solanezumab

... all'Analisi della Covarianza (ANCOVA)

- **Estensione dell'ANOVA (combinazione tra ANOVA e Regressione)**
- **Metodo per confrontare medie tra 2 o più gruppi aggiustando per COvariate quantitative**

$$X_t = f(X_0, \text{Trattamento}, \text{Covariate})$$

$$X_t - X_0 = f(X_0, \text{Trattamento}, \text{Covariate})$$

$$X_t - X_0 = f(X_0, \text{Trattamento}, \text{Visita}, \text{Covariate})$$

Misure di efficacia: Endpoint Time to Event (Analisi della Sopravvivenza)

Tipo di Endpoint	Misure di effetto - Analisi Statistica	
Time To Event	Confronto dei Tempi di Sopravvivenza	Analisi della sopravvivenza, Curve K-M, Hazard Ratio (HR), Modello di Cox

Si usa quando l'Endpoint misura il tempo che va dalla randomizzazione fino a un qualche ben definito evento (**Time to Event**)



Per ogni paziente bisogna conoscere:

EVENTO (es. tempo al Decesso, alla Progressione della malattia es. Recidiva, Scomparsa del dolore, etc...)

CENSORING (tempo di uscita dallo studio per fine del follow-up o perché il paziente abbandona)

Tipologia di eventi

Gli “**Eventi**” possono avere una accezione Positiva:

es. Cessazione del dolore, Termine dell’ospedalizzazione, ...

o una accezione Negativa:

es. Progressione della malattia, Decesso, ...

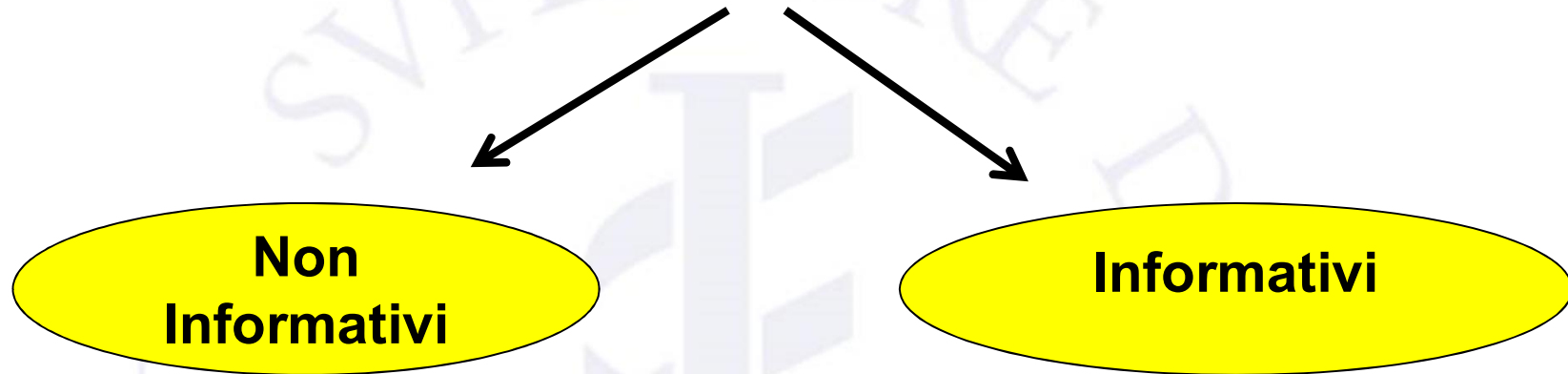
In tutti i casi si parla comunque in modo generale di:

Analisi della Sopravvivenza e Curve di Sopravvivenza

o

Survival Analysis

Tipologie di Censoring



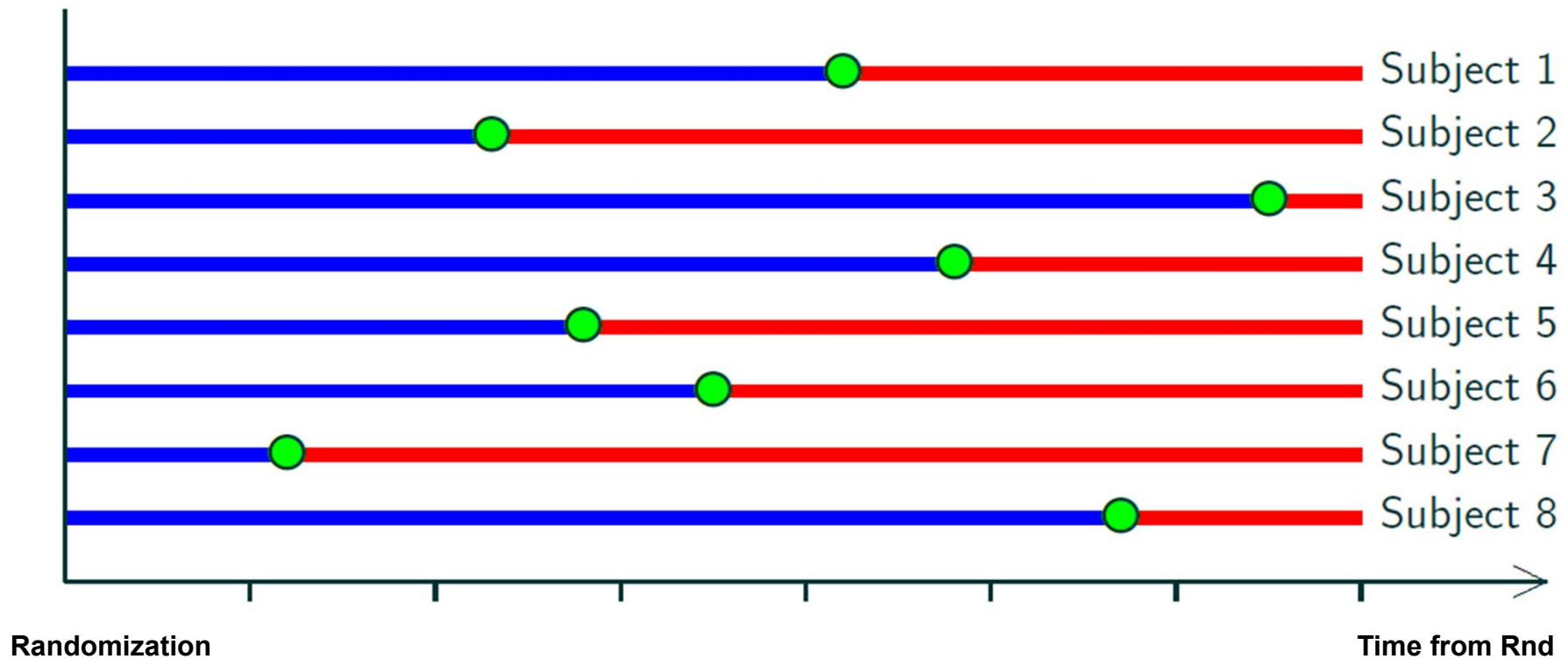
E' indipendente
(non è associato)
all'occorrenza
dell'evento

E' dipendente
(associato)
all'occorrenza
dell'evento

↓
Potenziale BIAS

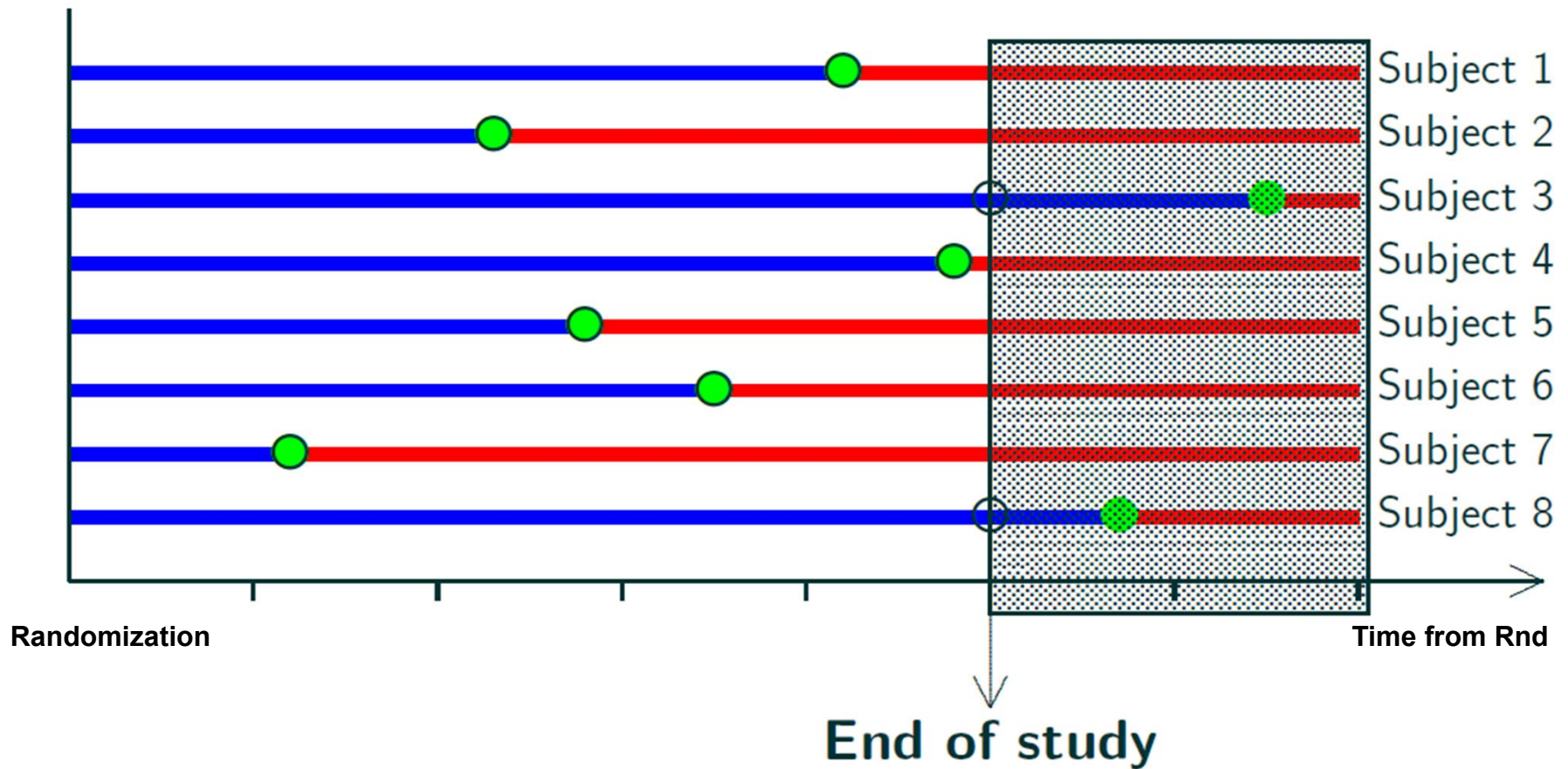
No Censoring

—: Before event —: After event ●: True event time ○: Observations



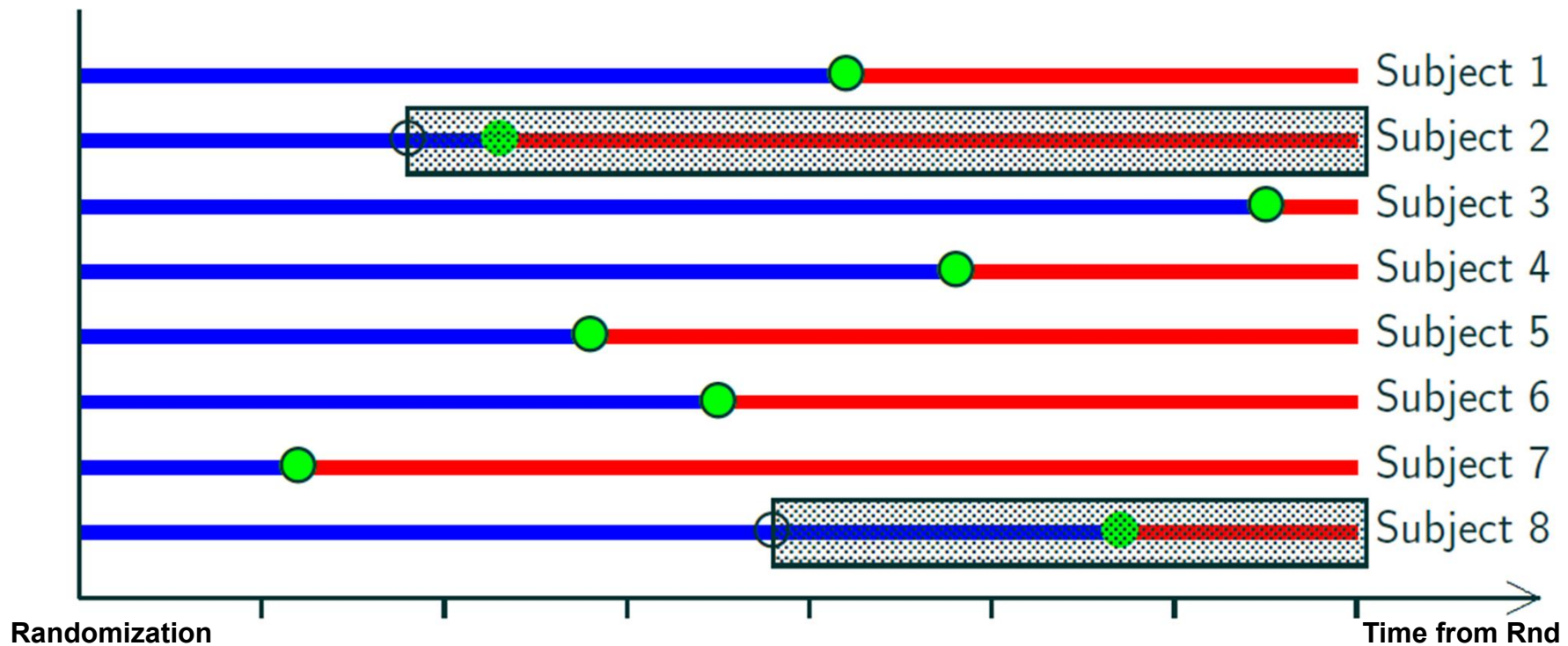
Right Censoring due to study end (non informative)

—: Before event —: After event ●: True event time ○: Observations



Right Censoring due to Withdrawal (informative)

—: Before event —: After event ●: True event time ○: Observations



Il problema dei casi Censored

Il fatto che ci siano dei tempi che sono Censored crea dei problemi ad analizzare i dati:

L'uso delle **medie** per confrontare i tempi di sopravvivenza di 2 gruppi non è appropriato (ancor più se decidiamo di ignorare i pazienti censored). In entrambi i casi si sottostima l'endpoint e introduciamo bias nel confronto

Potremmo invece usare le **mediane** che non risentono degli estremi (ma rimane comunque il problema che i censored sono delle sottostime del tempo ipotetico di sopravvivenza e quindi si collocano in posizione errata nella serie ordinata per il calcolo della mediana)

Metodo di Kaplan-Meier

- Definire un “**tempo iniziale**” (randomizzazione).
- Definire un “**tempo finale**” (evento/censoring).
 - i pazienti entrano in tempi diversi, vengono osservati per periodi diversi fino all’evento o al censoring.
- Calcolo dei “**tempi di sopravvivenza**” e degli “**event rate**”
- Stima, all’interno di ogni gruppo di trattamento, della **percentuale di pazienti sopravvissuti** per ogni periodo del follow-up.
- **Confronto delle curve di sopravvivenza**, costruite, separatamente per gruppi, in base ai vari tempi di sopravvivenza, per verificare se il trattamento abbia avuto o meno effetto nel diminuire l’incidenza dell’evento o nel prolungare la sopravvivenza.

Calcolo del Tempo di sopravvivenza



R: Pz Randomizzato (t=0)

E: Evento (es. Decesso)

C: Censored

Serie dei tempi di Sopravvivenza:
11 , 6 , 11 , 8+ , 12 , 9+ , 10 , 7

(+) sono tempi Censored

Ordiniamo e allineamo la serie:
6 , 7 , 8+ , 9+ , 10 , 11 , 11 , 12

(+) sono tempi Censored



Calcolo della Curva di Kaplan Meier

Serie dei tempi di Sopravvivenza 6 , 7 , 8+ , 9+ , 10 , 11 , 11 , 12 giorni

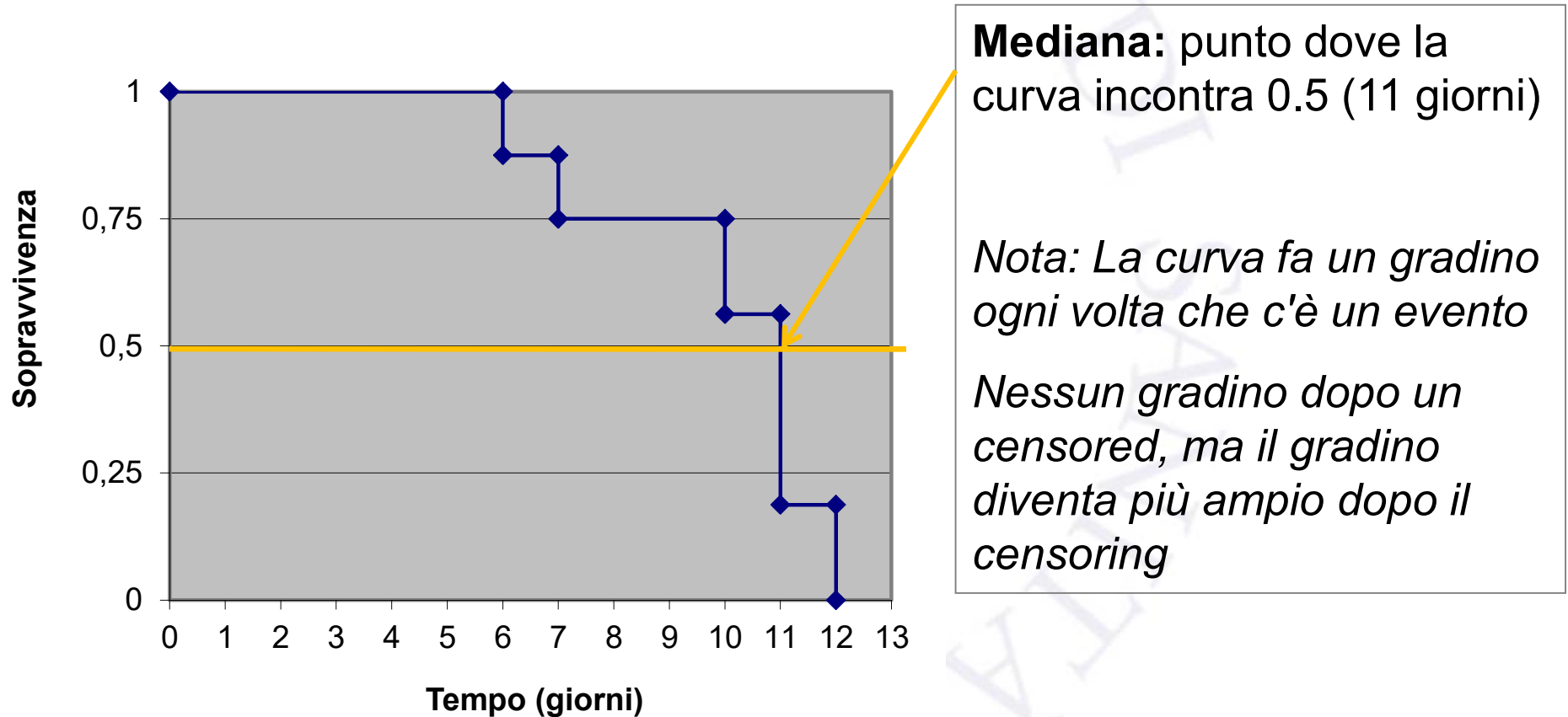
x (g)	n	e	p	Pc
0	8	0	1	1
6	8	1	$7/8=0,875$	$1 \times 0,875=0,875$
7	7	1	$6/7=0,857$	$0,875 \times 0,857=0,75$
10	4	1	$3/4=0,75$	$0,75 \times 0,75=0,5625$
11	3	2	$1/3=0,33$	$0,5625 \times 0,33=0,1875$
12	1	1	$0/1=0$	$0,1875 \times 0=0$

dove:

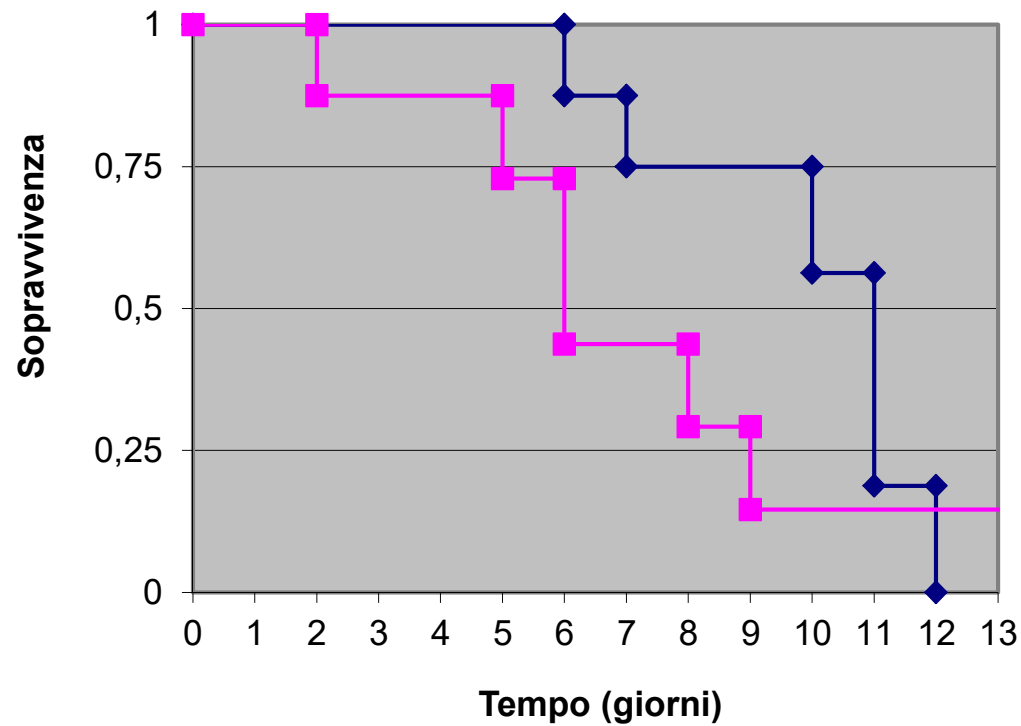
- x** tempo dall'ingresso nello studio
- n** numero di soggetti vivi e in osservazione all'inizio dello studio
- e** numero di soggetti con evento durante l'intervallo x
- p** Probabilità di sopravvivenza
- Pc** Probabilità cumulativa di sopravvivenza

Pc: probabilità di sopravvivere nell'intervallo da x a x+1, **supposto che il paziente sia sopravvissuto nei precedenti intervalli e sia vivo e sotto osservazione all'inizio dell'intervallo considerato.**

Grafico della Curva di sopravvivenza (Kaplan Meier)



Confronto tra 2 curve di sopravvivenza



Trattamento A

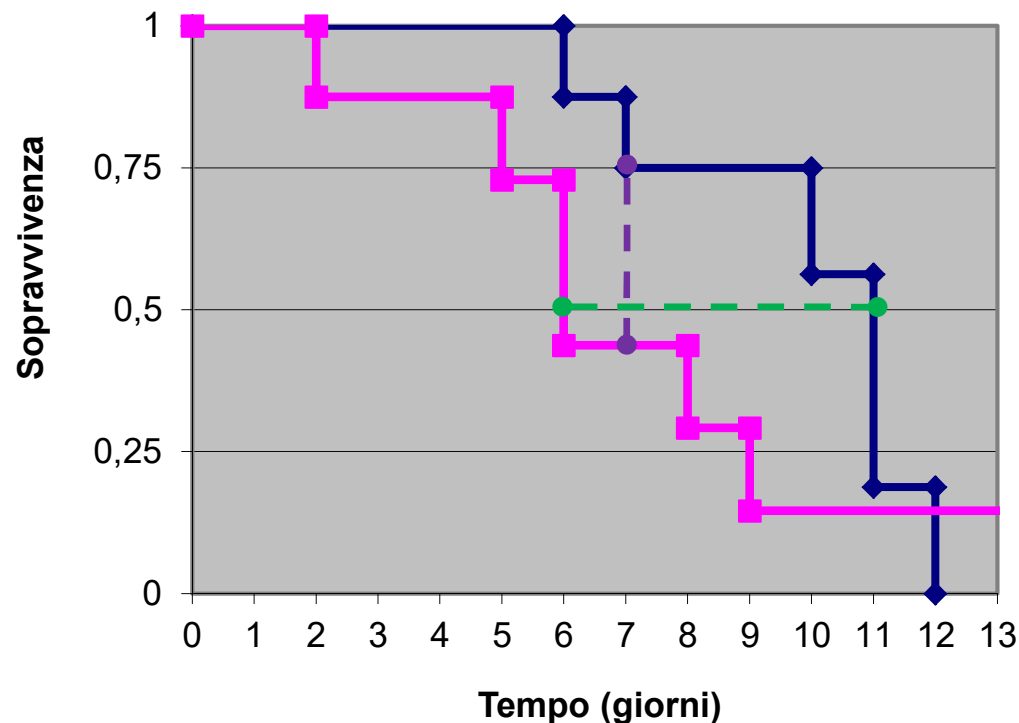
12, 11, 11, 10, 9*, 8*, 7, 6 gg

Trattamento B

10*, 9, 8, 6, 6, 5, 3*, 2 gg

Possiamo costruire e confrontare le 2 curve di sopravvivenza

Come misurare l'effetto del trattamento? Confronto attraverso mediane e/o proporzioni



Differenza in mediane

$$11 \text{ gg} - 6 \text{ gg} = 5 \text{ gg}$$

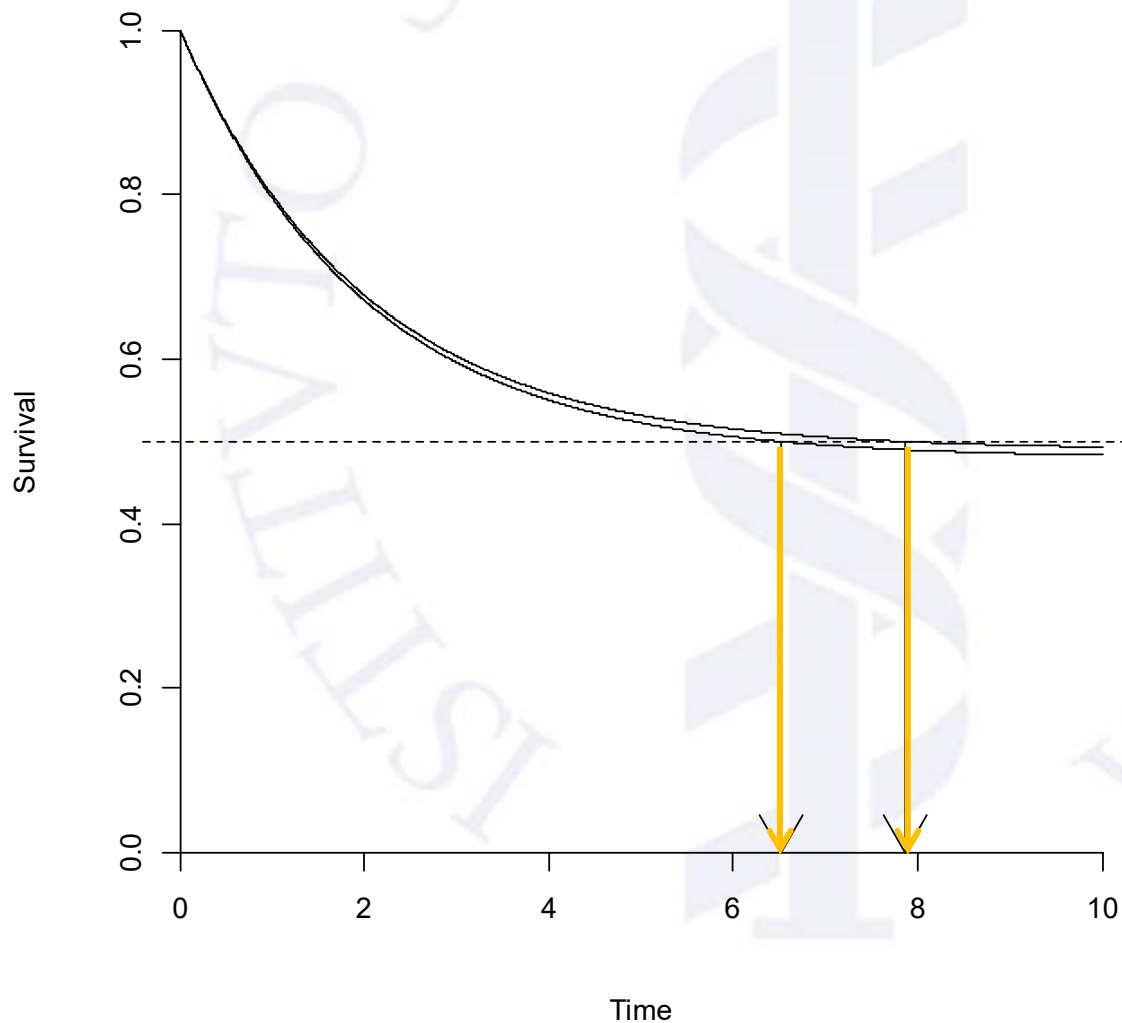
Differenza in % di "event free" a 1 settimana

$$0,75 - 0,44 = 0,31 \text{ (31\%)}$$



**Mediane o Proporzione
Quale è appropriata ??**

Situazione A

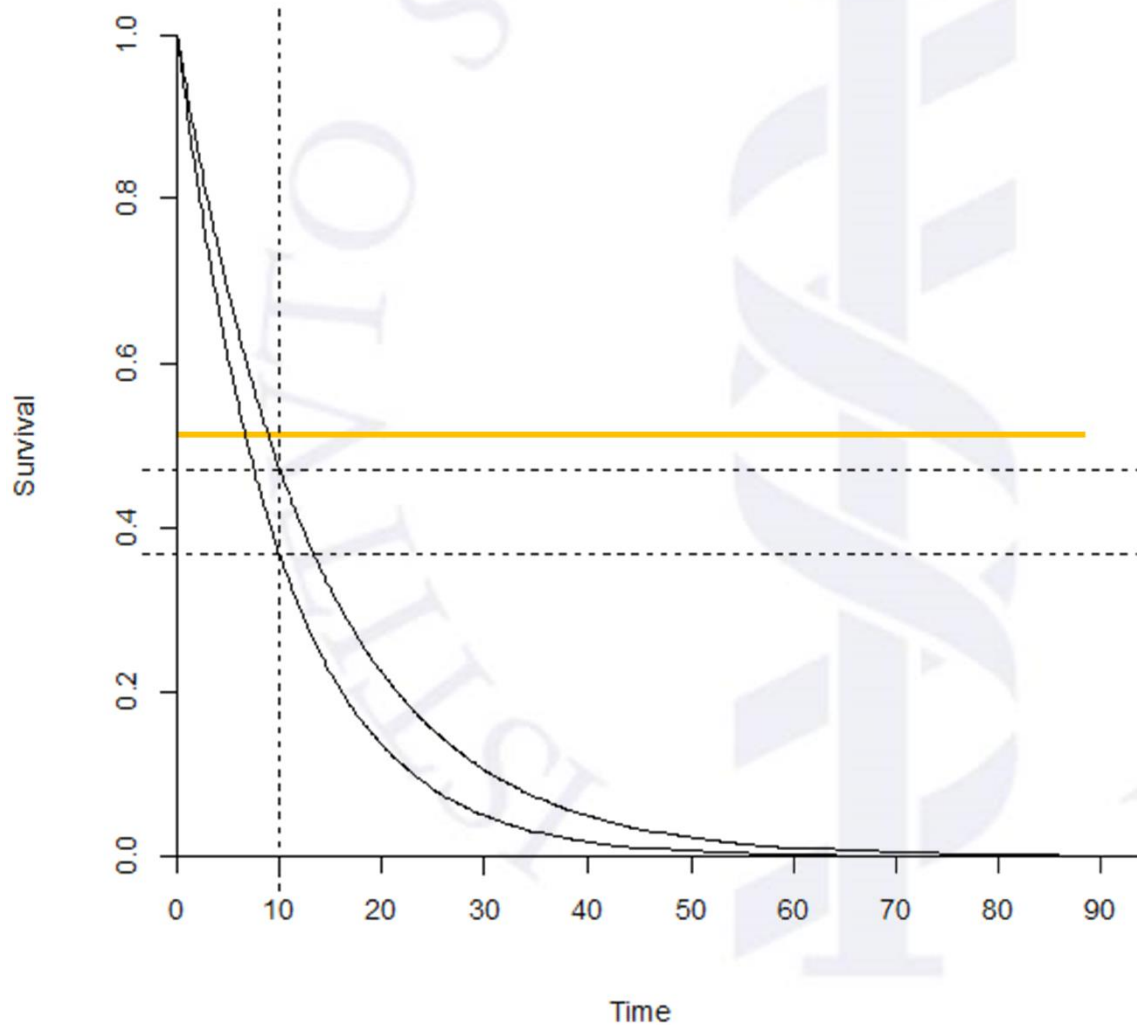


Slowing Declining Curves



Confronto attraverso
Proporzioni al tempo t
(event free rate)
è più realistico dell'uso
della mediana

Situazione B



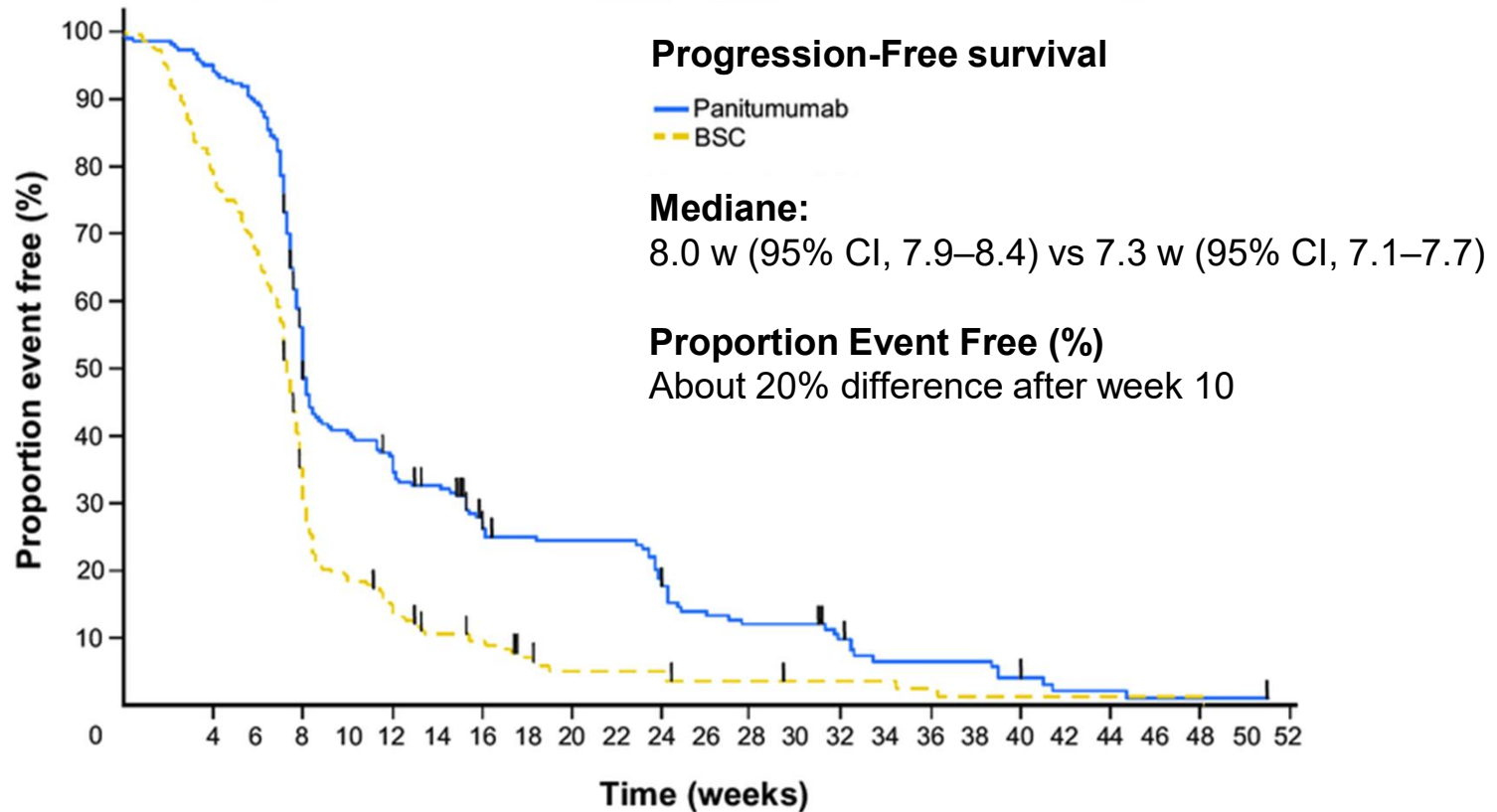
Quickly declining curves



Confronto attraverso
Mediane è più realistico
delle proporzioni al
tempo t

Nella realtà ...

Panitumumab + **Best Standard of Care** vs **BSC**
nel trattamento del tumore del Colon-retto Metastatico



Patients at risk

Panitumumab	231	209	118	76	49	40	31	19	13	8	5	2	1
BSC	232	175	75	31	17	7	7	4	3	2	1	1	1

Analisi di sopravvivenza: Log Rank test

H_0 : le due curve di sopravvivenza sono uguali

H_1 : le due curve di sopravvivenza sono differenti

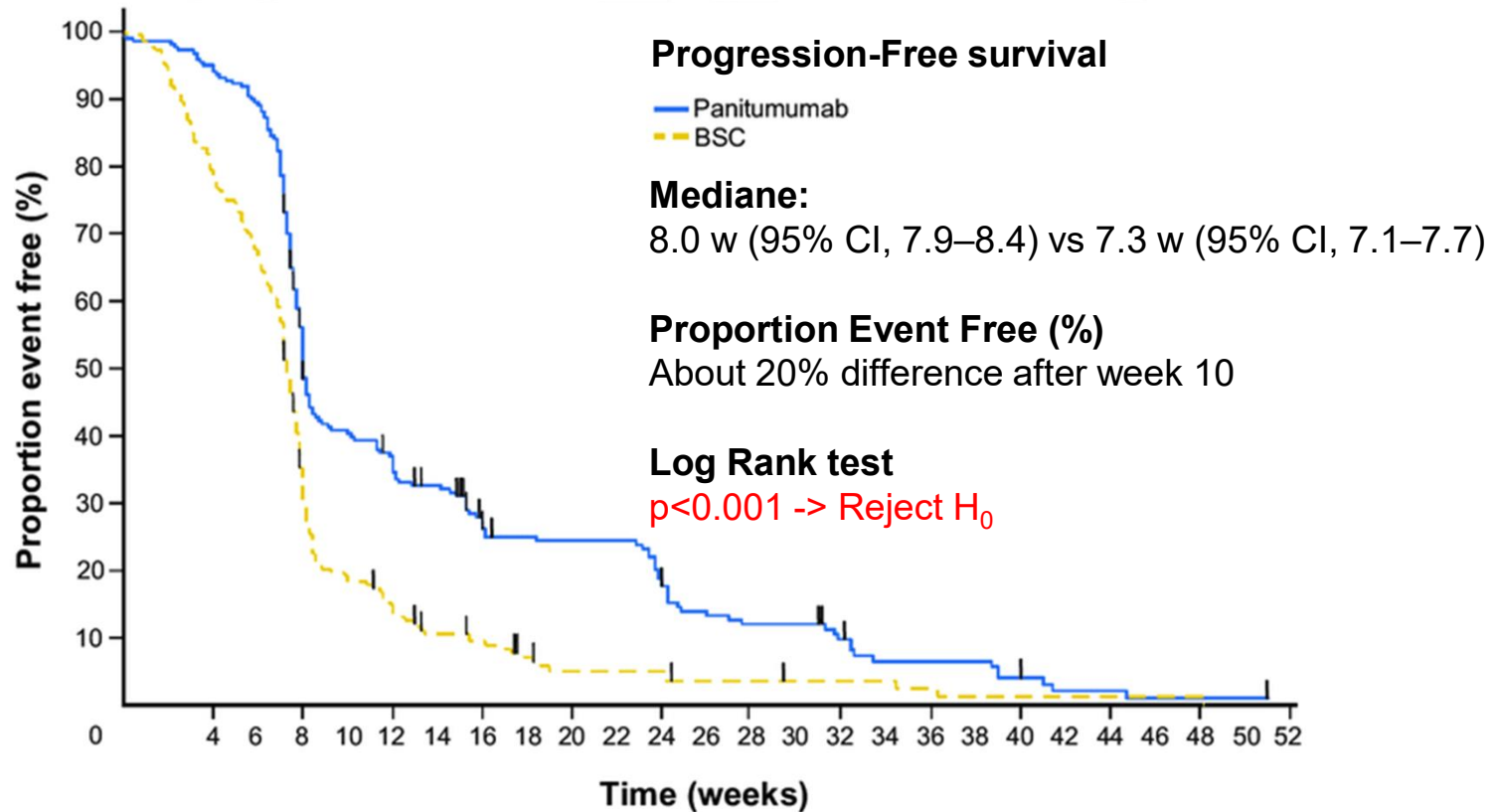
Uno dei test più usati per sottoporre a verifica questa ipotesi è il **LOG RANK TEST**, che permette di confrontare due curve di sopravvivenza.

due curve identiche → trattamenti ugualmente efficaci

**$p \leq 0,05$ → differenza statisticamente significativa tra le
2 curve di sopravvivenza**

... continua

Panitumumab + **Best Standard of Care** vs **BSC**
nel trattamento del tumore del Colon-retto Metastatico



Patients at risk

Panitumumab	231	209	118	76	49	40	31	19	13	8	5	2	1
BSC	232	175	75	31	17	7	7	4	3	2	1	1	1

Hazard (Rate) e Hazard Ratio

Hazard (Rate):

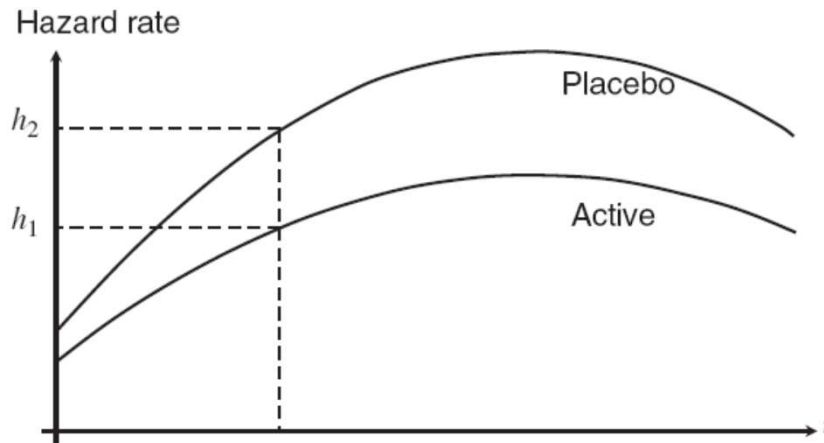
L'Hazard di un evento (es. il decesso, evento cardiovascolare, progressione malattia ...) rappresenta il tasso istantaneo dell'evento (misurato ad ogni istante) durante il periodo di follow up.

Hazard Ratio:

Anche detto Relative Hazard, è assimilabile al concetto di RR. Si ottiene dal rapporto tra gli Hazard di 2 gruppi di trattamento ed è usato per confrontare variabili Time to Event tra 2 gruppi

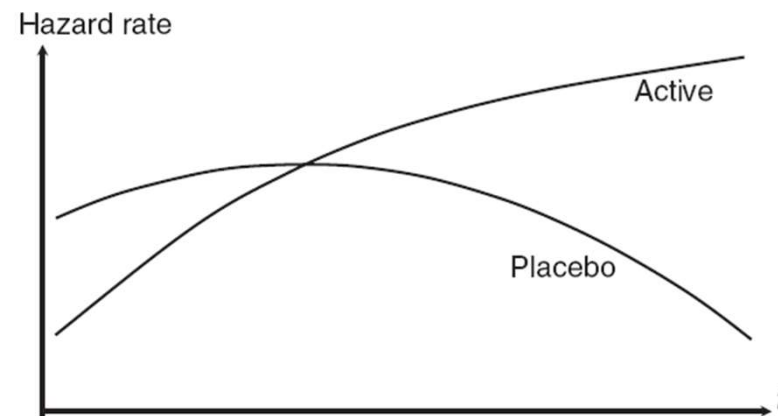
Modello di Cox

Se è soddisfatta l'ipotesi di proporzionalità dei rischi (Hazard Ratio Costante al variare di t) o almeno non vistosamente violata, si può applicare il modello di Cox e può essere stimato l'HR (Hazard Ratio) con il relativo intervallo di confidenza (95% CI)



Hazards variano nel tempo
Ma il loro rapporto è
sostanzialmente stabile.

Si può usare Cox Model



Hazards variano nel tempo
Ma la costanza del rapporto è
vistosamente violata.

Non si può usare Cox Model

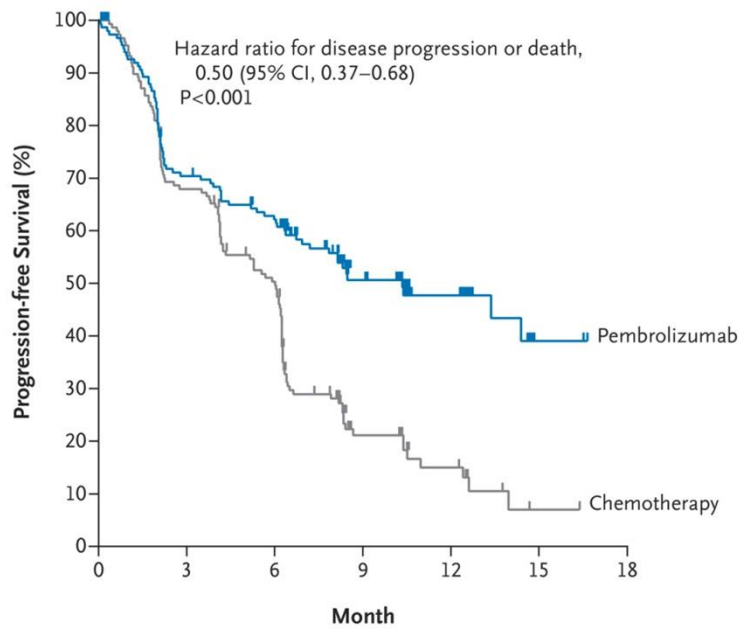
Interpretazione dell'Hazard Ratio

- HR=1 Stesso hazard di occorrenza di eventi tra 2 gruppi a confronto (indipendentemente dal tempo)
- HR<1 Minor hazard di occorrenza dell'evento nel gruppo trattato rispetto al comparator (indipendentemente dal tempo)
- HR>1 Maggior hazard di occorrenza di eventi nel gruppo trattato rispetto al comparator (indipendentemente dal tempo)

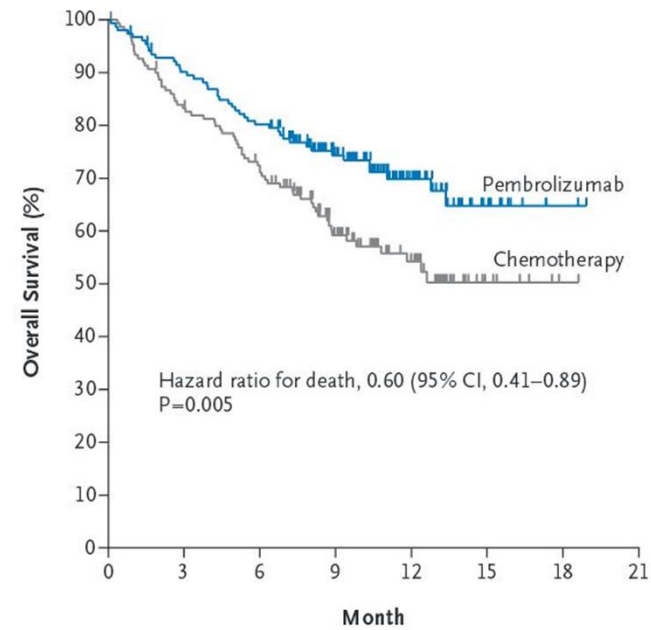
Nella realtà ...

Panitumumab + Best Standard of Care vs BSC nel trattamento del tumore del Colon-retto Metastatico

Progression-free Survival and Overall Survival in the Intention-to-Treat Population



No. at Risk	0	3	6	9	12	15	18
Pembrolizumab	154	104	89	44	22	3	1
Chemotherapy	151	99	70	18	9	1	0



No. at Risk	0	3	6	9	12	15	18	21
Pembrolizumab	154	136	121	82	39	11	2	0
Chemotherapy	151	123	106	64	34	7	1	0

Calcolo del Campione nel Trial Clinico

L'obiettivo principale è determinare il numero di pz. necessario a **catturare** un effetto del trattamento **cl clinicamente rilevante**

Se il numero di pz. è **troppo basso**, possiamo non essere in grado di evidenziare in modo statisticamente significativo l'esistenza di un effetto clinicamente importante \Rightarrow
(Under-Powered Trial)

Se il numero di pz. è **troppo alto** misure di effetto minime risulterebbero statisticamente significative (anche se non clinicamente rilevanti), oltre a richiedere tempo e risorse \Rightarrow
(Over-Powered)

Documentare il Sample Size in modo appropriato

Tutti gli elementi che concorrono al calcolo del Sample Size devono essere ben documentati nel protocollo, nel SAP, nella sottomissione di un progetto di ricerca, in una pubblicazione scientifica

CONSORT STATEMENT (<http://www.consort-statement.org/>)

α , β , Endpoint Primario o Co-primari, **Effetto Clinicamente Rilevante**,
Test Statistico Utilizzato (1 o 2 code), Parametri su cui si basa il calcolo del test con fonte, % di dropout, controllo della molteplicità.

Esempio

Sample Size

“To detect a reduction in PHS (postoperative hospital stay) of 3 days (SD 5 days), which is in agreement with the study of Lobo et al. 17 with a two-sided 5% significance level and a power of 80%, a sample size of 50 patients per group was necessary, given an anticipated dropout rate of 10%. To recruit this number of patients a 12-month inclusion period was anticipated.”(114)

Concludendo ...

- ✓ Protocollo, Statistical Analysis Plan (SAP) ⇨ Pre-Specificazione
 - ✓ Metodi di Randomizzazione
 - ✓ Definizione del Set di Analisi ⇨ Intention To Treat, Per Protocol
 - ✓ Metodi Statistici (Endpoint: Categorici, Continui, Time To Event)
 - ✓ Sample Size Calculation ⇨ Effetto Clinicamente Rilevante
 - Molteplicità ⇨ Type I Error Inflation
 - Gestione dei dati Mancanti ⇨ Metodi di Imputazione
 - Sensitivity Analyses
 - Analisi per Sottogruppi
 - Interim Analyses

Per Approfondire

- ✓ Linee Guida ICH (International Council for Harmonisation)
 - ICH E9 'Statistical Principles for Clinical Trials'
(www.ich.org/home.html)

- ✓ Linee Guida EMA e FDA
 - EMA Biostatistics
(www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000602.jsp&mid=WC0b01ac05807d91a4)
 - FDA Biostatistics
(www.fda.gov/RegulatoryInformation/Guidances/default.htm)

- ✓ Un buon libro:
 - A Concise Guide to Clinical Trials – Allan Hackshaw
(<https://onlinelibrary.wiley.com/doi/book/10.1002/9781444311723>)