

Con il Patrocinio di



STUDI CLINICI: METODOLOGIA

Coordinatore:

Dr.ssa Stefania Gori

Evento ECM MODULO 1
(formazione di base)

“A good foundation”



NEGRAR
22-23 Gennaio 2016

Centro Formazione
Ospedale Sacro Cuore
Don Calabria



STUDI CLINICI: METODOLOGIA

Coordinatore:
Dr.ssa Stefania Gori

*Evento ECM MODULO 1
(formazione di base)*

"A good foundation"



NEGRAR
22-23 Gennaio 2016

Centro Formazione
Ospedale Sacro Cuore
Don Calabria

...dove eravamo rimasti:

- Abbiamo un quesito clinico...
- Quale obiettivo?
 - ✓ *attività Vs efficacia*
- Quale disegno?
 - ✓ *sperimentale Vs osservazionale*
 - ✓ *randomizzato?*
- Quale endpoint?
 - ✓ *"vero" Vs surrogato*
 - ✓ *PRO?*
- Quale strategia di selezione?
 - ✓ *inclusiva Vs esclusiva*

**Ora si tratta di definire il piano
statistico dello studio...**

Con il Patrocinio di



STUDI CLINICI: METODOLOGIA

Coordinatore:

Dr.ssa Stefania Gori

*Evento ECM MODULO 1
(formazione di base)*

"A good foundation"



NEGRAR
22-23 Gennaio 2016

Centro Formazione
Ospedale Sacro Cuore
Don Calabria

Sabato 23 gennaio 2016

- Tipi di variabili statistiche
 - ✓ *indici di centralità e variabilità*
- Misure di effetto relative e assolute
 - ✓ *RR, OR, HR / ARR, NTT*
- Gli errori statistici
 - ✓ *alpha error* (tipo I)
 - ✓ *beta error* (tipo II)
- Il target di rilevanza clinica
 - ✓ *delta* (di superiorità)
 - ✓ *margin M* (di non inferiorità)
- Principi di dimensionamento campionario
 - ✓ per tipo di variabile di outcome
 - ✓ multiplicity e controllo di *alpha error*

Con il Patrocinio di



STUDI CLINICI: METODOLOGIA

Coordinatore:

Dr.ssa Stefania Gori

*Evento ECM MODULO 1
(formazione di base)*

"A good foundation"



NEGRAR
22-23 Gennaio 2016

Centro Formazione
Ospedale Sacro Cuore
Don Calabria

Sabato 23 gennaio 2016

- **Tipi di variabili statistiche**
 - ✓ ***indici di centralità e variabilità***
- Misure di effetto relative e assolute
 - ✓ *RR, OR, HR / ARR, NTT*
- Gli errori statistici
 - ✓ *alpha error* (tipo I)
 - ✓ *beta error* (tipo II)
- Il target di rilevanza clinica
 - ✓ *delta* (di superiorità)
 - ✓ *margin M* (di non inferiorità)
- Principi di dimensionamento campionario
 - ✓ per tipo di variabile di outcome
 - ✓ multiplicity e controllo di *alpha error*

VARIABILE DI RISPOSTA

- di tipo **quantitativo**
 - assume uno spettro continuo di valori e viene misurata in riferimento a una scala a intervalli costanti.
- di tipo qualitativo
 - esprime categorie di risposta del tipo successo / insuccesso (di un trattamento somministrato).
- del tipo “tempo a evento”
 - rappresenta il tempo trascorso fino al verificarsi (o meno) di un evento.

VARIABILE DI TIPO QUANTITATIVO

INDICI DI TENDENZA CENTRALE

	<i>definizione</i>	<i>vantaggi</i>	<i>svantaggi</i>
MEDIA	somma dei dati / numero dei dati	adatta a manipolazioni matematiche	influenzata dai valori estremi
MEDIANA	livello al di sotto del quale → metà dei dati	poco influenzata dai valori estremi	non adatta a manipolazioni matematiche
MODA	valore che ricorre con maggior frequenza	significato facilmente intuibile	possibilità di bi-modalità o tri-modalità



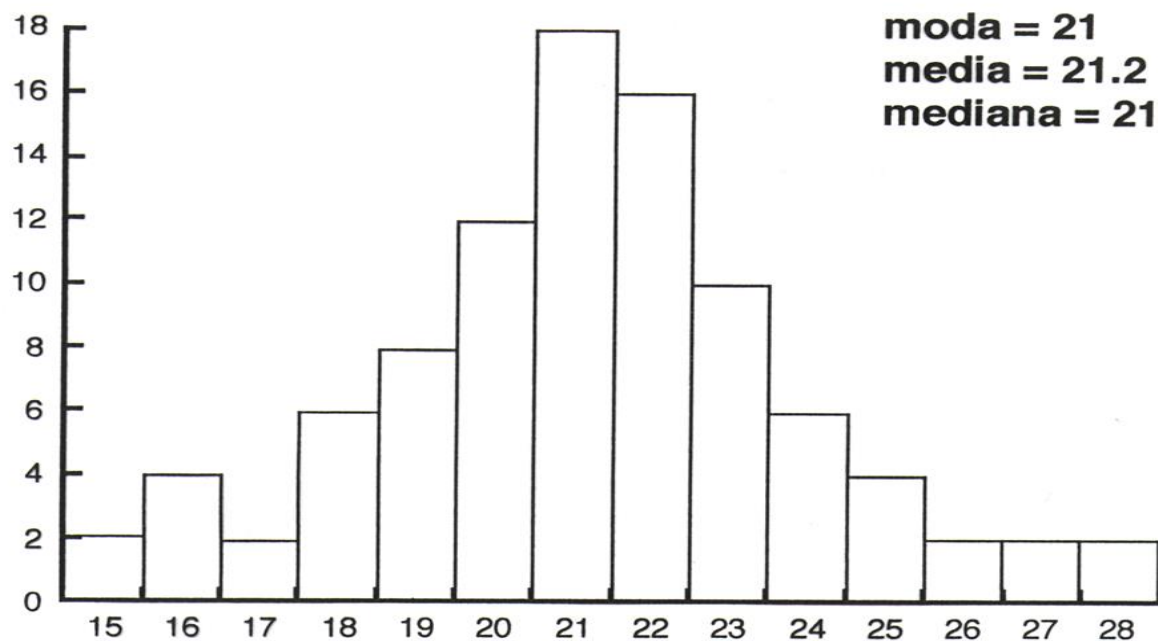
DISTRIBUZIONE DELLE FREQUENZE

<i>età</i>	<i>n° rilevazioni</i>
------------	-----------------------

15	2
16	4
17	2
18	6
19	8
20	12
21	18

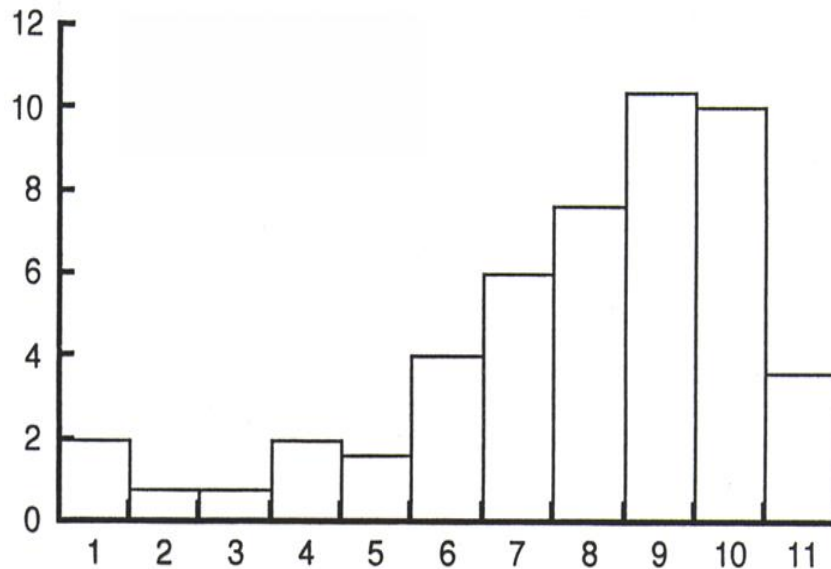
<i>età</i>	<i>n° rilevazioni</i>
------------	-----------------------

22	16
23	10
24	6
25	4
26	2
27	2
28	2

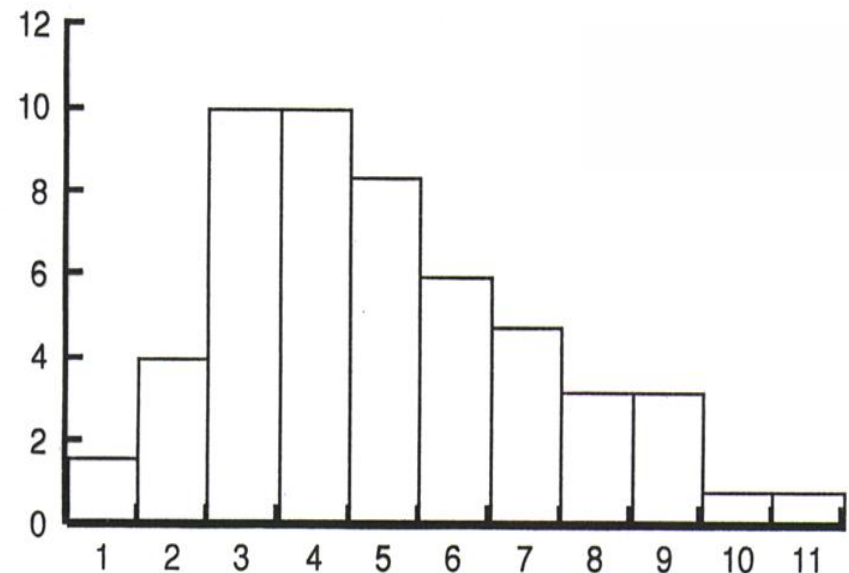


DISTRIBUZIONI ASIMMETRICHE

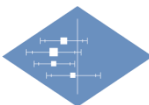
ASIMMETRIA NEGATIVA



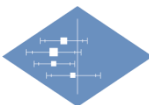
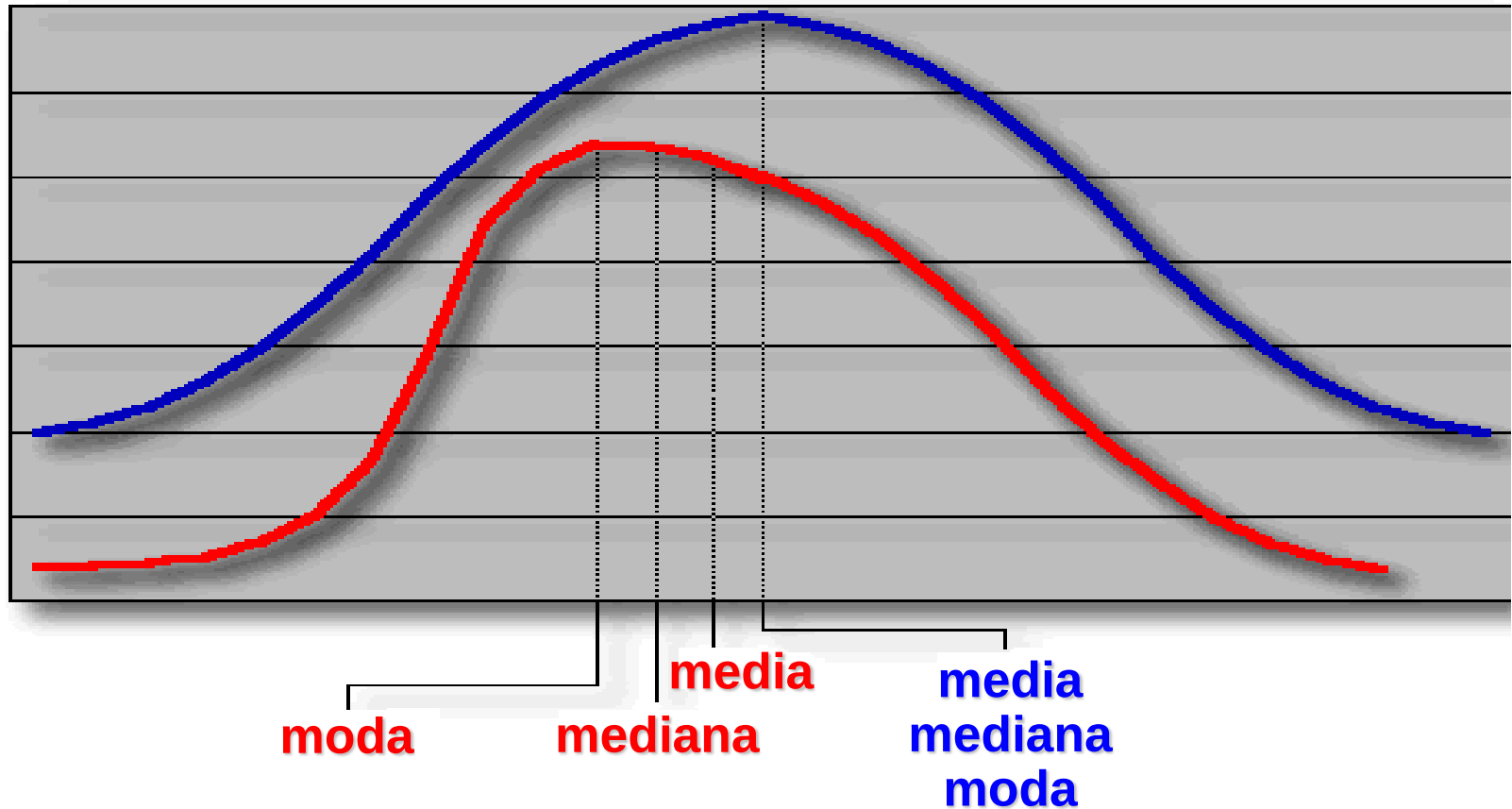
ASIMMETRIA POSITIVA



Indice di asimmetria (*skewness*) = misura della asimmetria della distribuzione rispetto alla media

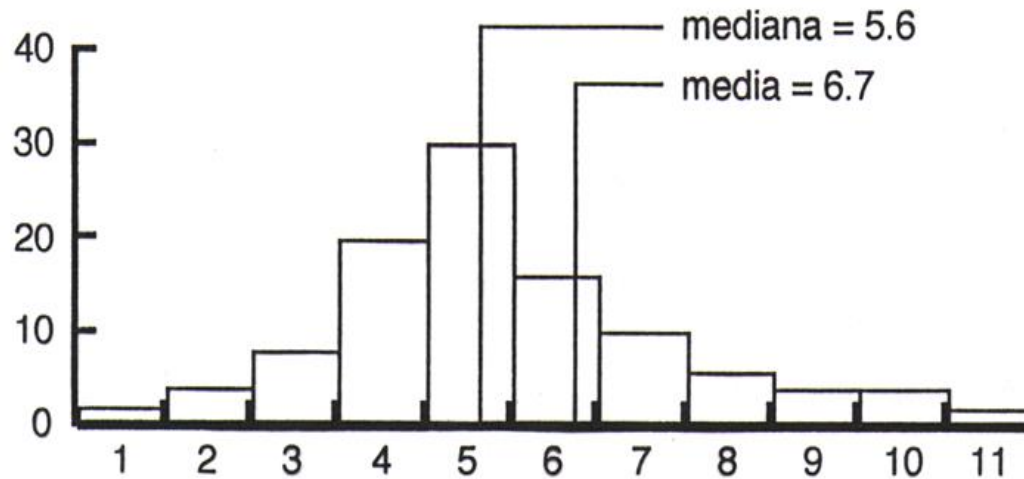


SIMMETRIA vs ASIMMETRIA

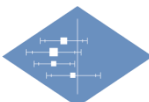


MEGLIO LA MEDIA O LA MEDIANA?

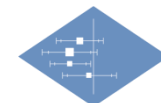
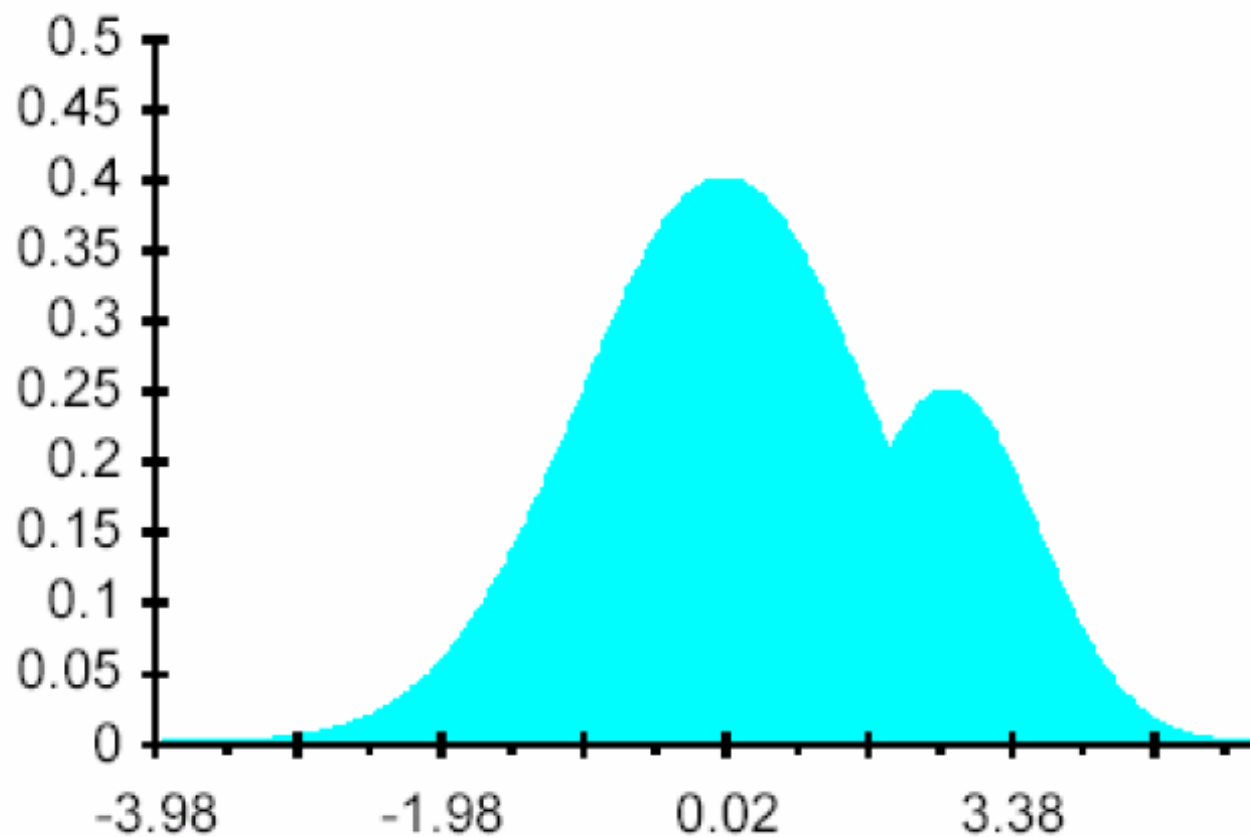
Nel caso specifico:



- la media è influenzata dai valori estremi (rispetto alla mediana è più distante dal valore tipico della distribuzione (**moda = 5**);
- la mediana fornisce una sintesi più realistica delle informazioni.



Esempio di distribuzione bimodale:



VARIABILE DI TIPO QUANTITATIVO

INDICI DI VARIAZIONE

definizione

vantaggi

svantaggi

**DEVIAZIONE
STANDARD**

**misura della
distanza media
dalla media**

**adatta a
manipolazioni
statistiche**

**non per
distribuzioni
asimmetriche**

RANGE

**valori più
basso e
più alto**

**comprende
tutti
i valori**

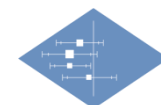
**influenzato
dai valori
stremi**

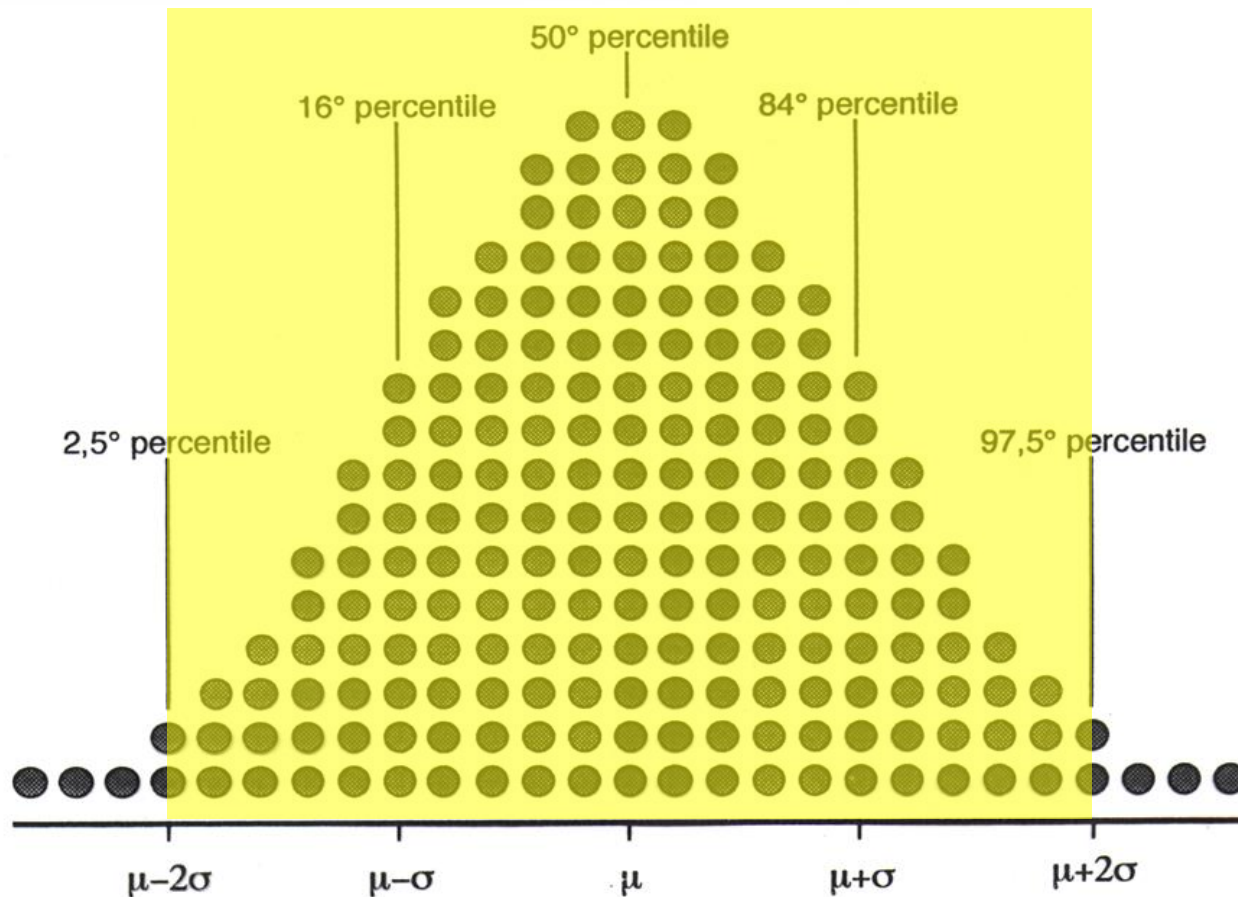
PERCENTILI

**% di osserva-
zioni → sopra
un certo valore**

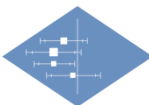
**OK per
distribuzioni
asimmetriche**

**non adatti a
manipolazioni
statistiche**





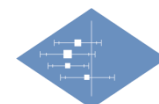
Quando le misurazioni appartengono a distribuzioni ragionevolmente simmetriche, il 95% di tali osservazioni cade entro 2 DS dalla media.



INDICI DI VARIAZIONE: DS vs. ESM

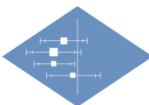
- La **deviazione standard** (DS) è la migliore misura di dispersione nel caso in cui (distribuzioni simmetriche) la media sia la migliore misura di centralità.
- La DS dell'insieme di tutte le possibili medie di campioni di una data numerosità n è detta **errore standard della media** (ESM):
 - ✓ esso è funzione sia della DS della popolazione sia della numerosità del campione.

$$\sigma_M = \frac{\sigma}{\sqrt{N}}$$



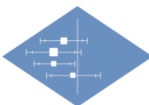
INDICI DI VARIAZIONE: DS vs. ESM

- La DS descrive la **variabilità di una misura.**
- L'ES descrive **l'incertezza nella stima della media.**
- *La maggior parte dei Ricercatori tende a riassumere i propri dati con l'ESM, poiché esso è sempre inferiore alla DS: ciò fa sì che i dati sembrino “migliori” di quanto lo siano nella realtà!*



INDICI DI VARIAZIONE E DISTRIBUZIONE ASIMMETRICA

- In caso di distribuzione asimmetrica è d'uso riportare:
 - ✓ il valore che separa il 25% inferiore della popolazione dal resto di essa (**25° percentile**);
 - ✓ il valore che separa il 25% superiore dal rimanente 75% (**75° percentile**);
 - ✓ l'intervallo di variazione dal valore più basso al valore più alto (**range**).



Punteggio	Frequenza	Freq. relativa %	Freq. cumulata	
0	1	4,2%	4,2%	
2	2	8,3%	12,5%	
4	1	4,2%	16,7%	
5	1	4,2%	20,8%	
7	1	4,2%	25,0%	= Q1
9	1	4,2%	29,2%	
10	1	4,2%	33,3%	
11	1	4,2%	37,5%	
12	2	8,3%	45,8%	
13	1	4,2%	50,0%	=Q2
14	1	4,2%	54,2%	
16	1	4,2%	58,3%	
19	1	4,2%	62,5%	
21	4	16,7%	79,2%	=Q3
22	1	4,2%	83,3%	
24	1	4,2%	87,5%	
25	1	4,2%	91,7%	
26	1	4,2%	95,8%	
29	1	4,2%	100,0%	=Q4

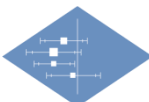
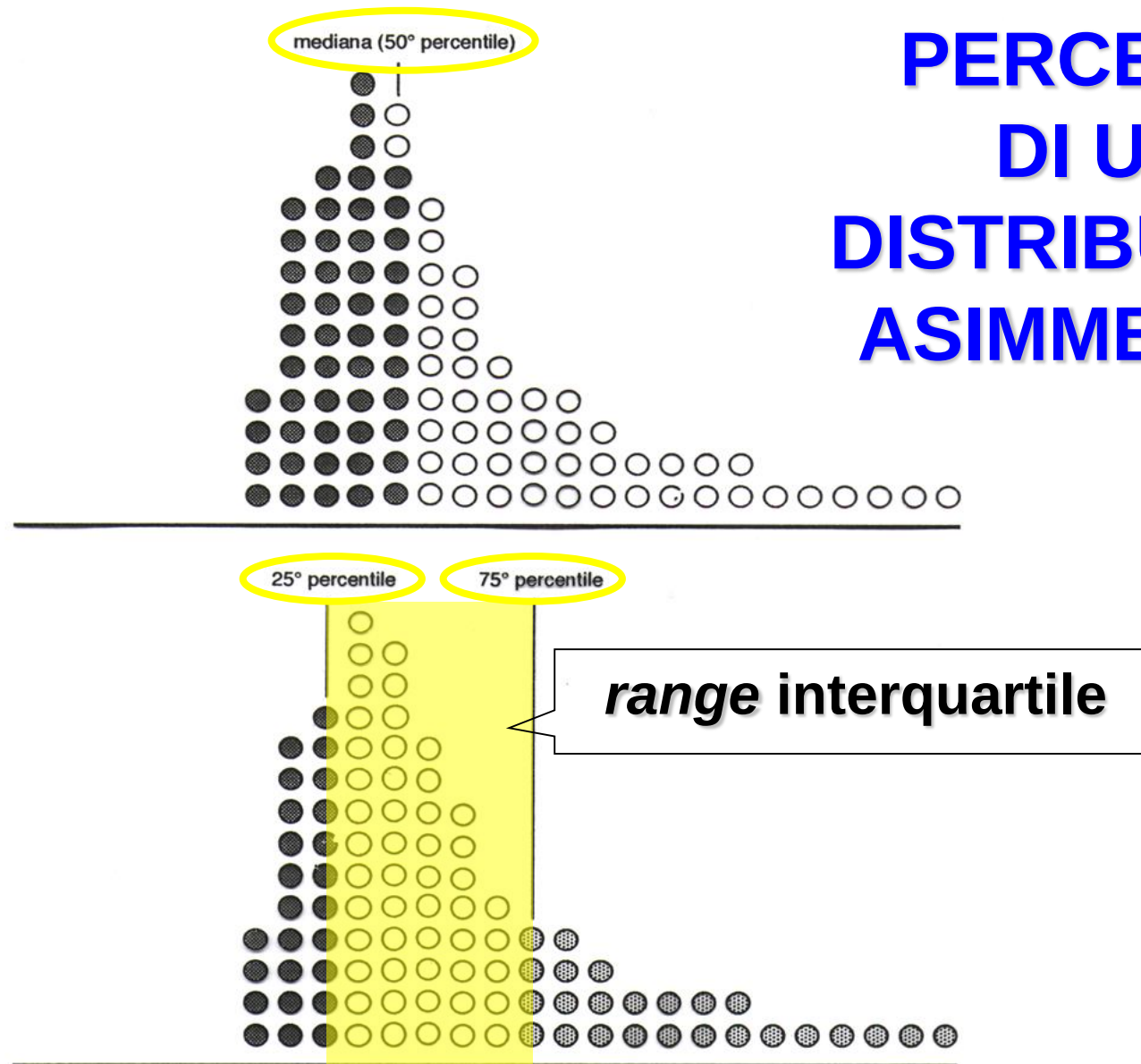
Primo quartile: si trova esattamente sul valore 7, dato che la percentuale cumulata corrispondente a tale punto è 25,0%

Secondo quartile: si trova esattamente sul valore 13, dato che la percentuale cumulata corrispondente a tale punto è 50,0%. Coincide sempre con la mediana

Terzo quartile: si trova all'incirca sul valore 21, dato che la percentuale cumulata corrispondente a tale punto è 79,2% (75,0%)

Quarto quartile: si trova sempre sull'ultimo valore, in questo caso è 29, dato che la percentuale cumulata corrispondente a tale punto è 100%

PERCENTILI DI UNA DISTRIBUZIONE ASIMMETRICA



VARIABILE DI RISPOSTA

- di tipo quantitativo
 - assume uno spettro continuo di valori e viene misurata in riferimento a una scala a intervalli costanti.
- di tipo **qualitativo**
 - esprime categorie di risposta del tipo successo / insuccesso (di un trattamento somministrato).
- del tipo “tempo a evento”
 - rappresenta il tempo trascorso fino al verificarsi (o meno) di un evento.

STUDIO DI UN FENOMENO QUALITATIVO

Per sintetizzare le informazioni raccolte è utile costruire una tabella che mostra i valori che possono essere presi da una variabile e il n° di soggetti (*frequenza assoluta*) aventi ciascuna modalità.

<i>tipo di neoplasia osservata</i>	<i>n° di casi*</i>
vescica	72
alte vie urinarie	16
prostata	86
rene	12
testicolo	14
<i>totale</i>	200

* frequenza assoluta

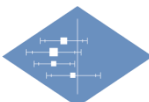


La **frequenza relativa** è il rapporto tra la frequenza assoluta di una modalità di una variabile ed il numero totale di osservazioni.

<i>tipo di neoplasia osservata</i>	<i>n° di casi*</i>	<i>frequenza relativa</i>
vescica	72	0.36
alte vie urinarie	16	0.08
prostata	86	0.43
rene	12	0.06
testicolo	14	0.07
<i>totale</i>	200	1.00

* frequenza assoluta

- Assume valori compresi tra 0 e 1.
- Spesso è espressa in percentuale.



ERRORE STANDARD DELLA STIMA

- La composizione del campione (in assenza di fonti di *bias*) è influenzata dal caso; per la misura del campione è influenzata dal *numero di campionamento*, di cui possiamo stimare i limiti.
- L'**errore standard** rappresenta un indice della variabilità (limiti probabili di oscillazione) della nostra misura.

$$\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

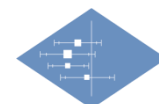
Stima della probabilità di risposta: $14/40 = 0.35$

Errore Standard (ES) della stima = 7.54



LIMITI DI CONFIDENZA

- Forniscono informazioni sulla precisione della stima con un livello di 'confidenza' prestabilito (tipicamente **95% ~ 2xES**).
- Sono interpretabili come la probabilità che l'intervallo comprenda l'effetto vero del trattamento.
- L'ampiezza dell'intervallo e quindi la precisione della stima, varia con la numerosità dello studio e con il livello di confidenza desiderato.

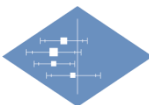


VARIABILE DI RISPOSTA

- di tipo quantitativo
 - assume uno spettro continuo di valori e viene misurata in riferimento a una scala a intervalli costanti.
- di tipo qualitativo
 - esprime categorie di risposta del tipo successo / insuccesso (di un trattamento somministrato).
- del tipo “**tempo a evento**”
 - rappresenta il tempo trascorso fino al verificarsi (o meno) di un evento.

STIMA DELLA FUNZIONE DI SOPRAVVIVENZA (Kaplan-Meier)

- **Stima della probabilità di sopravvivere in corrispondenza di ciascuno dei tempi in cui si verifica almeno un evento.**
- **Si assume che l'esperienza dei pazienti “usciti vivi” o “persi di vista” sia simile a quella dei pazienti rimasti sotto osservazione sino al momento in cui si è verificato l'evento.**
- **La sopravvivenza cumulativa si ottiene moltiplicando tra loro le probabilità così ottenute.**



**Analisi
della sopravvivenza
in sperimentazioni
cliniche controllate
e nelle osservazioni
pianificate**

E. Marubini - M.G. Valsecchi

A cura del «Centro Zambon»
dell'Università di Milano

DALL'ISTITUTO DI STATISTICA MEDICA E BIOMETRIA
DELLA FACOLTÀ DI MEDICINA E CHIRURGIA

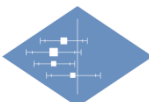
Tempi di risposta $t_{(j)}$	Tempi troncati* t^*	N° soggetti esposti a rischio n_j	N° eventi terminali d_j	Rischio istantaneo di "morte" $\hat{\lambda}(t_{(j)})$	Probabilità cumulativa di sopravvivere $t_{(j)}$ $\hat{P}_j$
9	177*	20	1	1/20 = .050	$(1 - 1/20) \times 1 = .9500$
13		19	1	1/19 = .053	$(1 - 1/19) \times .9500 = .8996$
20		18	1	1/18 = .055	$(1 - 1/18) \times .8996 = .8501$
26		17	1	1/17 = .059	$(1 - 1/17) \times .8501 = .7999$
27		16	1	1/16 = .062	$(1 - 1/16) \times .7999 = .7503$
28		15	1	1/15 = .067	$(1 - 1/15) \times .7503 = .7000$
30		14	1	1/14 = .071	$(1 - 1/14) \times .7000 = .6503$
32		13	2	2/13 = .154	$(1 - 2/13) \times .6503 = .5502$
75		11	1	1/11 = .091	$(1 - 1/11) \times .5502 = .5001$
79		10	1	1/10 = .100	$(1 - 1/10) \times .5001 = .4501$
91	1499*	9	1	1/9 = .111	$(1 - 1/9) \times .4501 = .4001$
		8	0	0/8 = .0	$(1 - 0/8) \times .4001 = .4001$
193		7	1	1/7 = .143	$(1 - 1/7) \times .4001 = .3429$
541		6	1	1/6 = .167	$(1 - 1/6) \times .3429 = .2856$
1129		5	1	1/5 = .200	$(1 - 1/5) \times .2856 = .2285$
	1499*	4	0	0/4 = .0	$(1 - 0/4) \times .2285 = .2285$
1585		3	1	1/3 = .333	$(1 - 1/3) \times .2285 = .1524$

TABELLA 10.

Calcolo secondo Kaplan e Meier della curva di sopravvivenza del rene trapiantato nei pazienti di tabella 6.



La precisione della stima per ogni periodo di tempo considerato dipende dal numero di soggetti che sopravvivono almeno quel tempo: se il numero di soggetti è esiguo, la variabilità della stima (sempre espressa in termini di ES) è elevata.



Con il Patrocinio di



STUDI CLINICI: METODOLOGIA

Coordinatore:

Dr.ssa Stefania Gori

*Evento ECM MODULO 1
(formazione di base)*

"A good foundation"



NEGRAR
22-23 Gennaio 2016

Centro Formazione
Ospedale Sacro Cuore
Don Calabria

Sabato 23 gennaio 2016

- Tipi di variabili statistiche
 - ✓ *indici di centralità e variabilità*
- **Misure di effetto relative e assolute**
 - ✓ ***RR, OR, HR / ARR, NNT***
- Gli errori statistici
 - ✓ *alpha error* (tipo I)
 - ✓ *beta error* (tipo II)
- Il target di rilevanza clinica
 - ✓ *delta* (di superiorità)
 - ✓ *margin M* (di non inferiorità)
- Principi di dimensionamento campionario
 - ✓ per tipo di variabile di outcome
 - ✓ multiplicity e controllo di *alpha error*

Quantification of Benefit

- Instruments & Tools -



- **Relative measures of risk reduction**
 - Odds Ratio (**OR**)
 - Risk Ratio (**RR**)
 - Hazard Ratio (**HR**)
- **Absolute measures of risk reduction**
 - **Absolute Differences** (%)
 - Number of patients needed to treat for 1 to benefit (**NNT**)

Odds Ratio (OR)

- Indice utilizzato per definire il **rapporto di causa-effetto tra due fattori**
 - Es. fattore di rischio – malattia
- **Confronto tra le frequenze di comparsa dell'evento** (ad esempio, malattia) rispettivamente nei soggetti esposti e in quelli non esposti al fattore di rischio in studio.
- Generalmente utilizzato negli **studi retrospettivi e caso-controllo**, dove non è necessaria la raccolta dei dati nel tempo,
 - non calcola un andamento
 - indipendente dal fattore durata.

$$OR = \frac{P(\text{malattia}|\text{esposti}) / (1 - P(\text{malattia}|\text{esposti}))}{P(\text{malattia}|\text{non esposti}) / (1 - P(\text{malattia}|\text{non esposti}))}$$

Rischio Relativo (RR)

- La probabilità che un soggetto, appartenente ad un gruppo esposto a determinati fattori, sviluppi la malattia, rispetto alla probabilità che un soggetto appartenente ad un gruppo non esposto sviluppi la stessa malattia.
- Questo indice è generalmente utilizzato negli studi prospettici e di coorte, dove l'esposizione è misurata nel tempo:

$$RR = I(esposti) / I(nonesposti)$$

- dove I = incidenza, che si definisce come:

$$I = n.nuoviammalati / (n.personetot - n.ammalati)$$

- *Se l'RR risulta uguale a 1 il fattore di rischio è influente sulla comparsa della malattia*
- *Se l'RR è maggiore di 1 il fattore di rischio è implicato nel manifestarsi della malattia*
- *Se l'RR è minore di 1 il fattore di rischio dipende dalla malattia (fattore di difesa).*

Hazard Ratios or Medians?

'Comments' upon Survival Curves

- **Hazard Ratio (HR)**

- Which model for survival analysis?
- HR: Principles, Assumptions and Limitations

- **Curves' Models**

- Median survivals and Late rates
- 'Adjusted' Data

- **'Visual' Maturity**

- **'Tricks' to enlarge HRs.....**

- **Correlation & Surrogates**

- **Implications for clinical trials' design**

- **The '*Two-Fingers*' Rule i.e. the '*quantity*' or**



Cancer Survival Analysis

- Survival tools to determine overall group survival, taking into account both ***censored*** and ***uncensored*** observations:

- **Life-table Method:**

- Actuarial method.
- Calculates the proportion surviving at fixed intervals.

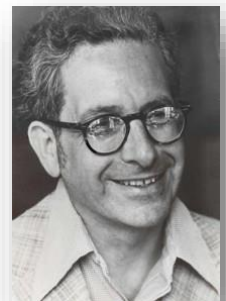
- **Kaplan-Meier Method:**

- Utilize individual survival times.
- Calculates the proportion surviving to each point that death occurs.
- **The (Kaplan-Meier) curve represents the odds of an endpoint having occurred at each point in time (the Hazard).**

E.L. Kaplan **P. Meier**



Died 2006



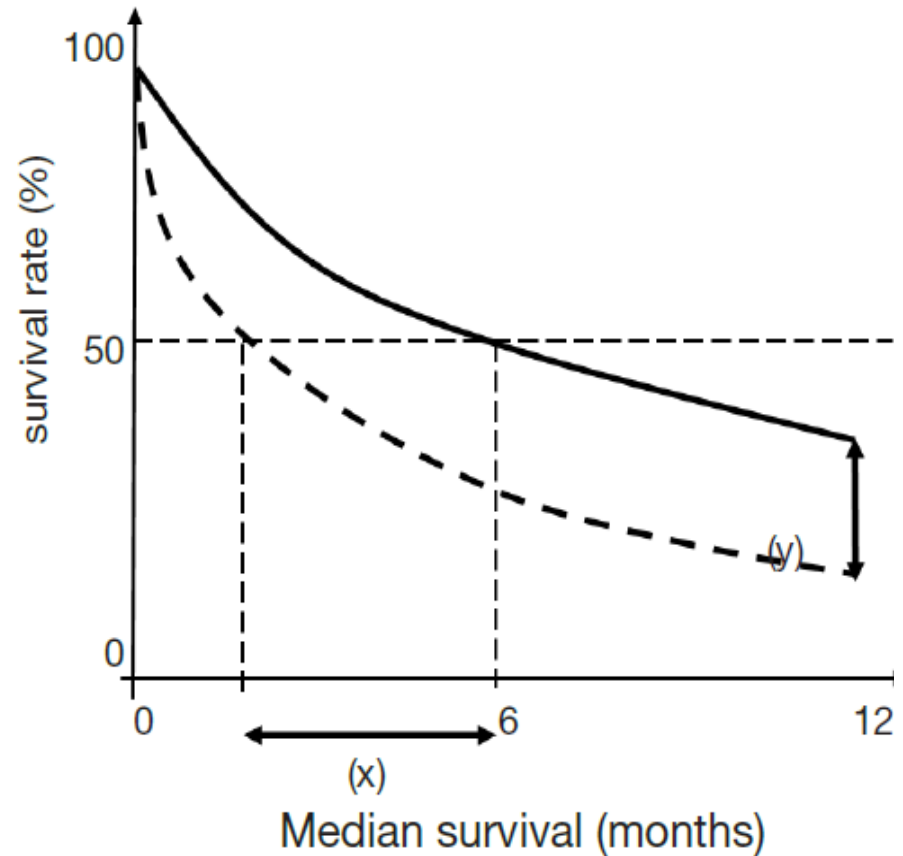
Died 2011

Hazard Ratio (HR)

- The **Hazard Ratio (HR)** summarized the chance of an event occurring in the treatment arm divided by the chance of the event occurring in the control arm.
 - HRs differ from relative risks in that the latter are cumulative over an entire study, while the former represent instantaneous risk over the study time period
- The **HR** represents:
 - The relationship between the instantaneous hazards in the 2 groups
 - The magnitude of distance between the Kaplan-Meier plots in a single number.

Hazard Ratio (HR)

- When **HR** is adopted, it is assumed that:
 - Difference between groups was proportional;
 - Graphically the K-M survival curves displayed a constant distance apart.
- **HR BECOMES MEANINGLESS WHEN THIS ASSUMPTION OF PROPORTIONALITY IS NOT MET!!!!!!**

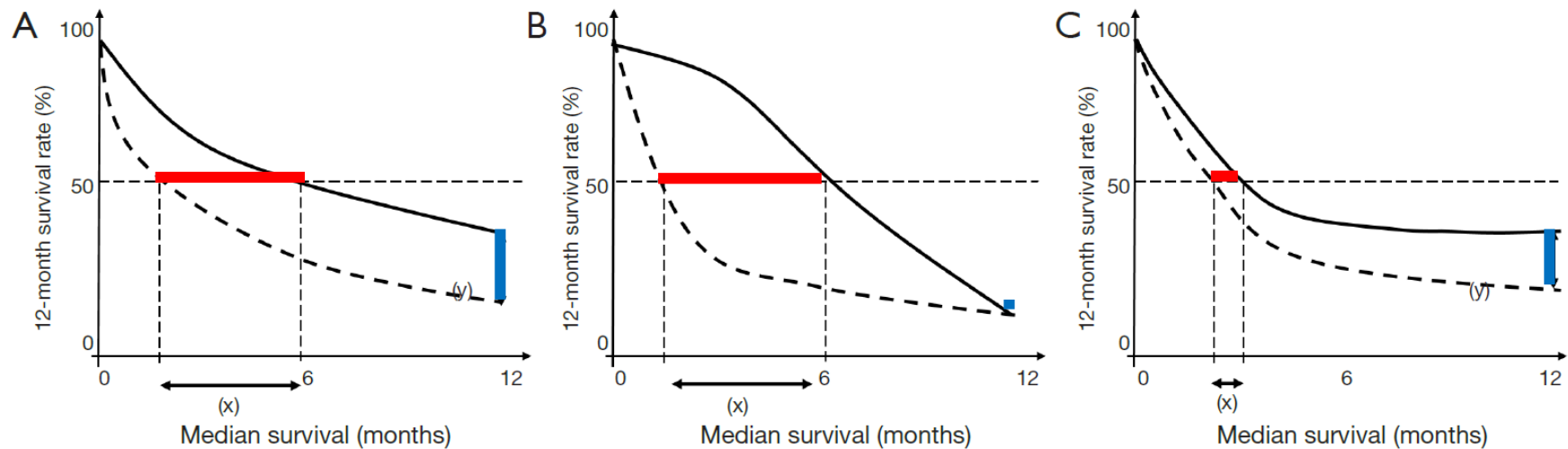


Moving towards a customized approach for drug development: lessons from clinical trials with immune checkpoint inhibitors in lung cancer

Typical survival curves (Kaplan-Meier model) observed in clinical trials

(x) difference in median survival;

(y) 12-month difference in survival rate.

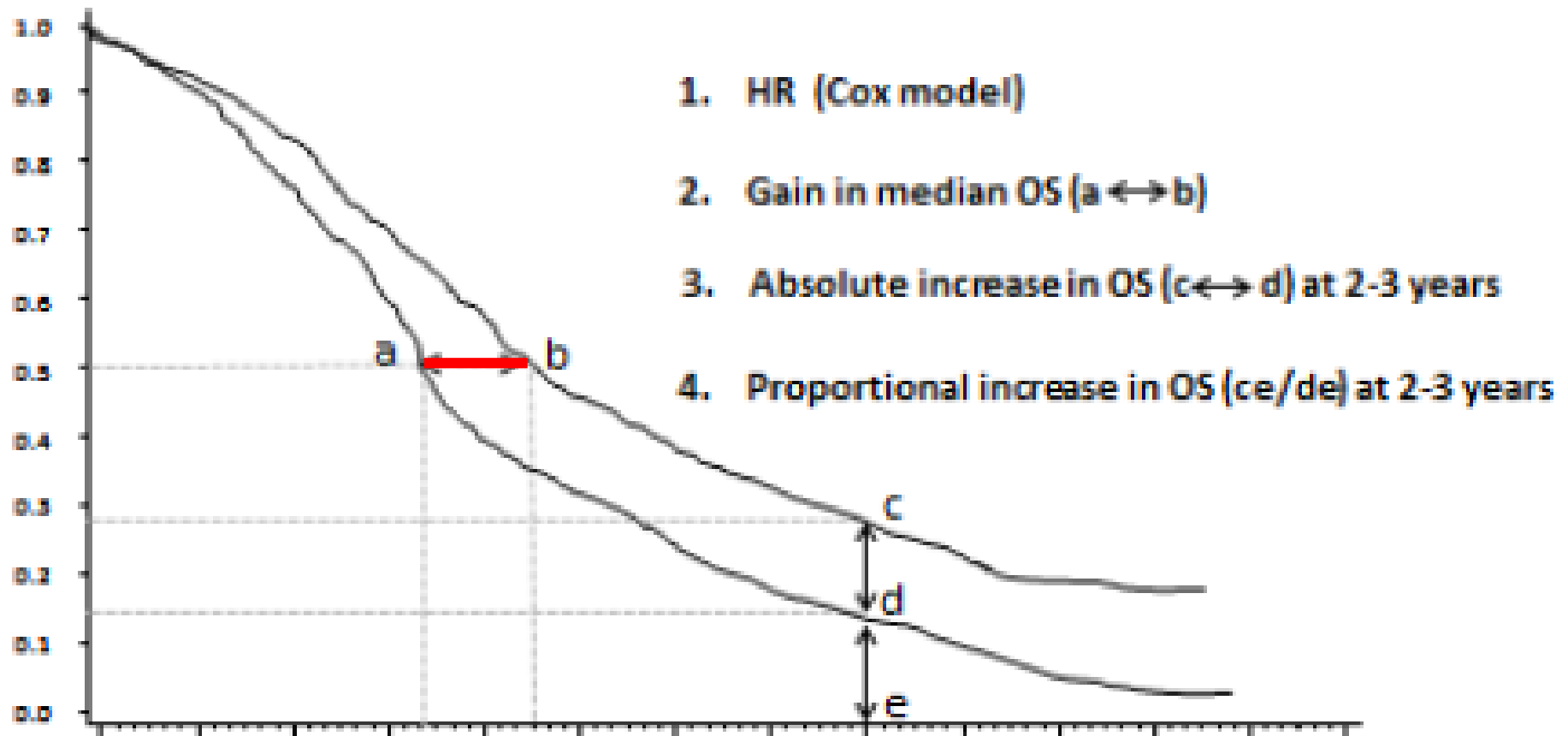


Early Stop for Futility	YES	YES	NO
Correlation with late benefit	YES	NO	NO

Raising the bar for antineoplastic agents: how to choose threshold values for superiority trials in advanced solid tumors

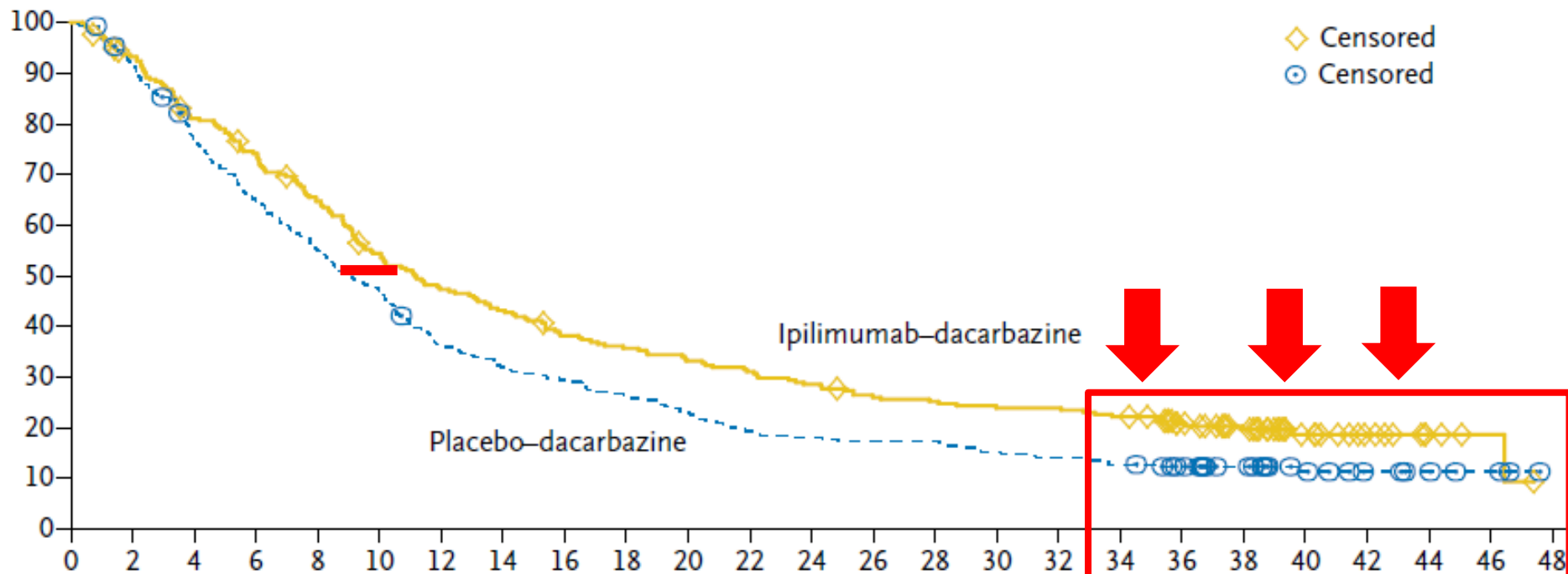
Alberto F Sobrero, Alessandro Pastorino, Daniel J. Sargent, et al.

To establish the concept of minimum *clinically meaningful outcome (mCMO)* of treatment in advanced solid tumors, to establish its threshold and evaluate how many superiority trials of new antineoplastic agents pass this threshold.



Sobrero A et al, CCR 2014

Ipilimumab plus Dacarbazine for Previously Untreated Metastatic Melanoma



- Late 'rates' better than HRs
- Absolute differences at specific time-points
- NNT

Events on the
'right' side

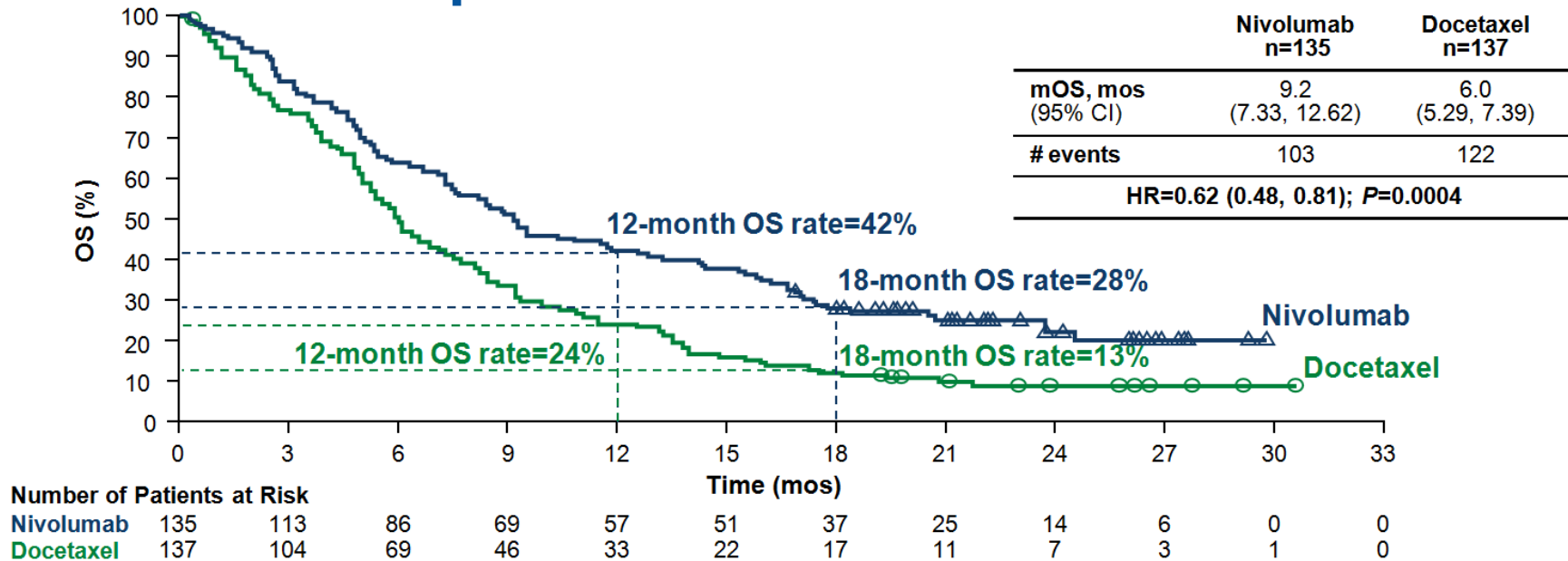
CheckMate 017: Updated OS Data



INTERNATIONAL ASSOCIATION FOR THE STUDY OF LUNG CANCER

16TH WORLD CONFERENCE ON LUNG CANCER
SEPTEMBER 6-9, 2015 → DENVER, COLORADO, USA

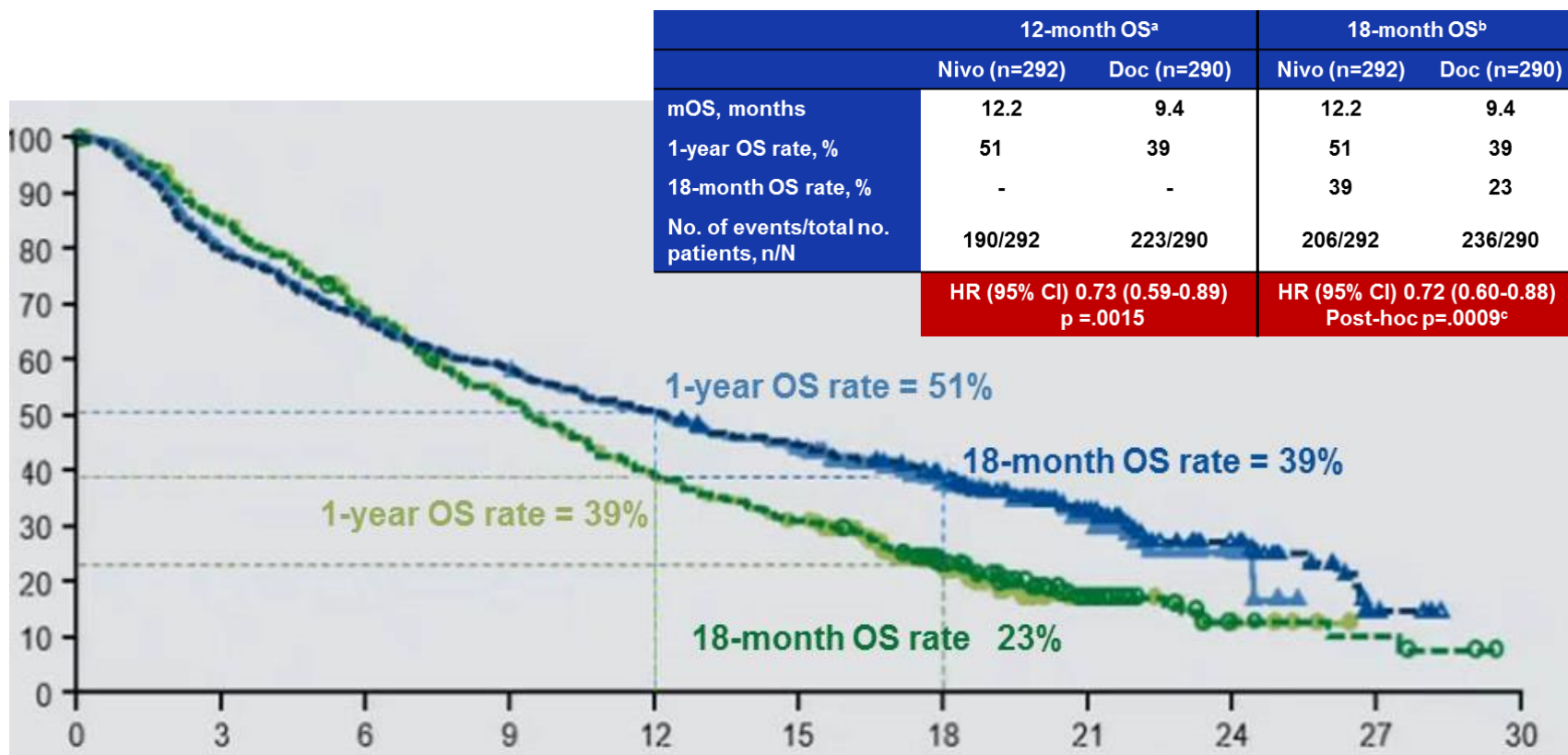
Updated Overall Survival



Minimum follow-up for survival: 18 months

Based on August 2015 DBL.
Symbols refer to censored observations.

CheckMate 057: Updated OS Data



No. at risk (12-month OS)^a

Nivolumab	292	232	194	169	146	123	62	32	9	0	0
Docetaxel	290	244	194	150	111	88	34	10	5	0	0

No. at risk (18-month OS)^b

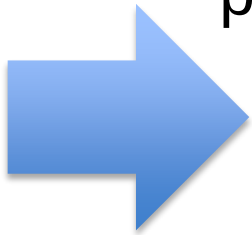
Nivolumab	292	233	195	171	148	128	107	55	27	4	0
Docetaxel	290	244	194	150	111	89	61	23	6	4	0

Minimum F.U. for 12-month OS rate, 13.2 months; for 18-month OS rate, 17.1 months

Horn L, et al, ECC 2015

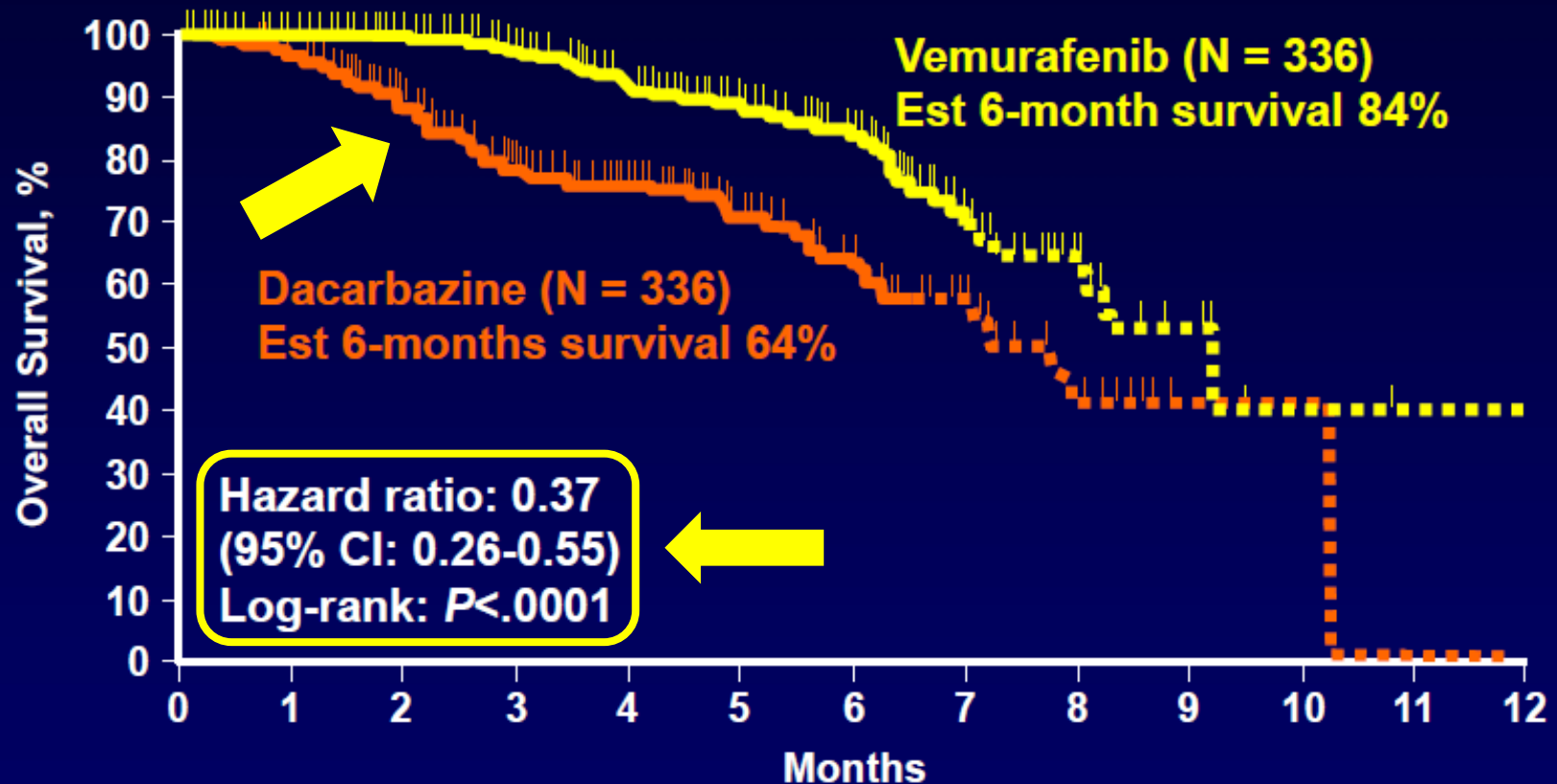
Cancer Survival Analysis

- Thus, patients should be considered:
 - Uncensored:
 - Subjects who are observed until they reach the endpoint of interest (i.e. recurrence or death).
 - Censored :
 - Those patients who survived beyond the end of the follow-up or who are lost to follow-up at some point.
 - Censoring: the loss of subjects from a study before the events of interest has occurred.



*Source: AJCC Cancer Staging Manual, 7th Edition 2010;
Everitt BS, Medical Statistics from A to Z, Ed. 2003*

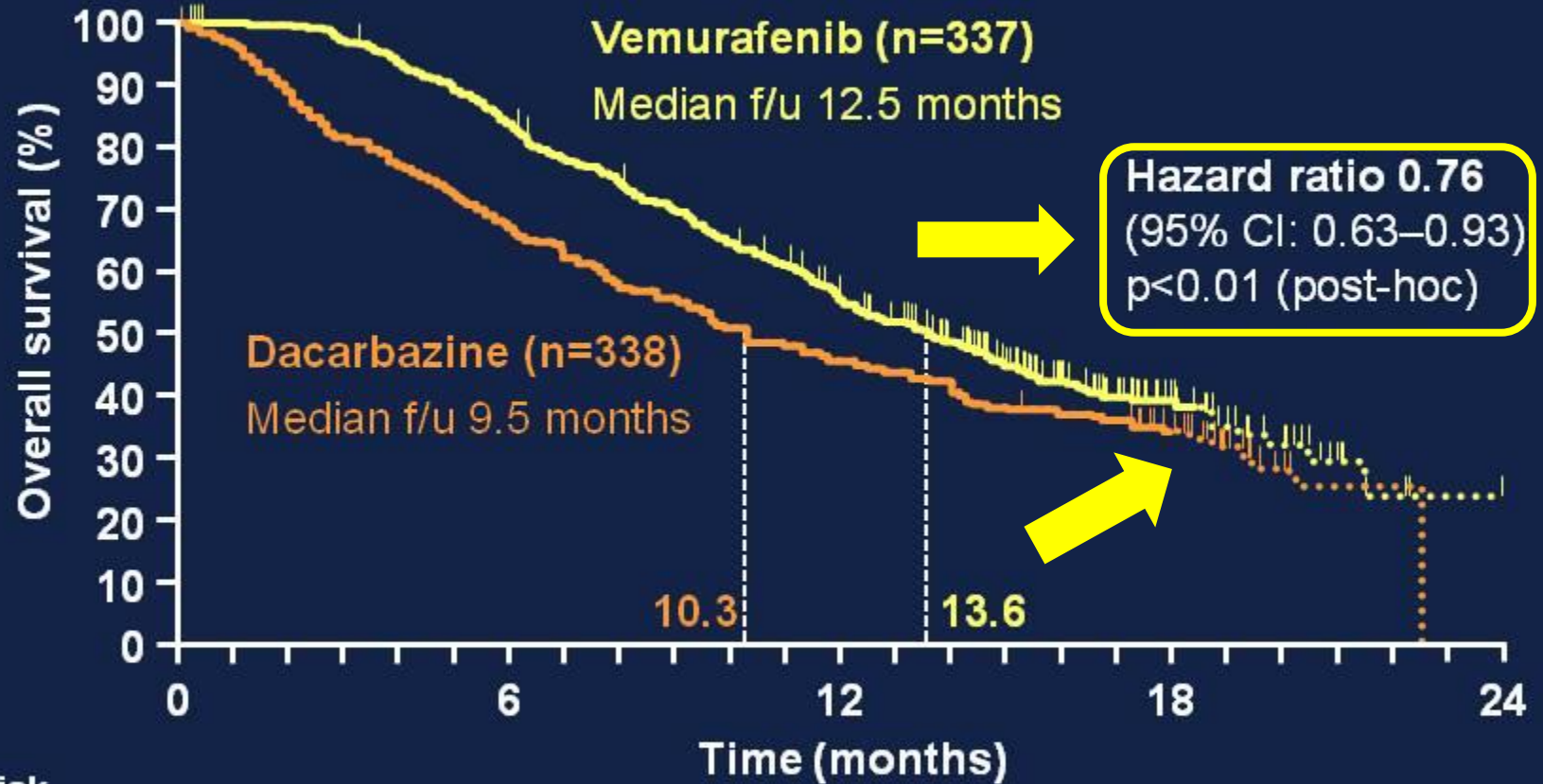
Phase III BRIM3 Trial: Overall Survival (Dec 30, 2010 Cutoff)



No. of patients in follow-up

Dacarbazine	336	283	192	137	98	64	39	20	9	1	1
Vemurafenib	336	320	266	210	162	111	80	35	14	6	1

Overall survival (February 01, 2012 cut-off)



No. at risk

Dacarbazine	338	255	211	173	136	81	34	6	0
Vemurafenib	337	326	280	231	178	109	44	7	1

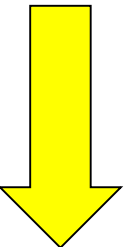
Implication of Meta-Analyses:

1. For Clinical Research

- Hypothesis generation for further RCTs
- Strength control arm in RCTs
- Surrogate end-point development

2. For Clinical Practice

- Relevance of the main question
- Magnitude of the eventual benefit
- Outcome results presentation
 - *Relative Risk (RR -12%)*
 - *Absolute benefits (AB -3-4%)*
 - *Number of patients needed to treat (NNT 25)*



Completeness of reporting trial results: effect on physicians' willingness to prescribe

Marco Bobbio, Brunella Demichelis, Guido Giustetto

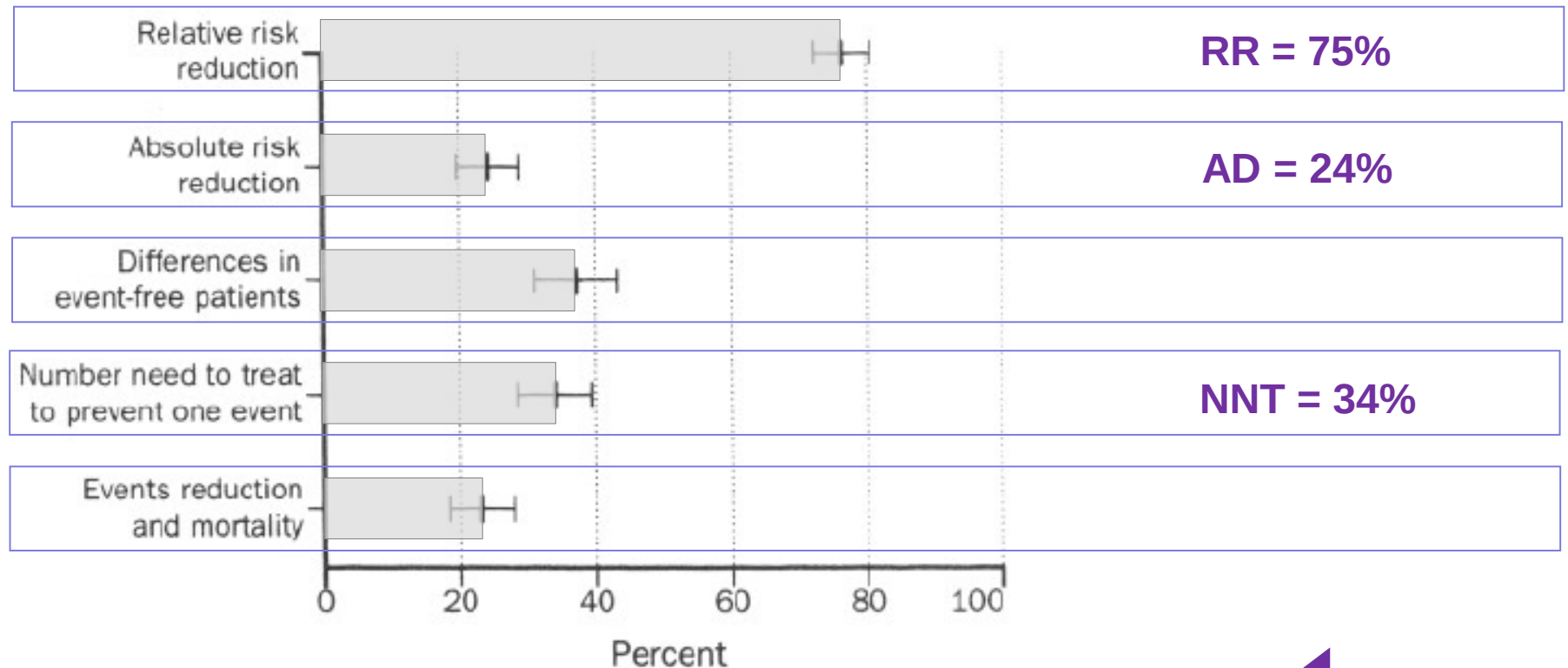


Figure: **Physicians' agreement to prescribe the drug when data are reported in different terms**

Means and 95 percent confidence intervals.

Adjuvant Chemotherapy for Breast Cancer: How Presentation of Recurrence Risk Influences Decision-Making

- **Purpose:**

- to examine the impact of 4 methods of communicating survival benefits on chemotherapy decisions.

- **Results:**

- Absolute survival benefit is the most easily understood method of conveying the information regarding benefit of treatment.
- These results support the hypothesis that the **method used to present information about chemotherapy influences treatment decisions.**

Con il Patrocinio di



STUDI CLINICI: METODOLOGIA

Coordinatore:

Dr.ssa Stefania Gori

*Evento ECM MODULO 1
(formazione di base)*

"A good foundation"



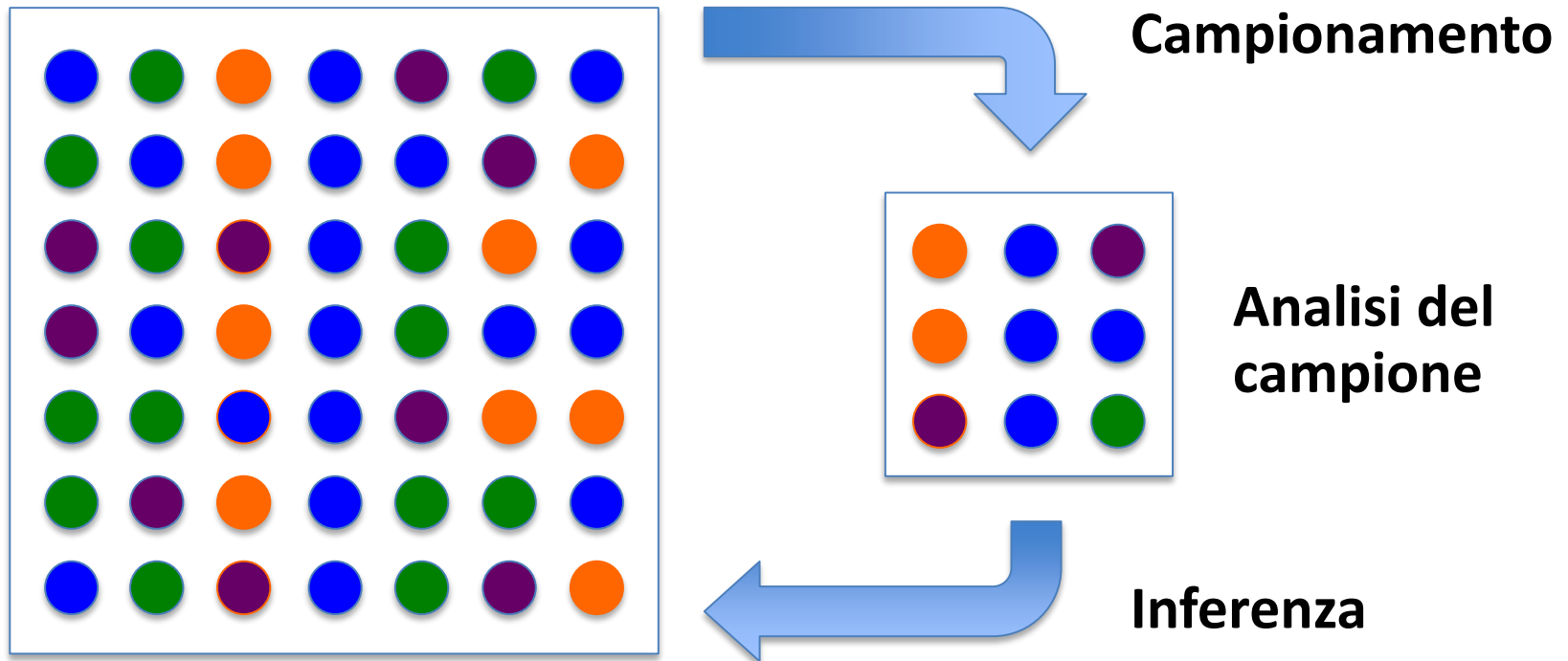
NEGRAR
22-23 Gennaio 2016

Centro Formazione
Ospedale Sacro Cuore
Don Calabria

Sabato 23 gennaio 2016

- Tipi di variabili statistiche
 - ✓ *indici di centralità e variabilità*
- Misure di effetto relative e assolute
 - ✓ *RR, OR, HR / ARR, NTT*
- **Gli errori statistici**
 - ✓ ***alpha error (tipo I)***
 - ✓ ***beta error (tipo II)***
- Il target di rilevanza clinica
 - ✓ *delta (di superiorità)*
 - ✓ *margin M (di non inferiorità)*
- Principi di dimensionamento campionario
 - ✓ per tipo di variabile di outcome
 - ✓ multiplicity e controllo di *alpha error*

Campione $\neq$ Popolazione



L'INFERENZA STATISTICA SI OCCUPA DI ESTENDERE AD UNA INTERA POPOLAZIONE LE INFORMAZIONI CHE SI RICAVALANO DALLE STATISTICHE CAMPIONARIE

POPOLAZIONE, CAMPIONE, INFERENZA

- ✓ La *Popolazione* comprende *tutti* i soggetti con le caratteristiche di interesse per un determinato studio
- ✓ Il *Campione* su cui lo studio verrà condotto potrà essere (o non essere) rappresentativo di tutte le caratteristiche della popolazione (la cosa ci è *ignota*)
- ✓ Lo Sperimentatore è naturalmente portato a trasferire alla *Popolazione* quanto osservato nel *Campione* (*Inferenza*)
- ✓ Il risultato del test statistico di confronto fra due trattamenti è influenzato (ma non sappiamo a favore di chi, e in che misura) dal fatto che lo studio è condotto su un campione

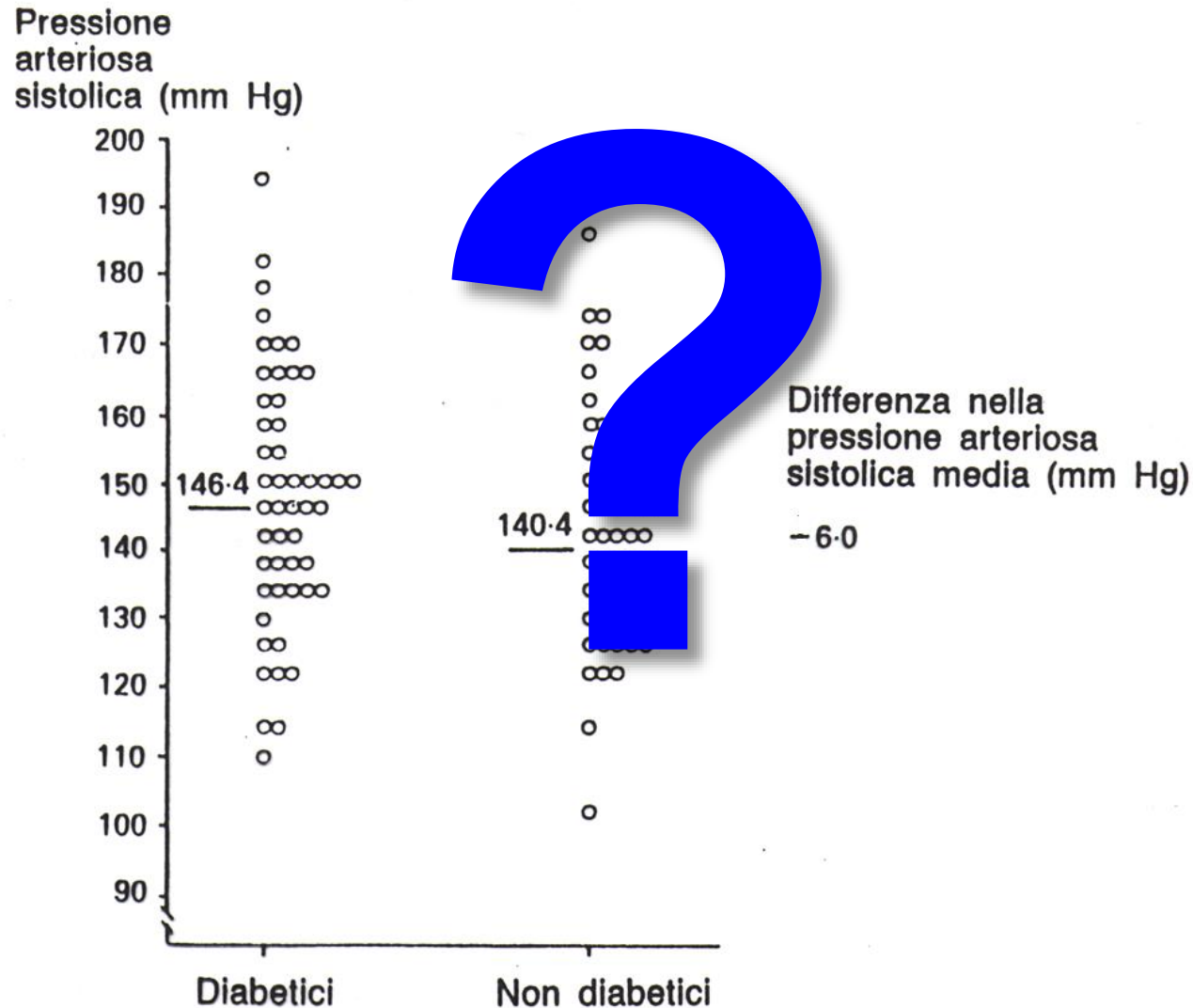
VARIABILE DI TIPO QUALITATIVO...

Table 2. Rates of Radiologic Response.*

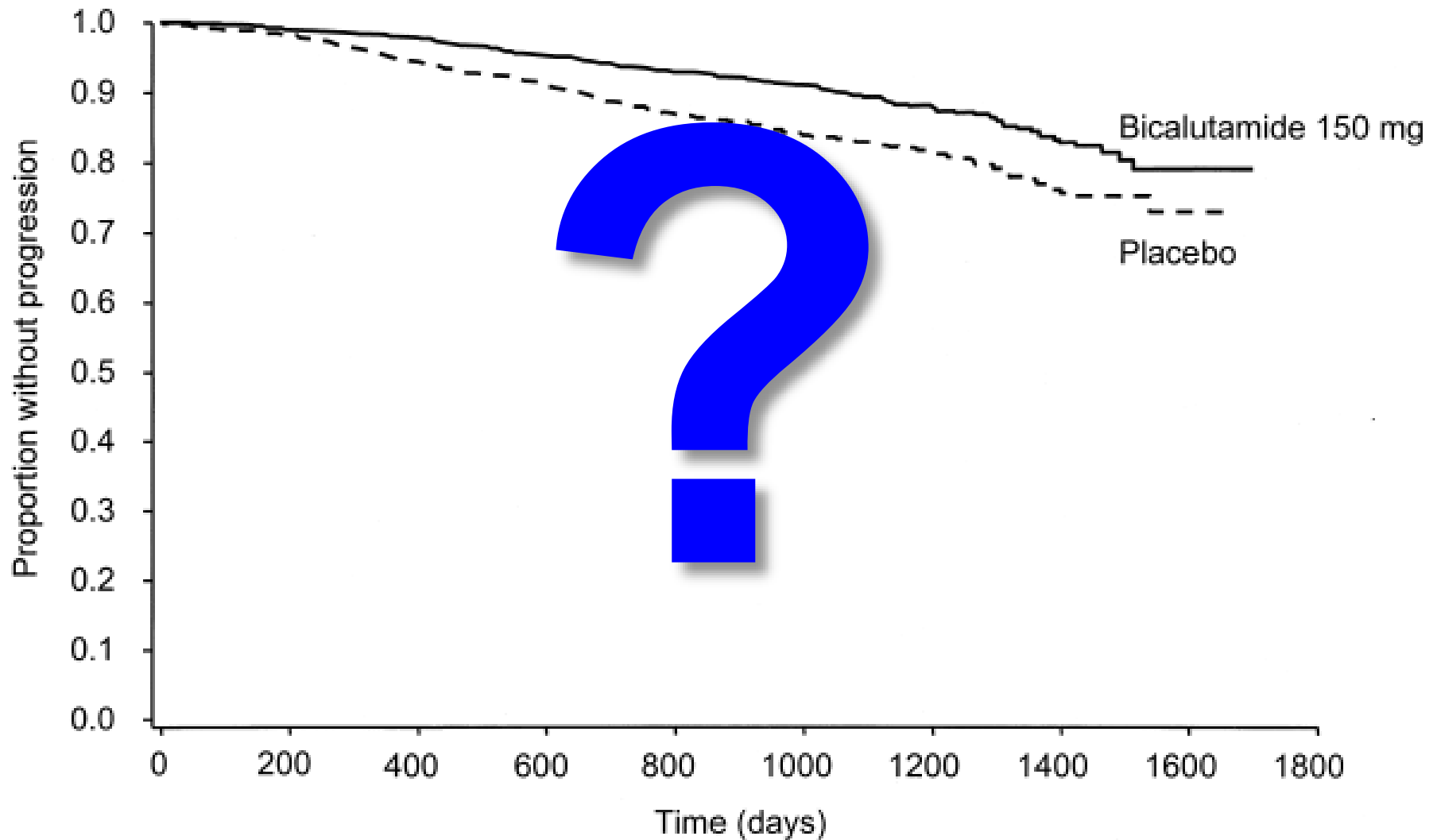


Subgroup and Variable	Cetuximab plus Irinotecan	Cetuximab
No. of patients	218	111
Response — no. (%)		
Complete response	0	0
Partial response	50 (22.9)	12 (10.8)
Stable disease	71 (32.6)	24 (21.6)
Progressive disease	68 (31.2)	59 (53.2)
Could not be evaluated	29 (13.3)	16 (14.4)
Overall response	50 (22.9)	12 (10.8)

VARIABILE DI TIPO QUANTITATIVO...



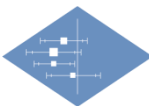
VARIABILE DI TIPO TEMPO A EVENTO...



- **L'effetto osservato è solo uno dei possibili risultati dello studio legati all'assegnazione casuale dei soggetti ai trattamenti.**
- **Se non c'è un reale effetto del trattamento, è comunque possibile che per caso si osservino differenze tra i gruppi.**
- **Se c'è un reale effetto del trattamento, è comunque possibile che per caso non si osservi quella differenza fra i trattamenti.**

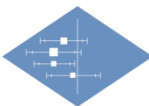
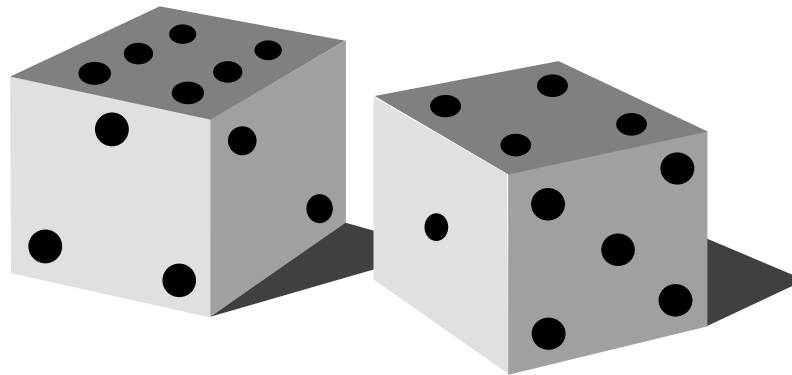
effetto osservato = effetto del trattamento + effetto del caso

(in assenza di errore sistematico!)



IL METODO STATISTICO

- Il metodo statistico ci permette di valutare quanto può essere affidabile il risultato osservato.
- Se i soggetti sono stati assegnati in modo casuale ai trattamenti, si è in grado di calcolare esattamente il ruolo probabile del caso.



Hypothesis Testing and Jury Trials

In tribunale

Test statistico

Presunzione di innocenza

Ipotesi di non differenza tra i trattamenti (H_0)

Evidenza di colpevolezza

Ipotesi di efficacia del trattamento in esame (H_1)

Ma noi non sappiamo quale delle due ipotesi sia quella vera nella realtà (cioè nella popolazione) ...

of Hypothesis

	H_0 is True	H_0 is NOT True
Accept H_0		
Reject H_0		

Il test statistico
può risultare
positivo o
negativo

Hypothesis Testing and Jury Trials

In tribunale

Test statistico

Presunzione di innocenza

Ipotesi di non differenza tra i trattamenti (H_0)

Evidenza di colpevolezza

Ipotesi di efficacia del trattamento in esame (H_1)

Assolvere l'innocente

Condannare il colpevole

Making Decisions: Test of Hypothesis

	H_0 is True	H_0 is NOT True
Accept H_0		
Reject H_0		

Il test potrebbe quindi dare
la risposta corretta...

Hypothesis Testing and Jury Trials

In tribunale	Test statistico
Presunzione di innocenza	Ipotesi di non differenza tra i trattamenti (H_0)
Evidenza di colpevolezza	Ipotesi di efficacia del trattamento in esame (H_1)
Assolvere l'innocente	Accettazione corretta di H_0
Condannare il colpevole	Corretto rifiuto di H_0
Condannare l'innocente	
Assolvere il colpevole	

Making Decisions: Test of Hypothesis

	H_0 is True	H_0 is NOT True
Accept H_0		
Reject H_0		

... oppure potrebbero essere stati commessi degli errori.

Hypothesis Testing and Jury Trials

In tribunale	Test statistico
Presunzione di innocenza	Ipotesi di non differenza tra i trattamenti (H_0)
Evidenza di colpevolezza	Ipotesi di efficacia del trattamento in esame (H_1)
Assolvere l'innocente	Accettazione corretta di H_0
Condannare il colpevole	Corretto rifiuto di H_0
Condannare l'innocente	Errore di 1° tipo
Assolvere il colpevole	Errore di 2° tipo

Making Decisions: Test of Hypothesis

	H_0 is True	H_0 is NOT True	
Accept H_0		 	Type II error
Reject H_0	 		

La probabilità di commettere un errore statistico è strettamente dipendente dalle dimensioni del campione

ERRORI STATISTICI

- ✓ **Errore di 1° tipo** (errore α)
 - quando si conclude per un'efficacia del trattamento sperimentale, quando non lo è nella realtà; *lo studio è falsamente positivo.*
- ✓ **Errore di 2° tipo** (errore β)
 - quando si conclude per una non efficacia del trattamento sperimentale, quando invece lo è nella realtà; *lo studio è falsamente negativo.*
- ✓ Il calcolo delle dimensioni del campione mira a contenere la dimensione degli errori statistici entro valori accettabili (5% per l'errore di 1° tipo e 20% per l'errore di 2° tipo)

Con il Patrocinio di



STUDI CLINICI: METODOLOGIA

Coordinatore:

Dr.ssa Stefania Gori

*Evento ECM MODULO 1
(formazione di base)*

"A good foundation"



NEGRAR
22-23 Gennaio 2016

Centro Formazione
Ospedale Sacro Cuore
Don Calabria



Riflessioni e Sintesi



WHAT?

Cosa è emerso di particolarmente saliente e rilevante?

(indicare almeno 2 risposte condivise)



SO WHAT?

Perché le cose emerse sono così rilevanti?

(indicare almeno 2 risposte condivise)



NOW WHAT?

Quali ricadute nell'immediato per la mia professione?

(indicare almeno 2 risposte condivise)

1. Riflettete da soli per 10 min.

STUDI CLINICI: METODOLOGIA Brescia ECM MODULO 1 (formazione di base) "56 giorni di formazione" INTEGRAB - 20/21 Gennaio 2016 Centro Formazione - Ospedale Santa Croce Don Calisto		Verifica Apprendimento: Sessione n°
nome e cognome		
 RIFLESSIONI E SINTESI sui temi della Sessione		
	WHAT? Cosa è emerso di particolarmente saliente / rilevante?	
		
	SO WHAT? Per quale motivo le cose emerse sono così rilevanti?	
		
	NOW WHAT? Quali ricadute nell'immediato per la mia professione?	
		



WHAT?

Cosa è emerso di particolarmente saliente e rilevante?

(indicare almeno 2 risposte condivise)



SO WHAT?

Perché le cose emerse sono così rilevanti?

(indicare almeno 2 risposte condivise)



NOW WHAT?

Quali ricadute nell'immediato per la mia professione?

(indicare almeno 2 risposte condivise)

1. Riflettete da soli per 10 min.
2. Confrontatevi con i Colleghi del Vostro tavolo per 15 min., declinate un W^3 condiviso e delegate un portavoce
3. Riportate sulla lavagna il Vostro W^3 condiviso su almeno due aspetti ritenuti rilevanti e impattanti sulla professione (in 5 min.)
4. Presentate ai Colleghi degli altri tavoli il Vostro W^3 condiviso

Con il Patrocinio di



STUDI CLINICI: METODOLOGIA

Coordinatore:

Dr.ssa Stefania Gori

*Evento ECM MODULO 1
(formazione di base)*

"A good foundation"



NEGRAR
22-23 Gennaio 2016

Centro Formazione
Ospedale Sacro Cuore
Don Calabria

Sabato 23 gennaio 2016

- Tipi di variabili statistiche
 - ✓ *indici di centralità e variabilità*
- Misure di effetto relative e assolute
 - ✓ *RR, OR, HR / ARR, NTT*
- Gli errori statistici
 - ✓ *alpha error (tipo I)*
 - ✓ *beta error (tipo II)*
- **Il target di rilevanza clinica**
 - ✓ ***delta* (di superiorità)**
 - ✓ ***margin* *M* (di non inferiorità)**
- Principi di dimensionamento campionario
 - ✓ per tipo di variabile di outcome
 - ✓ multiplicity e controllo di *alpha error*

Phase III trials: questions

- The new drug is more effective than other drugs?
 - *Superiority trial (is 'statistically' relevant?)*
- The use of the new drug determines a therapeutic benefit for patients?
 - *Amount of the benefit (is 'clinically' relevant?)*
- The new drug is as effective as other drugs but with fewer side effects (or minor discomfort or lower costs)?
 - *Non-inferiority trial*
- What categories of patients may derive more benefit from the new drug?
 - *Subgroups analysis*

'Sizing' a trial for:.....

Arm A [*Experimental*] vs. Arm B [*Control*]



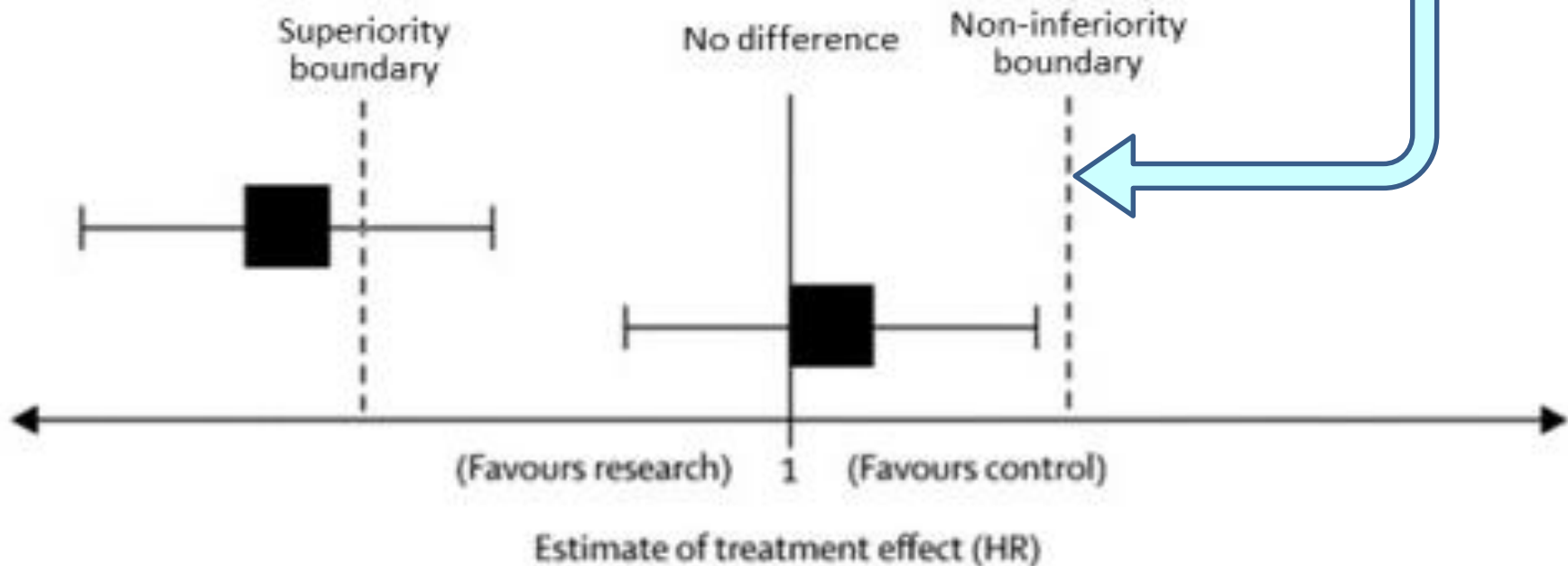
- **Superiority:**
 - A is better than B in terms of the primary outcome, and *this expected difference must be clinically relevant [Minimum Clinically Relevant Benefit: Δ]*
- **Non-inferiority:**
 - A is *inferior* to B in terms of the primary outcome, but *not lower than a clinically pre-specified 'margin' [M] which is considered clinically relevant.*
 - **Ex. You can accept a small degree of lower benefit, if patients experience better tolerability**

Superiority & Non-inferiority

GUIDELINE ON THE CHOICE OF THE NON-INFERIORITY MARGIN

The choice of the non-inferiority margin must always be justified on both clinical and statistical grounds.

It always needs to be tailored specifically to the particular clinical context and no rule can be provided that covers all clinical situations.



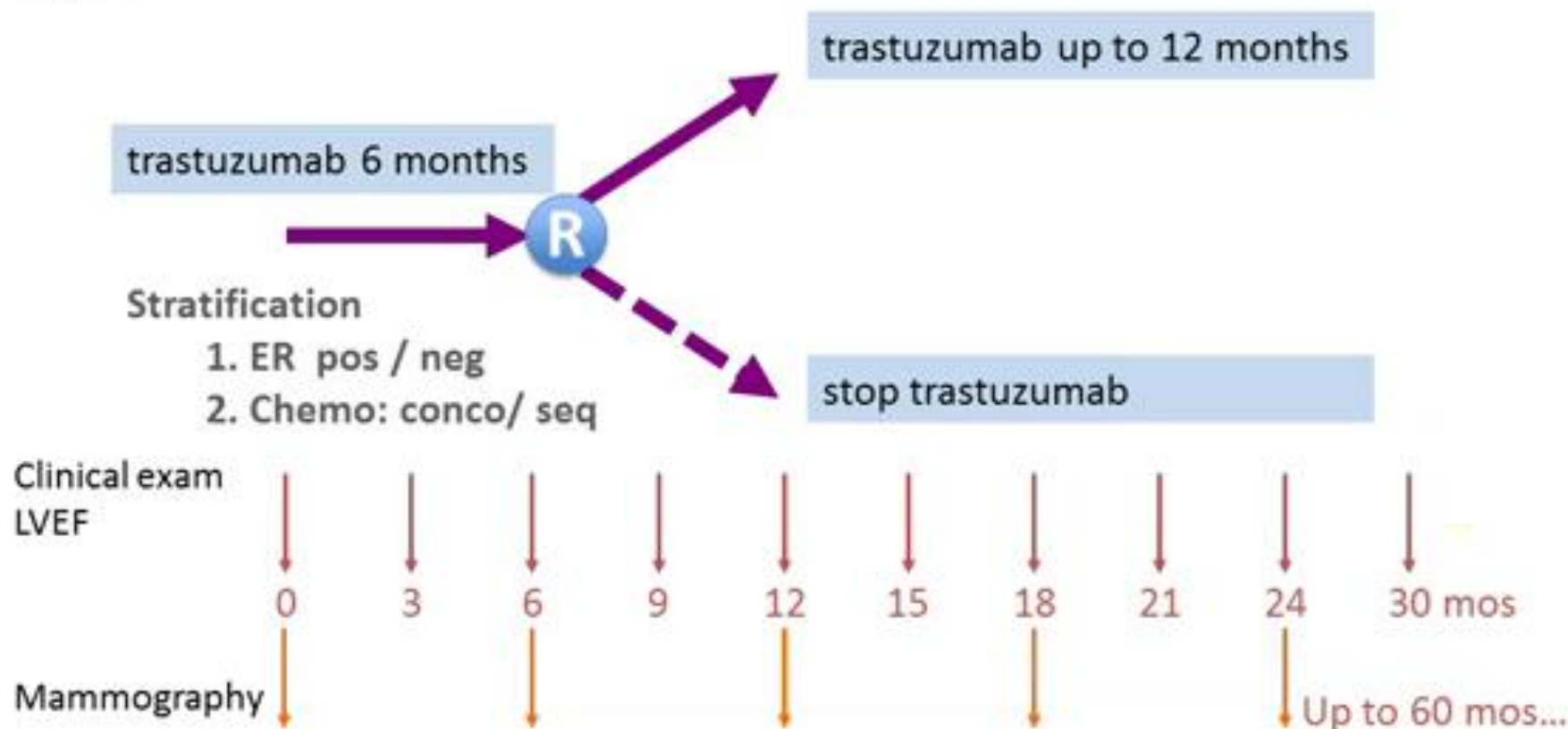
[Minimum Clinically
Relevant Benefit]

M

[Non-inferiority
Margin]



Study design





Objectives

- **Primary objective**

- Compare the effect of 6 months versus 12 months of treatment with trastuzumab in terms of disease-free survival
- Disease-Free Survival (DFS) from randomization to first event
 - local, regional, or distant recurrence, contralateral breast, second non-breast malignancies, or death from any cause

- **Non inferiority randomized trial**

- 2% variation in terms of absolute difference of recurrence
- The 95% CI HR margins should not cross the 1.15 boundary
- 1040 DFS events required for 80% power at 5% level

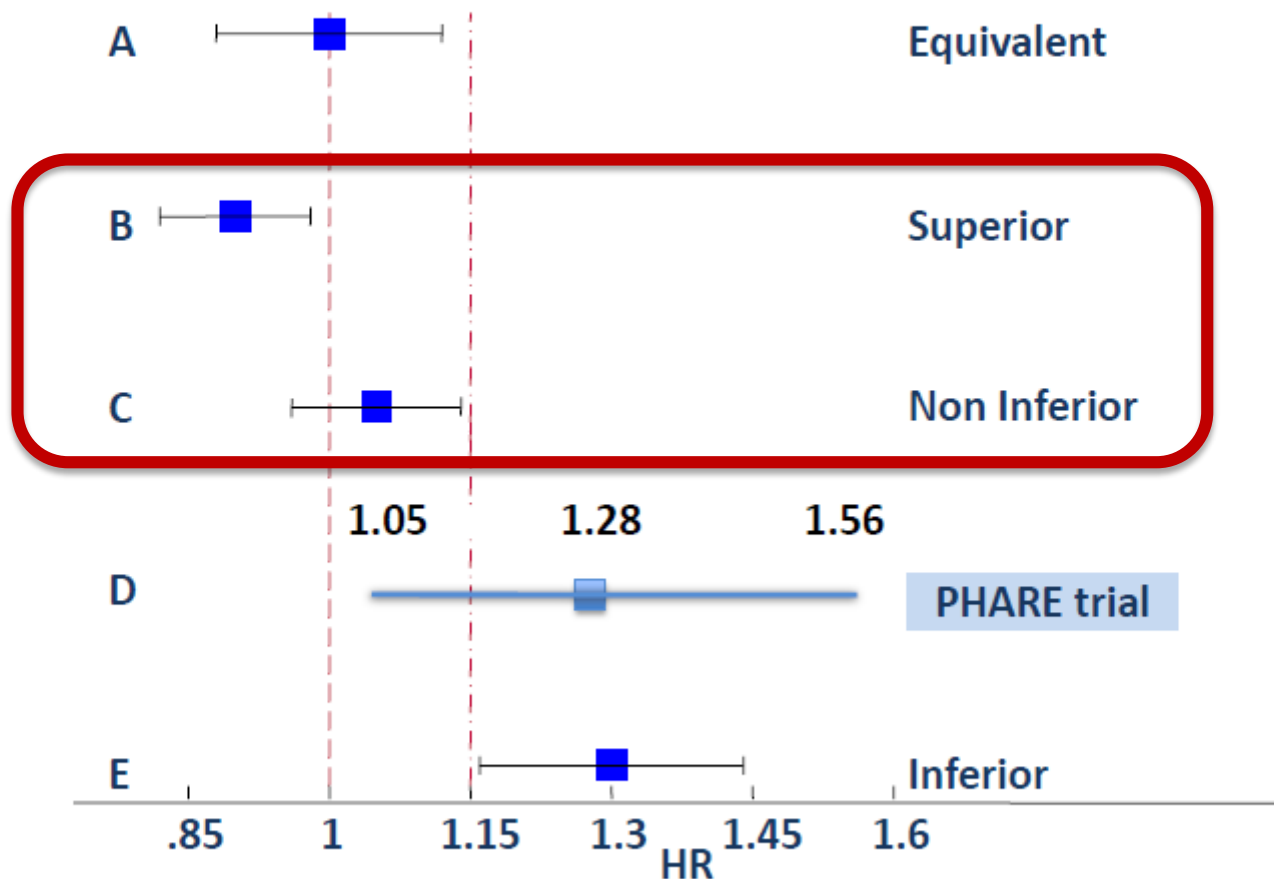
or

- 4 years of accrual and at least 2 years of follow-up
- HR were estimated from the stratified Cox model

- **Accrual target: 3400 patients**



Primary endpoint scenarii



Presentation Outline



- Quality

- Choice of the End-point

- **Quantity**

- Clinical (not simply statistical) Benefit
 - How do we ‘best’ measure that?

- **Value**

‘Shared’ Opininion & Ideas with a series of friends: F. Perrone, M. Di Maio, G. Pappagallo, P. Bruzzi, V. Torri, M. Cinquini, D. Giannarelli, I. Sperduti, F. Roila

Key Elements of Quality Health Care Delivery

- **Safety**
- **Effectiveness**
- **Cost**
- **Readily measured, ascertainable from high-quality medical evidence, and central to the mission of the clinical oncologist.**
- **Patient centeredness**
- **Timeliness**
- **Efficiency**
- **Equity**
- **Not as easily measured and are only rarely reported as outcomes of clinical trials.**

What do we assess in clinical trials?

- Activity:

- ability of the treatment to induce modifications of the disease thanks to which it is assumed that the patient may have a benefit **[Phase II]**

- Efficacy:

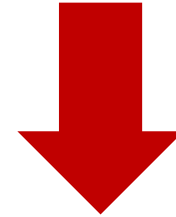
- ability of the treatment to induce a clinical benefit in patients who are administered *in an experimental context* **[Phase III]**

- Effectiveness:

- ability of a treatment to be effective in a *non-experimental, concrete and coincident with the clinical practice* **[are Phase IV, 'Real World' Data]**

American Society of Clinical Oncology Perspective: Raising the Bar for Clinical Trials by Defining Clinically Meaningful Outcomes

Recommended Targets for Meaningful Clinical Trial Goals



Cancer Type	Patient Population	Current Baseline Median OS (months)	Primary End Point		Secondary End Point	
			Improvement Over Current OS That Would Be Clinically Meaningful (months)	Target HRs	Improvement in 1-Year Survival Rate (%)*	Improvement in PFS (months)
Pancreatic cancer	FOLFIRINOX-eligible patients	10 to 11 ¹⁹	4 to 5	0.67 to 0.69	48 → 63	4 to 5
Pancreatic cancer	Gemcitabine or gemcitabine/nab-paclitaxel- eligible patients	8 to 9 ^{20,21}	3 to 4	0.6 to 0.75	35 → 50	3 to 4
Lung cancer	Nonsquamous cell carcinoma	13 ²²	3.25 to 4	0.76 to 0.8	53 → 61	4
Lung cancer	Squamous cell carcinoma	10 ²³	2.5 to 3	0.77 to 0.8	44 → 53	3
Breast cancer	Metastatic triple negative, previously untreated for metastatic disease	18 ^{24,25}	4.5 to 6	0.75 to 0.8	63 → 71	4
Colon cancer	Disease progression with all prior therapies (or not a candidate for standard second- or third-line options)	4 to 6 ²⁶	3 to 5	0.67 to 0.67	25 → 35	3 to 5



**On the basis of HR and
Medians, and correlation
between PFS and OS!!!!**

Ellis LM et al, JCO 2014

American Society of Clinical Oncology Statement: A Conceptual Framework to Assess the Value of Cancer Treatment Options

THE ASCO VALUE FRAMEWORK: ADVANCED DISEASE

Step 1: Determine the regimen's CLINICAL BENEFIT							
1.A. Is <u>Overall Survival</u> (OS) reported?	YES. Assign an <u>OS Score</u> (1 through 5 as shown below) and multiply by 16. Write this number in the box labeled, "OS Score." Proceed to 1.D.						OS Score
	OS Score	1	2	3	4	5	
	Improvement in median OS (% change in median OS)	> 0%-24%	25%-49%	50%-75%	76%-100%	At double the median OS of new regimen, there is a 50% improvement in the fraction of patients surviving	
	NO. Proceed to 1.B.						
1.B. If OS is not reported, is <u>Progression-Free Survival</u> (PFS) reported?	YES. Assign a <u>PFS Score</u> (1 through 5 as shown below) and multiply by 11. Write this number in the box labeled, "PFS Score." Proceed to 1.D.						PFS Score
	PFS Score	1	2	3	4	5	
	Improvement in median PFS (% change in median PFS)	> 0%-24%	25%-49%	50%-75%	76%-100%	At double the median PFS of new regimen, there is a 50% improvement in the fraction of patients without progression or death	
	NO. Proceed to 1.C.						
1.C. If neither OS nor PFS is reported, is <u>Response Rate</u> (RR) reported?	YES. Assign an <u>RR Score</u> (1 through 5 as shown below) and multiply by 8. RR should be calculated by adding the complete response (CR) and partial response (PR) rates. Write this number in the box labeled, "RR Score." Proceed to 1.D.						RR Score
	RR Score	1	2	3	4	5	
	What was the reported response rate (CR + PR)?	> 0%-20%	21%-40%	41%-60%	61%-80%	81%-100%	
1.D. Calculate the <u>Clinical Benefit Score</u>	Insert the OS, PFS, or RR Score. Note: You should have EITHER an OS Score OR a PFS score OR an RR score, NOT MORE THAN ONE. Write the total in the box labeled "Clinical Benefit Score." The maximum allowable points are 80. Proceed to Step 2.						Clinical Benefit Score

A standardised, generic, validated approach to stratify the magnitude of clinical benefit that can be anticipated from anti-cancer therapies: the European Society for Medical Oncology Magnitude of Clinical Benefit

Table 2. Maximal preliminary scores

Treatments with curative intent (form 1)

>5% improvement of survival at ≥ 3 -year follow-up

Improvements in DFS alone HR < 0.60 (primary end point) in studies without mature survival data

Treatments with non-curative intent (form 2)

Primary outcome OS (form 2a)

Control ≤ 12 months

HR ≤ 0.65 AND gain ≥ 3 months OR

Increase in 2-year survival alone $\geq 10\%$

Control > 12 months

HR ≤ 0.70 AND gain ≥ 5 months OR

Increase in 3-year survival alone $\geq 10\%$

Primary outcome PFS (form 2b)

Control ≤ 6 months

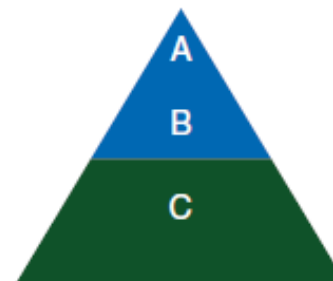
HR ≤ 0.65 AND gain ≥ 1.5 months

Control > 6 months

HR ≤ 0.65 AND gain ≥ 3 months

ESMO MCBS evaluation

Curative



Non-curative



Curative-Evaluation form 1: for new approaches to adjuvant therapy or new potentially curative therapies

Non-curative-Evaluation forms 2a, b or c: for therapies that are not likely to be curative

Figure 3. Visualisation of ESMO-MCB scores for curative and non-curative setting. A & B and 5 and 4 represent the grades with substantial improvement.

Con il Patrocinio di



STUDI CLINICI: METODOLOGIA

Coordinatore:

Dr.ssa Stefania Gori

*Evento ECM MODULO 1
(formazione di base)*

"A good foundation"



NEGRAR
22-23 Gennaio 2016

Centro Formazione
Ospedale Sacro Cuore
Don Calabria

Sabato 23 gennaio 2016

- Tipi di variabili statistiche
 - ✓ *indici di centralità e variabilità*
- Misure di effetto relative e assolute
 - ✓ *RR, OR, HR / ARR, NTT*
- Gli errori statistici
 - ✓ *alpha error (tipo I)*
 - ✓ *beta error (tipo II)*
- Il target di rilevanza clinica
 - ✓ *delta (di superiorità)*
 - ✓ *margin M (di non inferiorità)*
- **Principi di dimensionamento campionario**
 - ✓ **per tipo di variabile di outcome**
 - ✓ **multiplicity e controllo di *alpha error***

DIMENSIONAMENTO CAMPIONARIO

Punti caratterizzanti:

- ✓ definizione del **livello basale di efficacia** (stima dell'effetto nel gruppo di controllo)
- ✓ **beneficio minimo** di interesse clinico (a favore del trattamento sperimentale) che si intende dimostrare

DIMENSIONAMENTO CAMPIONARIO

Punti caratterizzanti:

- ✓ definizione del **livello basale di efficacia** (stima dell'effetto nel gruppo di controllo)
- ✓ **beneficio minimo** di interesse clinico (a favore del trattamento sperimentale) che si intende dimostrare
- ✓ **dimensione (accettabile) dell'errore di 1° tipo = livello di significatività statistica**
 - quando il valore di **p** risultante dal test di significatività è più piccolo del valore soglia (usualmente 5%), si considera lo studio positivo;
 - se il valore di **p** è maggiore del 5%, si considera lo studio negativo
 - non ha niente a che vedere con l'importanza clinica dell'effetto osservato
- ✓ **potenza statistica** (1 meno errore di 2° tipo) che si intende perseguire
 - usualmente fissata all'80%

VARIABILE DI RISPOSTA

- di tipo **quantitativo**
 - assume uno spettro continuo di valori e viene misurata in riferimento a una scala a intervalli costanti.
- di tipo qualitativo
 - esprime categorie di risposta del tipo successo / insuccesso (di un trattamento somministrato).
- del tipo “tempo a evento”
 - rappresenta il tempo trascorso fino al verificarsi (o meno) di un evento.

DIMENSIONAMENTO CAMPIONARIO PER VARIABILI DI TIPO QUANTITATIVO:

- Per una **variabile di risposta di tipo quantitativo**, occorre precisare:
 - media e deviazione standard del gruppo di controllo;
 - differenza di interesse clinico;
 - l'errore di 1° tipo;
 - la potenza del test statistico.

Effect of a monoclonal antibody to PCSK9, REGN727/
SAR236553, to reduce low-density lipoprotein cholesterol in
patients with heterozygous familial hypercholesterolaemia
on stable statin dose with or without ezetimibe therapy:
a **phase 2** randomised controlled trial

*Evan A Stein, Dan Gipe, Jean Bergeron, Daniel Gaudet, Robert Weiss, Robert Dufour, Richard Wu, Robert Pordy
Lancet 2012; 380: 29–36*

Sample size and power calculations were based on the ability to detect a treatment difference in LDL-C of **40%** from baseline to week 12 with a standard deviation range of 20–30% from the completed phase 1 studies.¹⁴

Monotherapy with the PCSK9 inhibitor alirocumab versus ezetimibe in patients with hypercholesterolemia: Results of a 24 week, double-blind, randomized **Phase 3** trial

Eli M. Roth ^{a,*}, Marja-Riitta Taskinen ^b, Henry N. Ginsberg ^c, John J.P. Kastelein ^d, Helen M. Colhoun ^e, Jennifer G. Robinson ^f, Laurence Merlet ^g, Robert Pordy ^h, Marie T. Baccara-Dinet ⁱ

International Journal of Cardiology 176 (2014) 55–61

A sample size of 45 patients per treatment arm was calculated to have 95% power to detect a mean difference between alirocumab and ezetimibe of **20%** in LDL-C percent change from baseline to week 24 using a 2-sided *t*-test with 5% significance, assuming a common standard deviation (SD) of 25% based on a previous alirocumab trial [1] and with an expected rate of exclusion of 5%.

VARIABILE DI RISPOSTA

- di tipo quantitativo
 - assume uno spettro continuo di valori e viene misurata in riferimento a una scala a intervalli costanti.
- di tipo **qualitativo**
 - esprime categorie di risposta del tipo successo / insuccesso (di un trattamento somministrato).
- del tipo “tempo a evento”
 - rappresenta il tempo trascorso fino al verificarsi (o meno) di un evento.

DIMENSIONAMENTO CAMPIONARIO PER VARIABILI DI TIPO QUALITATIVO:

- Per una **variabile di risposta di tipo qualitativo**, occorre precisare:
 - frequenza di “successi” nel gruppo di controllo;
 - differenza di interesse clinico;
 - l’errore di 1° tipo;
 - la potenza del test statistico.

Randomized Trial of Oral Teriflunomide for Relapsing Multiple Sclerosis

Paul O'Connor, M.D., Jerry S. Wolinsky, M.D., Christian Confavreux, M.D.,
Giancarlo Comi, M.D., Ludwig Kappos, M.D., Tomas P. Olsson, M.D., Ph.D.,
Hadj Benzerdjeb, M.D., Philippe Truffinet, M.D., Lin Wang, Ph.D.,
Aaron Miller, M.D., and Mark S. Freedman, M.D., for the TEMSO Trial Group*

N Engl J Med 2011;365:1293-303

A sample of 360 randomly assigned patients per group was required to provide 95% statistical power to detect relative risk reductions of 25% in the annualized relapse rate after 2 years, assuming an annualized relapse rate of 0.74 for the group receiving placebo and a standard deviation of 0.626.

VARIABILE DI RISPOSTA

- di tipo quantitativo
 - assume uno spettro continuo di valori e viene misurata in riferimento a una scala a intervalli costanti.
- di tipo qualitativo
 - esprime categorie di risposta del tipo successo / insuccesso (di un trattamento somministrato).
- del tipo “**tempo a evento**”
 - rappresenta il tempo trascorso fino al verificarsi (o meno) di un evento.

DIMENSIONAMENTO CAMPIONARIO PER VARIABILI TEMPO-A-EVENTO:

- Per una **variabile di risposta di tipo “tempo a evento”**, occorre precisare:
 - stima di sopravvivenza al tempo t nel braccio di controllo;
 - differenza di interesse clinico;
 - l'errore di 1° tipo;
 - la potenza del test statistico;
 - durata di arruolamento e follow-up;
 - max % di pazienti persi al follow-up.

High-Dose Atorvastatin vs Usual-Dose Simvastatin for Secondary Prevention After Myocardial Infarction

The IDEAL Study: A Randomized Controlled Trial

Terje R. Pedersen, MD, PhD

Matti J. Tikkanen, MD, PhD

Christina Lindahl, MD

Ole Faergeman, MD, DMSc

Ingar Holme, PhD

Michael Szarek, MS

John J. P. Kastelein, MD, PhD

Mogens Lytken Larsen, MD, DMSc

John Tsai, MD

Anders G. Olsson, MD, PhD

Fredrik S. Bendiksen, MD

JAMA. 2005;294:2437-2445

The primary clinical outcome was time to first occurrence of a major coronary event, defined as coronary death, hospitalization for nonfatal acute myocardial infarction, or cardiac arrest with resuscitation.

The trial was designed to have 90% power to detect an anticipated 21% relative risk reduction (from 10% to 7.9%) in the primary outcome variable with atorvastatin over 5 years using a 2-tailed α level of .05.

When performing 20 independent tests, we shall expect to have one significant result even though no difference exists.

$$1/20 = 0.05$$

Probability of at least one significant result

Number of tests	Probability
1	0.050
2	0.098
5	0.226
10	0.401
50	0.923

Nel caso di >1 test statistico sullo stesso campione, viene superata la probabilità (convenzionalmente accettata) che il risultato osservato sia falsamente positivo

Apixaban for Extended Treatment of Venous Thromboembolism

Giancarlo Agnelli, M.D., Harry R. Buller, M.D., Ph.D., Alexander Cohen, M.D.,
Madelyn Curto, D.V.M., Alexander S. Gallus, M.D., Margot Johnson, M.D.,
Anthony Porcari, Ph.D., Pharm.D., Gary E. Raskob, Ph.D.,
and Jeffrey I. Weitz, M.D., for the AMPLIFY-EXT Investigators*

N Engl J Med 2012.

STATISTICAL ANALYSIS

The study was designed to test the hypothesis that each dose of apixaban would be superior to placebo with respect to the primary efficacy outcome. Assuming an estimated incidence in the placebo group of 6.8% at 12 months and a decrease in the

A sharper Bonferroni procedure for multiple tests of significance

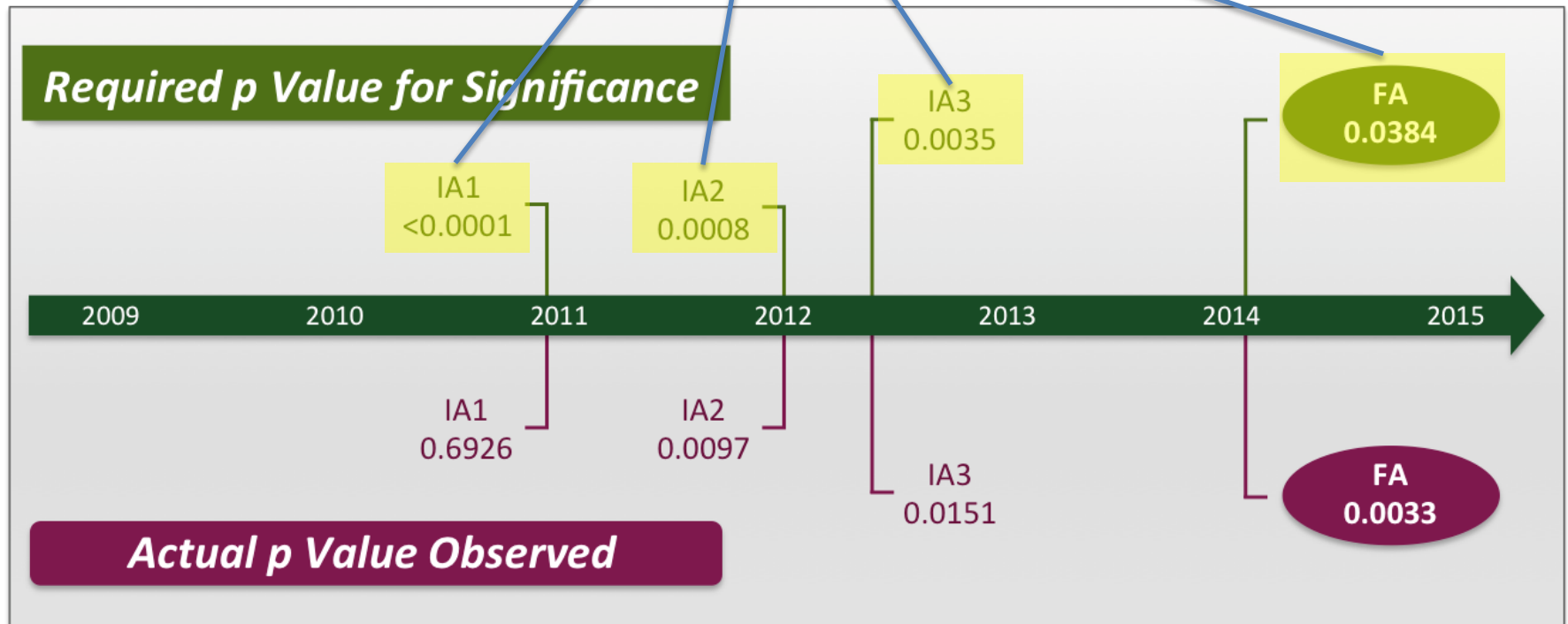
BY YOSEF HOCHBERG

Biometrika (1988), **75**, 4, pp. 800–2

for the study to have 90% power to detect the superiority of apixaban over placebo, at a two-sided alpha level of 0.05, with the use of the Hochberg multiple-testing method.²⁰

Il piano di analisi dello studio COU-302...

Dimensione globale dell'errore di 1° tipo <0.05



**COMMITTEE FOR PROPRIETARY MEDICINAL PRODUCTS
(CPMP)**

**POINTS TO CONSIDER ON MULTIPLICITY ISSUES IN CLINICAL
TRIALS**

2. ADJUSTMENT FOR MULTIPLICITY – WHEN IS IT NECESSARY AND WHEN IS IT NOT?

Sometimes a series of related objectives is pursued in the same trial, where one objective is of greatest importance but convincing results in others would clearly add to the value of the treatment.

In these situations, there is no intention or opportunity to select the most favourable result and, consequently, the individual type I error levels are set equal to the overall type I error level α , i.e. no reduction is necessary.

In such cases the hypotheses may be tested (and confidence intervals may be provided) according to a hierarchical strategy. The hierarchical order may be a natural one (e.g. hypotheses are ordered in time or with respect to the seriousness of the considered variables) or may result from the particular interests of the investigator. Again, no reduction or splitting of α is necessary. The hierarchical order for testing null hypotheses, however, has to be pre-specified in the study protocol.

Effect of a monoclonal antibody to PCSK9, REGN727/
SAR236553, to reduce low-density lipoprotein cholesterol in
patients with heterozygous familial hypercholesterolaemia
on stable statin dose with or without ezetimibe therapy:
a phase 2 randomised controlled trial

Evan A Stein, Dan Gipe, Jean Bergeron, Daniel Gaudet, Robert Weiss, Robert Dufour, Richard Wu, Robert Pordy
Lancet 2012; 380: 29–36

To address the multiple comparisons of the four treatment groups compared with placebo for the primary efficacy endpoint analysis, we applied a hierarchical testing procedure to ensure strong control of the overall type-1 error rate at the two-sided 0·05 significance level.

Effect of a monoclonal antibody to PCSK9, REGN727/
SAR236553, to reduce low-density lipoprotein cholesterol in
patients with heterozygous familial hypercholesterolaemia
on stable statin dose with or without ezetimibe therapy:
a phase 2 randomised controlled trial

*Evan A Stein, Dan Gipe, Jean Bergeron, Daniel Gaudet, Robert Weiss, Robert Dufour, Richard Wu, Robert Pordy
Lancet 2012; 380: 29-36*

To address the multiple comparisons of the four treatment groups compared with placebo for the primary efficacy endpoint analysis, we applied a hierarchical testing procedure to ensure strong control of the overall type-1 error rate at the two-sided 0·05 significance level.

The order used was REGN727 150 mg every 2 weeks versus placebo first; REGN727 300 mg every 4 weeks versus placebo second; REGN727 200 mg every 4 weeks versus placebo third; and finally, REGN727 150 mg every 4 weeks versus placebo.

The hierarchical testing sequence continued only when the higher-order test was statistically significant at the two-sided 5% significant level.

Monotherapy with the PCSK9 inhibitor alirocumab versus ezetimibe in patients with hypercholesterolemia: Results of a 24 week, double-blind, randomized Phase 3 trial

Eli M. Roth ^{a,*}, Marja-Riitta Taskinen ^b, Henry N. Ginsberg ^c, John J.P. Kastelein ^d, Helen M. Colhoun ^e, Jennifer G. Robinson ^f, Laurence Merlet ^g, Robert Pordy ^h, Marie T. Baccara-Dinet ⁱ

International Journal of Cardiology 176 (2014) 55–61

Percent change from baseline in secondary lipid parameters (ITT and on-treatment analysis).

LS mean (SE) % change from baseline to week 24	Alirocumab versus ezetimibe		
	LS mean difference (SE) %	95% CI	p-Value
ITT			
Apo B	−25.8 (3.3)	−32.3 to −19.2	<0.0001 ^a
Non-HDL-C	−25.5 (4.1)	−33.5 to −17.4	<0.0001 ^a
Total cholesterol	−18.7 (3.0)	−24.7 to −12.7	<0.0001 ^a
Lp(a) ^b	−4.4 (5.3)	−14.8 to 5.9	0.4013
TGs ^b	−1.2 (5.9)	−12.7 to 10.3	0.8433
HDL-C	4.4 (2.7)	−1.0 to 9.8	0.1116
Apo A-1	5.3 (2.2)	0.9 to 9.8	0.0196

^a Statistically significant according to the fixed hierarchical approach used to control overall type-I error rate.

Con il Patrocinio di



STUDI CLINICI: METODOLOGIA

Coordinatore:

Dr.ssa Stefania Gori

*Evento ECM MODULO 1
(formazione di base)*

"A good foundation"



NEGRAR
22-23 Gennaio 2016

Centro Formazione
Ospedale Sacro Cuore
Don Calabria



Riflessioni e Sintesi



WHAT?

Cosa è emerso di particolarmente saliente e rilevante?

(indicare almeno 2 risposte condivise)



SO WHAT?

Perché le cose emerse sono così rilevanti?

(indicare almeno 2 risposte condivise)



NOW WHAT?

Quali ricadute nell'immediato per la mia professione?

(indicare almeno 2 risposte condivise)

1. Riflettete da soli per 10 min.

STUDI CLINICI: METODOLOGIA Brescia ECM MEDUSO 1 (Formazione di base) "56 giorni di formazione" INTEGRAB - 20/21 Gennaio 2016 Centro Formazione - Ospedale Santa Croce Don Calisto		Verifica Apprendimento: Sessione n° nome e cognome
 RIFLESSIONI E SINTESI sui temi della Sessione		
	WHAT? Cosa è emerso di particolarmente saliente / rilevante? 	
	SO WHAT? Per quale motivo le cose emerse sono così rilevanti? 	
	NOW WHAT? Quali ricadute nell'immediato per la mia professione? 	



WHAT?

Cosa è emerso di particolarmente saliente e rilevante?

(indicare almeno 2 risposte condivise)



SO WHAT?

Perché le cose emerse sono così rilevanti?

(indicare almeno 2 risposte condivise)



NOW WHAT?

Quali ricadute nell'immediato per la mia professione?

(indicare almeno 2 risposte condivise)

1. Riflettete da soli per 10 min.
2. Confrontatevi con i Colleghi del Vostro tavolo per 15 min., declinate un W^3 condiviso e delegate un portavoce
3. Riportate sulla lavagna il Vostro W^3 condiviso su almeno due aspetti ritenuti rilevanti e impattanti sulla professione (in 5 min.)
4. Presentate ai Colleghi degli altri tavoli il Vostro W^3 condiviso