



Metodologia della Ricerca Clinica

**Corso teorico pratico:
“Come si scrive un protocollo
di ricerca clinica”**

Coordinatore:
Dr.ssa Stefania Gori

*Evento ECM MODULO 4
(formazione avanzata)*



**NEGRAR
24-25 Maggio 2016**

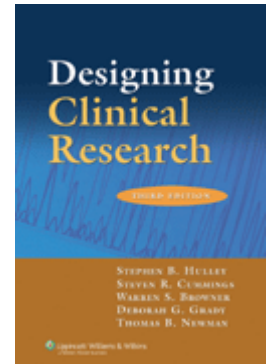
Centro Formazione
Ospedale Sacro Cuore
Don Calabria

- ✓ **Fattibilità e Rilevanza**
- ✓ Obiettivi
- ✓ Disegno di studio
- ✓ Endpoints
- ✓ Criteri di Selezione
- ✓ Protocollo diagnostico
terapeutico
- ✓ Dimensionamento
Campionario

FATTORI DA CONSIDERARE SULL'OPPORTUNITA' DI UNA SPERIMENTAZIONE CLINICA

- **Gravità dell'affezione.**
- **Efficacia (?) delle terapie disponibili.**
- **Tossicità (*scomodità*) delle terapie disponibili rispetto a quelle alternative.**
- **Presumibile superiorità delle terapie sperimentali.**

FATTORI DA CONSIDERARE SULL'OPPORTUNITA' DI UNA SPERIMENTAZIONE CLINICA



- Gravità dell'affezione
- Efficacia (benefici)
- Tossicità e rischi
disponibili.
- Presumibile
sperimentazione

Il criterio FINER

- Feasible = "fattibile"
- Interesting = "interessante" (per il ricercatore)
- Novel = "nuova"
- Ethical = "etica"
- Relevant = "rilevante" (per la conoscenza)

The FINER Criteria for Research Questions

- F Feasible**
- I Interesting**
- N Novel**
- E Ethical**
- R Relevant**

Hulley SB, Cummings SR, Browner WS, Grady DG, Newman TB. *Designing Clinical Research*, 3rd ed. Lippincott Williams & Wilkins: Philadelphia, 2007.

The FINER Criteria for Research Questions

F Feasible
I Interesting
N Novel
E Ethical
R Relevant

Hulley SB, Cummings SR, Browner WS, Grady DG, Newman TB.
Designing Clinical Research, 4th ed. Lippincott Williams & Wilkins:
Philadelphia, 2007.

27

Feasible

- ✓ Conducive Environment
- ✓ Time Frame (availability)
- ✓ Adequate Sampling Pool (statistical power, directness)
- ✓ Expertise (clinical, methodological)
- ✓ Resources

The FINER Criteria for Research Questions

F Feasible
I Interesting
N Novel
E Ethical
R Relevant

Hulley SB, Cummings SR, Browner WS, Grady DG, Newman TB.
Designing Clinical Research, 4th ed. Lippincott Williams & Wilkins:
Philadelphia, 2007.

27

Interesting

- ✓ Serves Immediate Patients Needs
- ✓ Funding (industry, institutional, public)
- ✓ Personal (dedication from inception to publication)
- ✓ Professional (reputation, promotion)
- ✓ Colleagues and Peers (feedback, recruitment)

The FINER Criteria for Research Questions

F Feasible
I Interesting
N Novel
E Ethical
R Relevant

Hulley SB, Cummings SR, Browner WS, Grady DG, Newman TB.
Designing Clinical Research, 4th ed. Lippincott Williams & Wilkins:
Philadelphia, 2007.

27

Novel

- ✓ Meaningful Contribution
- ✓ Master the Empirical Literature
- ✓ Yield Higher Level of Clinical Evidence
- ✓ Validate Results in New Populations
- ✓ Improve Generalizability

The FINER Criteria for Research Questions

F Feasible
I Interesting
N Novel
E Ethical
R Relevant

Hulley SB, Cummings SR, Browner WS, Grady DG, Newman TB.
Designing Clinical Research, 4th ed. Lippincott Williams & Wilkins:
Philadelphia, 2007.

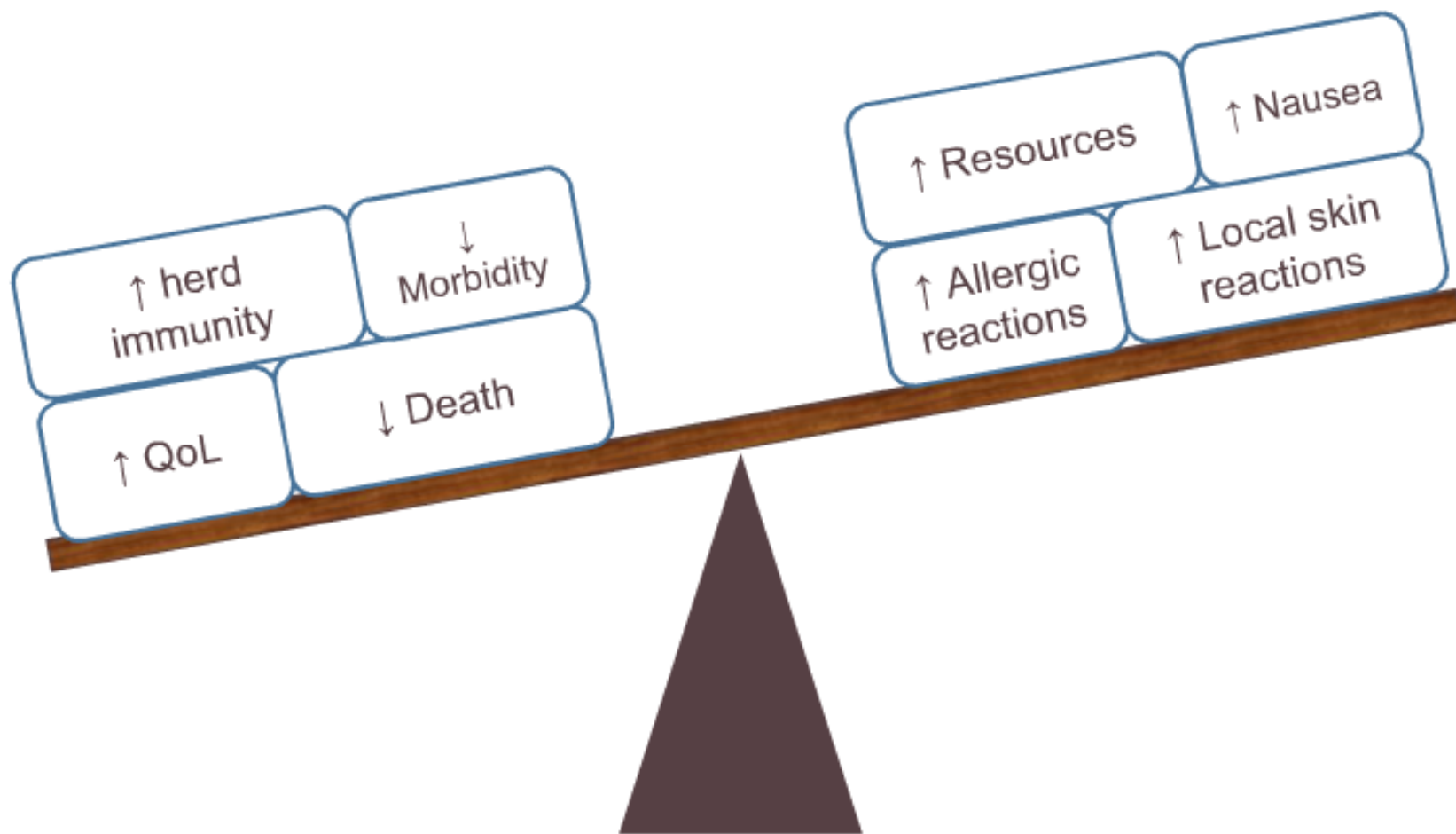
27

Ethical

- ✓ Recruitment / Consent
- ✓ Safety
- ✓ Do NOT Overstate Benefits
- ✓ Do NOT Understate Risks
- ✓ Data Security / Confidentiality



Consequences: desirable and undesirable



The FINER Criteria for Research Questions

F Feasible
I Interesting
N Novel
E Ethical
R Relevant

Hulley SB, Cummings SR, Browner WS, Grady DG, Newman TB.
Designing Clinical Research, 4th ed. Lippincott Williams & Wilkins:
Philadelphia, 2007.

27

Relevant

- ✓ Improve Patient Outcomes / Patient Safety
- ✓ Meets the Needs of Patient Populations
- ✓ Guides Future Research
- ✓ Influences Standard of Care
- ✓ “So What?”

Premessa e stato dell'arte (1)

- **Frequenza e prognosi della malattia oggetto di sperimentazione**
 - ✓ **breve inquadramento della rilevanza epidemiologica e clinica della problematica**
- **Efficacia e tollerabilità dei trattamenti disponibili**
 - ✓ **revisione completa, aggiornata e obiettiva delle possibili opzioni terapeutiche**

Premessa e stato dell'arte (2)

- **Caratteristiche del trattamento in sperimentazione**
 - ✓ **revisione completa delle conoscenze disponibili sul farmaco/trattamento in studio e sui farmaci/trattamenti analoghi**
- **Razionale, obiettivi generali e giustificazione della sperimentazione proposta**



Metodologia della Ricerca Clinica
Corso teorico pratico:
**“Come si scrive un protocollo
di ricerca clinica”**

Coordinatore:
Dr.ssa Stefania Gori

*Evento ECM MODULO 4
(formazione avanzata)*



NEGRAR
24-25 Maggio 2016

Centro Formazione
Ospedale Sacro Cuore
Don Calabria

- ✓ Fattibilità e Rilevanza
- ✓ **Obiettivi**
- ✓ Disegno di studio
- ✓ Endpoints
- ✓ Criteri di Selezione
- ✓ Protocollo diagnostico
terapeutico
- ✓ Dimensionamento
Campionario

Scopi Specifici (Obiettivi)

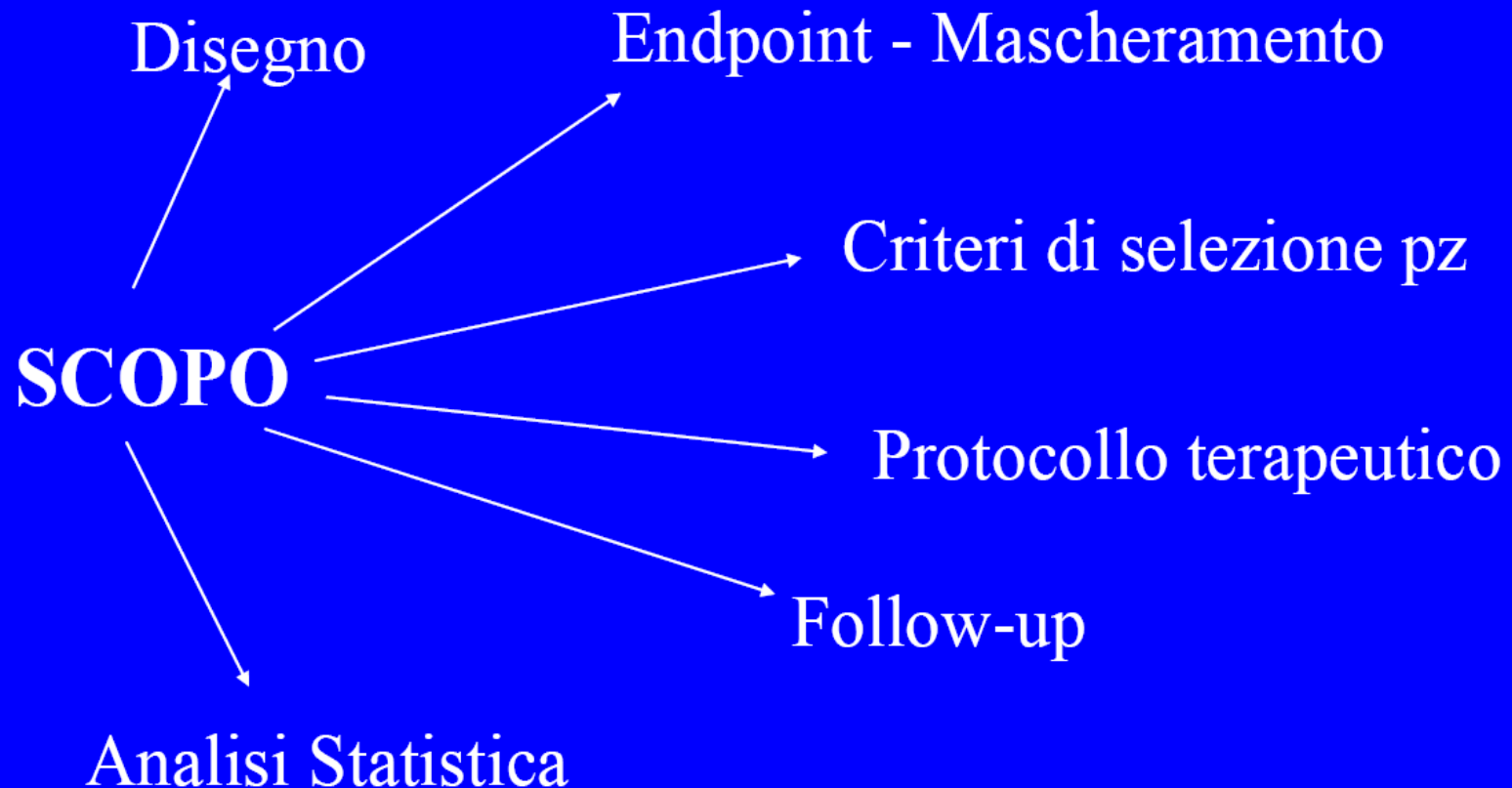
- **Scopo (obiettivo) Primario**

- ✓ quesito cui gli sperimentatori sono più interessati a rispondere, e al quale lo studio vuole dare una risposta;
- ✓ lo scopo primario determina il disegno dello studio e le dimensioni del campione

- **Scopi (obiettivi) Secondari**

- ✓ altri quesiti di interesse, in qualche modo correlati al quesito primario

L'aspetto piu' importante di uno studio e' il suo scopo primario



The 'clinical question' should specify the types of population (participants), types of interventions (and comparisons), and the types of outcomes that are of interest.

The acronym PICO (**P**articipants, **I**nterventions, **C**omparisons and **O**utcomes) helps to serve as a reminder of these.

P • Population

Used to first develop the health care question

I • Intervention

C • Comparison

O • Outcomes

Used to determine if the evidence found directly answers the health care question



Important Questions

Should be
from practice
NOT
evidence driven

Primary and Secondary Questions/Objectives

- Common error – Sinking ship: Avoid overloading the study with too many objectives and too much data collection
- A single primary question around which to focus the development of the protocol and sample size estimates
- Secondary research questions: can be related to the primary question or to other hypotheses





Hypotheses and Objectives

- KISS – keep it simple, stupid
- Too many objectives compromise a trial
 - A single hypothesis and a few secondary hypotheses
 - Can't study everything
- If you can't power an endpoint, it shouldn't be a primary or secondary objective

**Fasi
(tradizionali)**

III-IV

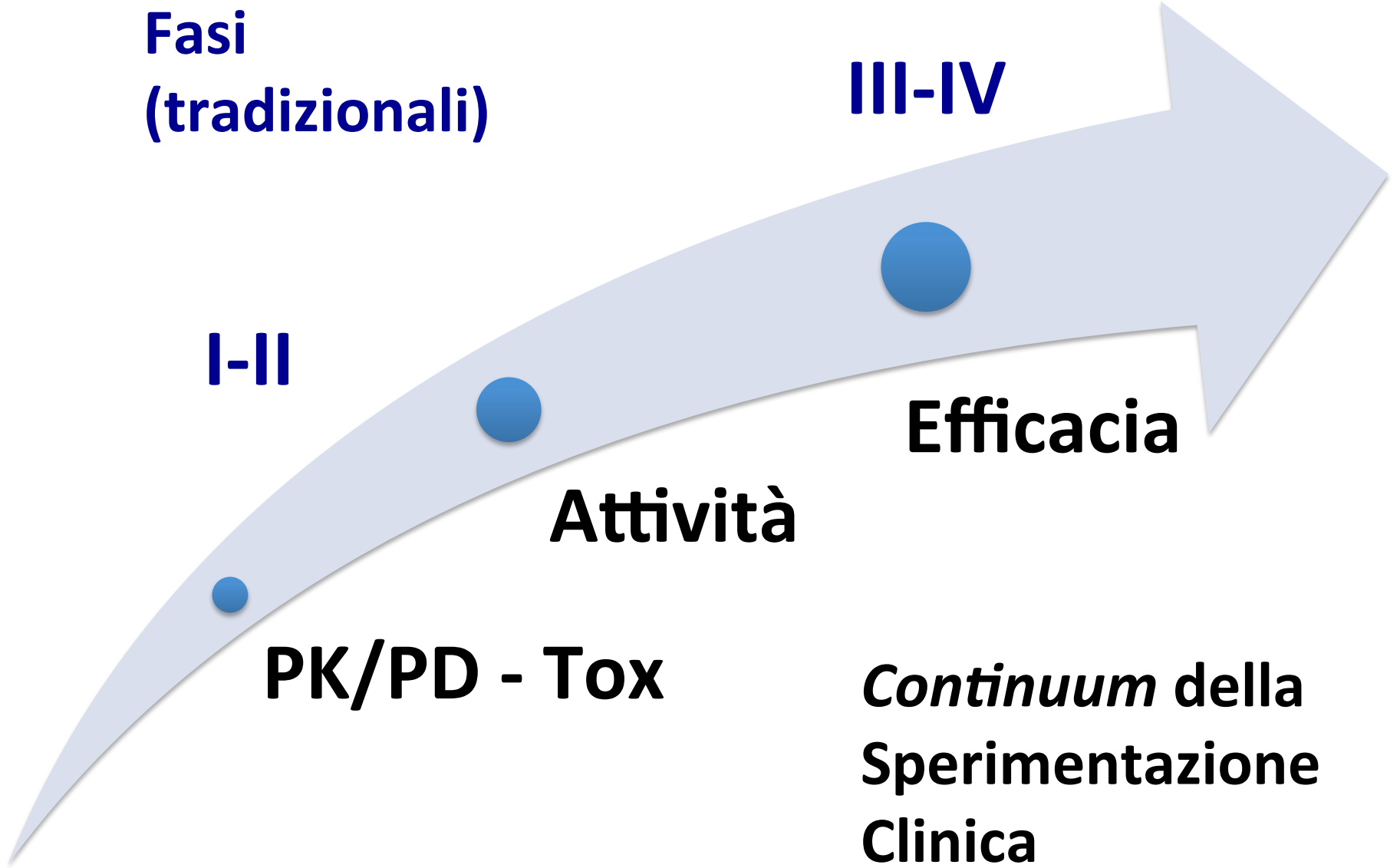
I-II

Attività

Efficacia

PK/PD - Tox

***Continuum della
Sperimentazione
Clinica***



Attività vs Efficacia

- **Attività**

- capacità di un trattamento di indurre le modificazioni attraverso le quali *si presume* di indurre dei benefici

- **Efficacia**

- capacità di un trattamento di indurre i benefici per ottenere i quali esso viene somministrato



Metodologia della Ricerca Clinica

**Corso teorico pratico:
“Come si scrive un protocollo
di ricerca clinica”**

Coordinatore:
Dr.ssa Stefania Gori

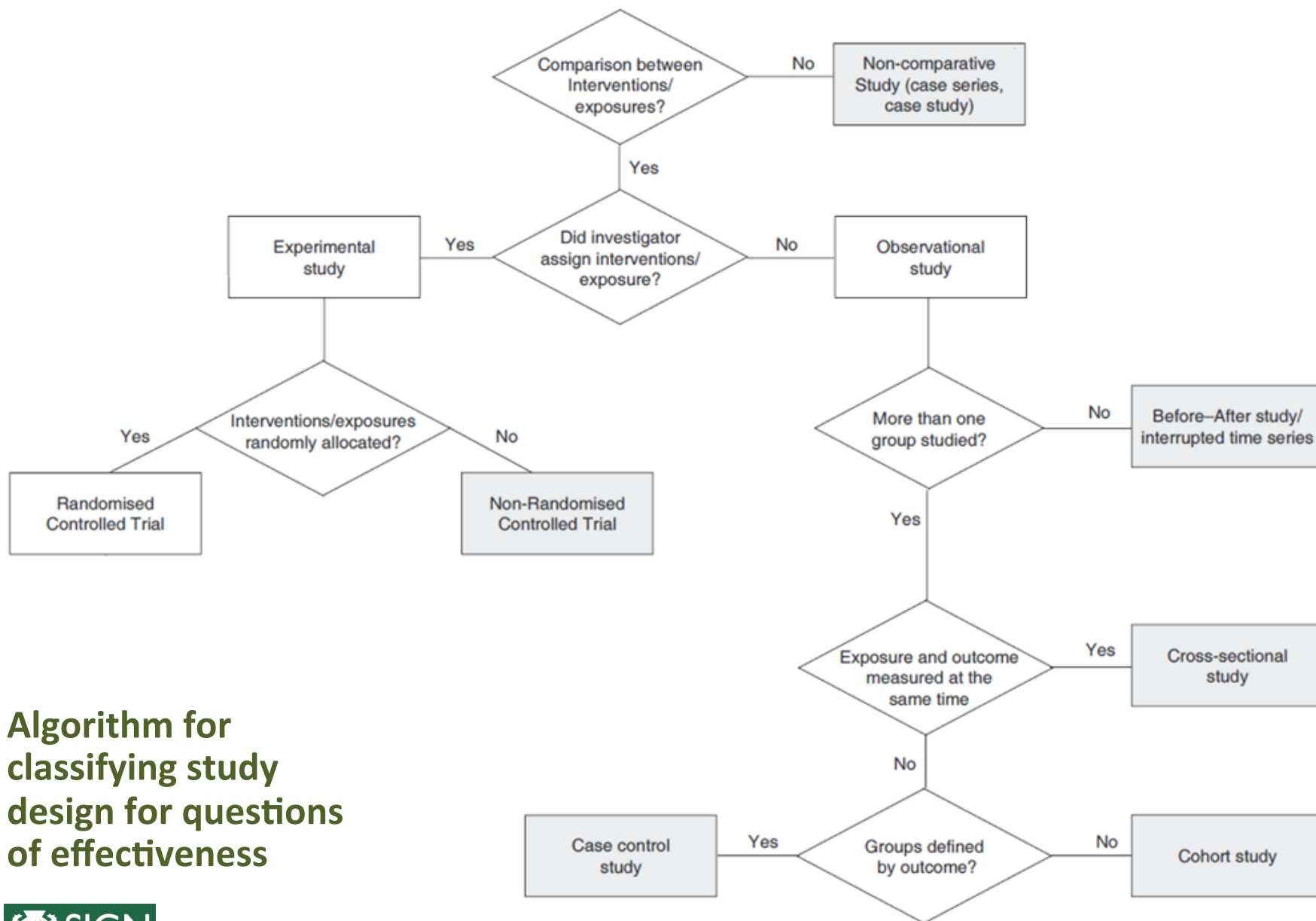
*Evento ECM MODULO 4
(formazione avanzata)*



**NEGRAR
24-25 Maggio 2016**

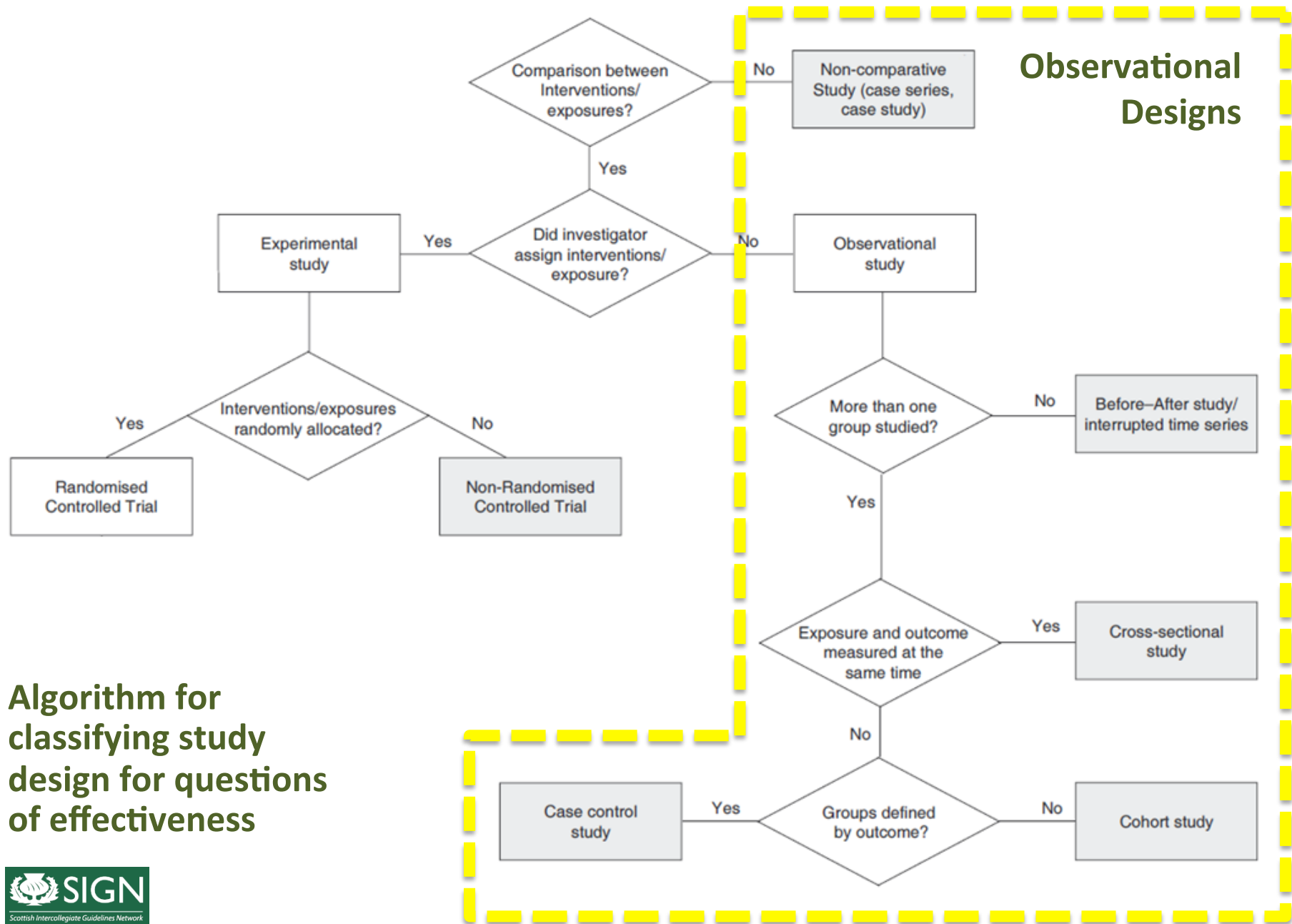
Centro Formazione
Ospedale Sacro Cuore
Don Calabria

- ✓ Fattibilità e Rilevanza
- ✓ Obiettivi
- ✓ **Disegno di studio**
- ✓ Endpoints
- ✓ Criteri di Selezione
- ✓ Protocollo diagnostico
terapeutico
- ✓ Dimensionamento
Campionario



Algorithm for classifying study design for questions of effectiveness

Algorithm for classifying study design for questions of effectiveness



OBSERVATIONAL STUDY: A DEFINITION

An observational study draws inferences from a sample to a population where the independent variable is **not under the control** of the researcher.

The term observational study covers a wide range of study designs, a common feature of which is that they are noninterventive, in the sense that the **study protocol does not determine the precise features of any therapy** given to the participants in the study.

Observational Studies

“EPIDEMIOLOGIC” Vs “THERAPEUTIC”

Before-After

Surveys disease status before and after an intervention

Cross-Sectional

Provide information on prevalence of a particular condition at a single time point (time window)

Case-Control

Identify predictors of a particular outcome

Cohort

Identify the incidence of a particular outcome over time

Non-comparative case series

Report outcomes of patients who received a specific intervention

Comparative case series

Compare outcomes between patients who received different interventions

Observational Studies

“EPIDEMIOLOGIC” Vs “THERAPEUTIC”

Before-After

Surveys disease status before and after an intervention

Cross-Sectional

Provide information on prevalence of a particular condition at a single time point (time window)

Case-Control

Identify predictors of a particular outcome

Cohort

Identify the incidence of a particular outcome over time

Non-comparative case series

Report outcomes of patients who received a specific intervention

Comparative case series

Compare outcomes between patients who received different interventions

The Value of Observational Cohort Studies for Cancer Drugs

Randomized controlled trials — the gold standard for clinical drug evaluation — can't always predict adverse events in real-world settings. For the new cancer therapies, observational cohort studies (OCSs) can help evaluate their effects in broader populations and provide valuable information for future clinical trials.

BY DAVID R. SPIGEL, MD **BIOTECHNOLOGY HEALTHCARE** · SUMMER 2010

WHAT IS AN OCS?

An OCS is an analysis of a group of individuals who have specific features in common and who are followed over a defined period of time.

Prospective OCSs are designed to examine pre-defined primary outcomes.

Post-approval OCSs generally follow a single cohort, although patient subgroups may be analyzed separately.

To represent a broad and diverse patient base and to detect rare adverse events, large community-based, multicenter OCSs are useful in the post-approval setting for new therapeutics.

Observational Studies

“EPIDEMIOLOGIC” Vs “THERAPEUTIC”

Before-After

Surveys disease status before and after an intervention

Cross-Sectional

Provide information on prevalence of a particular condition at a single time point (time window)

Case-Control

Identify predictors of a outcome

Cohort

Identify the incidence of outcome over time

Non-comparative case series

Report outcomes of patients who received a specific intervention

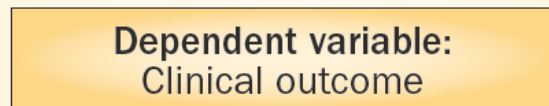
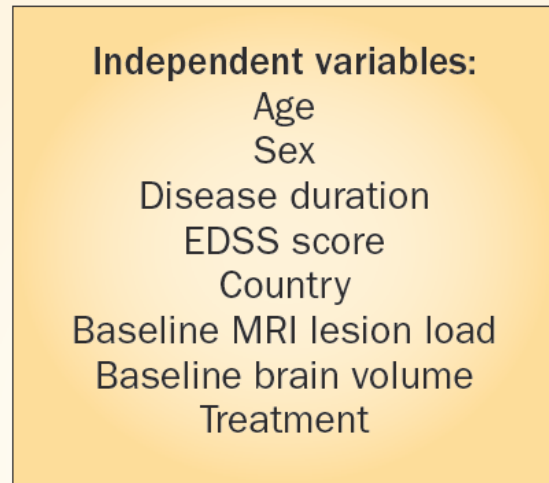
Comparative case series

Compare outcomes between patients who received different interventions

Any observed difference between the outcomes of study arms may be attributable to baseline differences rather than to a true treatment effect.

Standard, multivariate Vs Propensity Score adjustment

Standard adjustment for baseline factors
Multivariate analysis



INTERFERON BETA AND DISABILITY PROGRESSION IN MS

To address the methodological limitations of earlier studies, we used a comparable baseline for the treated and untreated cohorts, adjusted for potential confounders, increased treatment exposure accuracy, and addressed immortal time bias by treating exposure as a time-dependent variable.

JAMA. 2012;308(3):247-256

Comparative Effectiveness Research in Oncology Methodology: Observational Data

Dawn L. Hershman and Jason D. Wright

J Clin Oncol 30:4215-4222. © 2012 by American Society of Clinical Oncology

Propensity Score Analysis

Propensity score analyses attempt to balance covariates between experimental groups. Using multivariable modeling, the characteristics of a cohort are used to calculate the probability of receiving the intervention. This probability is the propensity score.

Le caratteristiche della coorte vengono usate per calcolare la probabilità di (*propensità* a) ricevere l'uno o l'altro dei trattamenti a confronto. Tale probabilità è espressa dal *propensity score*.

Standard, multivariate Vs Propensity Score adjustment

Standard adjustment for baseline factors Multivariate analysis

A score is assigned to each patient, estimating their probability of being treated or untreated in the future.

This score is calculated using a standard logistic regression model with all the unbalanced baseline characteristics as covariates.

Once the propensity score has been estimated, the comparison of the primary outcome between treated and untreated patients is adjusted by the propensity score.

Propensity score adjustment Multivariate analysis

Independent variables:

Age
Sex
Disease duration
EDSS score
Country
Baseline MRI lesion load
Baseline brain volume



Dependent variable:
Treatment status

Independent variables:

Propensity score
Treatment



Dependent variable:
Clinical outcome

Standard, multivariate Vs Propensity Score adjustment

Standard adjustment for baseline factors Multivariate analysis

Independent variables:

Age
Sex

Propensity score adjustment cannot solve the problem of unknown or unmeasurable confounders, and a substantial selection bias can never be ruled out in any observational study, no matter how well planned or designed.

Clinical outcome

Propensity score adjustment Multivariate analysis

Independent variables:

Age
Sex
Disease duration
EDSS score
Country
Baseline MRI lesion load
Baseline brain volume

Step 1:
Propensity
score
creation

Dependent variable:
Treatment status

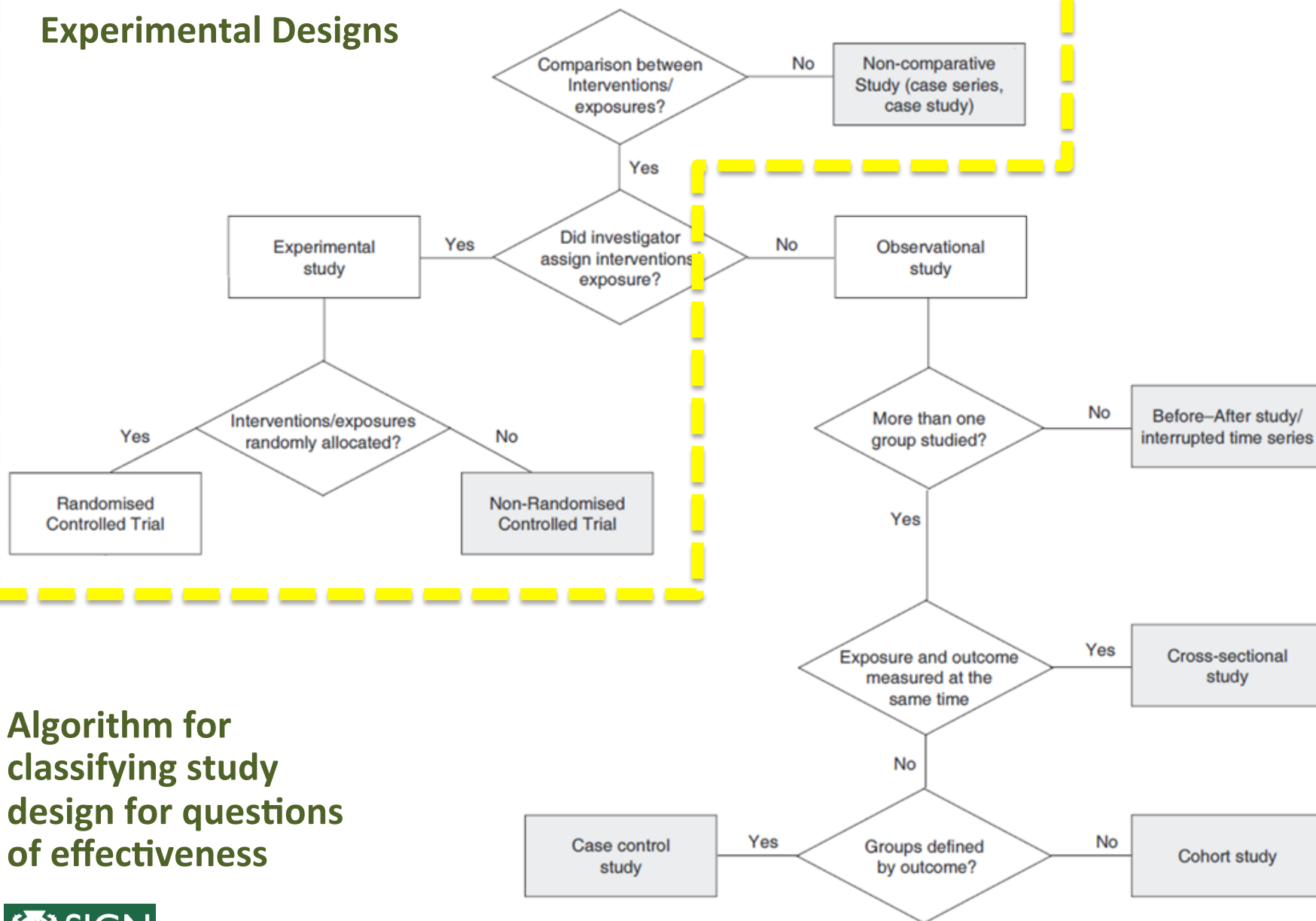
Step 2:
Propensity
score
adjustment

Independent variables:

Propensity score
Treatment

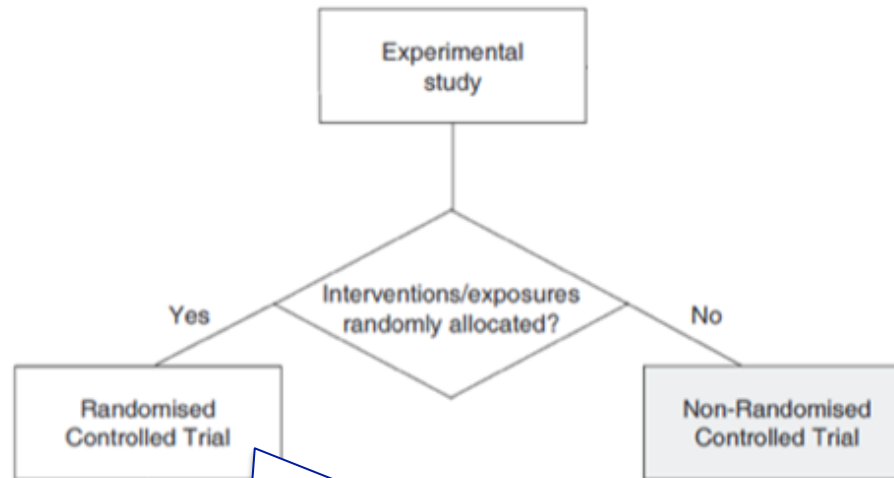
Dependent variable:
Clinical outcome

Experimental Designs



Algorithm for
classifying study
design for questions
of effectiveness

Randomization



Minimizes *allocation bias*,
balancing both known and
unknown prognostic factors, in the
assignment of treatments.

RANDOMIZZAZIONE

Assegnazione casuale dei pazienti al gruppo sperimentale o di controllo, al fine di assicurare che tutti i fattori prognostici - noti e sconosciuti - si distribuiscano omogeneamente nei due gruppi.

Tutti i requisiti della randomizzazione hanno lo scopo di assicurare che il **processo con cui vengono creati i due gruppi a confronto segua le leggi del caso**, e che **nessun fattore possa interferire** con la sua casualità.

Lachin, 2000

RANDOMIZATION COMPONENTS

Item	Descriptor
Sequence generation	Method used to generate the random allocation sequence, including details of any restriction (eg, blocking, stratification)
Allocation concealment	Method used to implement the random allocation sequence (eg, numbered containers or central telephone), clarifying whether the sequence was concealed until interventions were assigned
Implementation	Who generated the allocation sequence, who enrolled participants, and who assigned participants to their groups

Documentazione dell'assegnazione

456

Tabella A6 – Numeri casuali

I

PETER ARMITAGE

**STATISTICA
MEDICA**
METODI STATISTICI PER
LA RICERCA IN MEDICINA

FELTRINELLI

03	47	43	73	86	36	96	47	36	61	46	98	63	71	62	33	26	16	80	45	60	11	14	10	95
97	74	24	67	62	42	81	14	57	20	42	53	32	37	32	27	07	36	07	51	24	51	79	89	73
16	76	62	27	66	56	50	26	71	07	32	90	79	78	53	13	55	38	58	59	88	97	54	14	10
12	56	85	99	26	96	96	68	27	31	05	03	72	93	15	57	12	10	14	21	88	26	49	81	76
55	59	56	35	64	38	54	82	46	22	31	62	43	09	90	06	18	44	32	53	23	83	01	30	30
16	22	77	94	39	49	54	43	54	82	17	37	93	23	78	87	35	20	96	43	84	26	34	91	64
84	42	17	53	31	57	24	55	06	88	77	04	74	47	67	21	76	33	50	25	83	92	12	06	76
63	01	63	78	59	16	95	55	67	19	98	10	50	71	75	12	86	73	58	07	44	39	52	38	79
33	21	12	34	29	78	64	56	07	82	52	42	07	44	38	15	51	00	13	42	99	66	02	79	54
57	60	86	32	44	09	47	27	96	54	49	17	46	09	62	90	52	84	77	27	08	02	73	43	28
18	18	07	92	46	44	17	16	58	09	79	83	86	19	62	06	76	50	03	10	55	23	64	05	05
26	62	38	97	75	84	16	07	44	99	83	11	46	32	24	20	14	85	88	45	10	93	72	88	71
23	42	40	64	74	82	97	77	77	81	07	45	32	14	08	32	98	94	07	72	93	85	79	10	75
52	36	28	19	95	50	92	26	11	97	00	56	76	31	38	80	22	02	53	53	86	60	42	04	53
37	85	94	35	12	83	39	50	08	30	42	34	07	96	88	54	42	06	87	98	35	85	29	48	39
70	29	17	12	13	40	33	20	38	26	13	89	51	03	74	17	76	37	13	04	07	74	21	19	30
56	62	18	37	35	96	83	50	87	75	97	12	25	93	47	70	33	24	03	54	97	77	46	44	80
99	49	57	22	77	88	42	95	45	72	16	64	36	16	00	04	43	18	66	79	94	77	24	21	90
16	08	15	04	72	33	27	14	34	09	45	59	34	68	49	12	72	07	34	45	99	27	72	95	14
31	16	93	32	43	50	27	89	87	19	20	15	37	00	49	52	85	66	60	44	38	68	88	11	80
68	34	30	13	70	55	74	30	77	40	44	22	78	84	26	04	33	46	09	52	68	07	97	06	67
74	57	25	65	76	59	29	97	68	60	71	91	38	67	54	13	58	18	24	76	15	54	55	95	52
27	42	37	86	53	48	55	90	65	72	96	57	69	36	10	96	46	92	42	45	97	60	49	04	91

Statistica medica

- **A partire dalla prima riga, proseguire verso destra e quindi alle righe successive**
- **Numeri dispari: braccio A; numeri pari: braccio B.**

Generatore di numeri casuali compresi fra 0 e un numero n a scelta

	A	B	C	D	E	F	G	H	I	J	K
1	Generazione di numeri casuali compresi fra 1 e n										
2											
3	Numero complessivo di elementi nella popolazione:					200					
4	(premi F9 per ottenere una nuova serie di numeri casuali)										
5											
6	18	14	19	14	13	18	7	13	7	1	
7	7	1	16	13	12	16	14	10	5	17	
8	1	13	13	6	14	8	3	6	19	15	
9	10	20	2	11	10	7	14	18	16	15	
10	12	13	16	5	12	6	14	11	2	2	
11	17	4	7	3	16	17	7	13	8	2	
12	18	18	5	14	3	4	9	6	3	12	
13	8	14	11	17	4	4	19	18	10	8	
14	15	18	13	4	5	17	8	10	4	13	
15	7	16	16	14	4	15	15	16	3	15	
16	19	18	3	12	20	11	19	13	16	5	
17	15	1	15	11	6	17	11	15	10	9	
18	7	17	7	16	17	7	15	17	5	10	
19	11	12	2	16	7	18	16	19	10	19	
20	11	8	15	10	1	19	16	8	11	20	
21	5	15	15	1	17	10	20	14	14	5	
22	5	14	3	15	14	8	1	20	16	19	
23	1	13	16	16	17	4	19	5	6	5	
24	7	8	10	2	2	14	9	14	19	7	
25	18	6	7	1	2	14	20	18	12	19	

questo numero può essere modificato a piacere

ATTENZIONE: MANIPOLABILE E NON VERIFICABILE!

RANDOMIZZAZIONE A BLOCCHI

- La sequenza totale delle assegnazioni previste viene divisa in un certo numero di blocchi successivi.
- Il blocco rappresenta un gruppo di assegnazioni all'interno del quale vi è bilanciamento nel numero di pazienti assegnati ai due trattamenti, in modo da rispettare il rapporto di assegnazione previsto.
- I blocchi dovrebbero essere di dimensione variabile, in dipendenza dalle dimensioni campionarie e dal numero di strati)
 - es. blocco di 4: ABAB
 - es. blocco di 6: ABABAB
 - es. blocco di 8: ABABABAB

RANDOMIZZAZIONE A BLOCCHI

- ✓ Se $n = 4$ (dimensione del blocco)
- ✓ Se $x = 2$ (numero dei trattamenti)
- ✓ Se $A:B = 1:1$ (rapporto di assegnazione)

Quante (e quali) sono le possibili permutazioni?

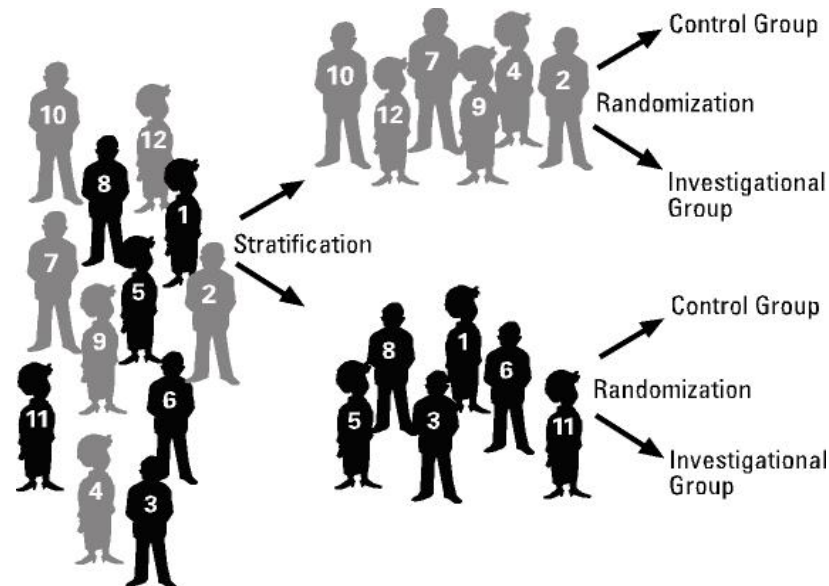
$$\frac{n!}{x! (n-x!)} = \frac{4!}{2! 2!} = 6$$


AABB
ABAB
BABA
ABBA
BAAB
BBAA

RANDOMIZZAZIONE STRATIFICATA

Allestimento di liste di randomizzazione separate per una o più caratteristiche pre-trattamento:

- ✓ misura atta a evitare sbilanciamenti fra i trattamenti a confronto per specifici fattori prognostici;
- ✓ possibili vantaggi di tipo gestionale e organizzativo (es. stratificazione per Centro);
- ✓ considerare solo fattori di stratificazione oggettivamente definibili.



RANDOMIZZAZIONE STRATIFICATA

Il numero di liste random che si viene a formare con la stratificazione è uguale al prodotto del numero degli strati di ogni fattore di stratificazione:

RAPID axSpA

- ✓ Site * 104
- ✓ mNY status * 2
- ✓ Prior TNF inhibition * 2

416 Liste di Randomizzazione
(325 pazienti)

RAPID PsA

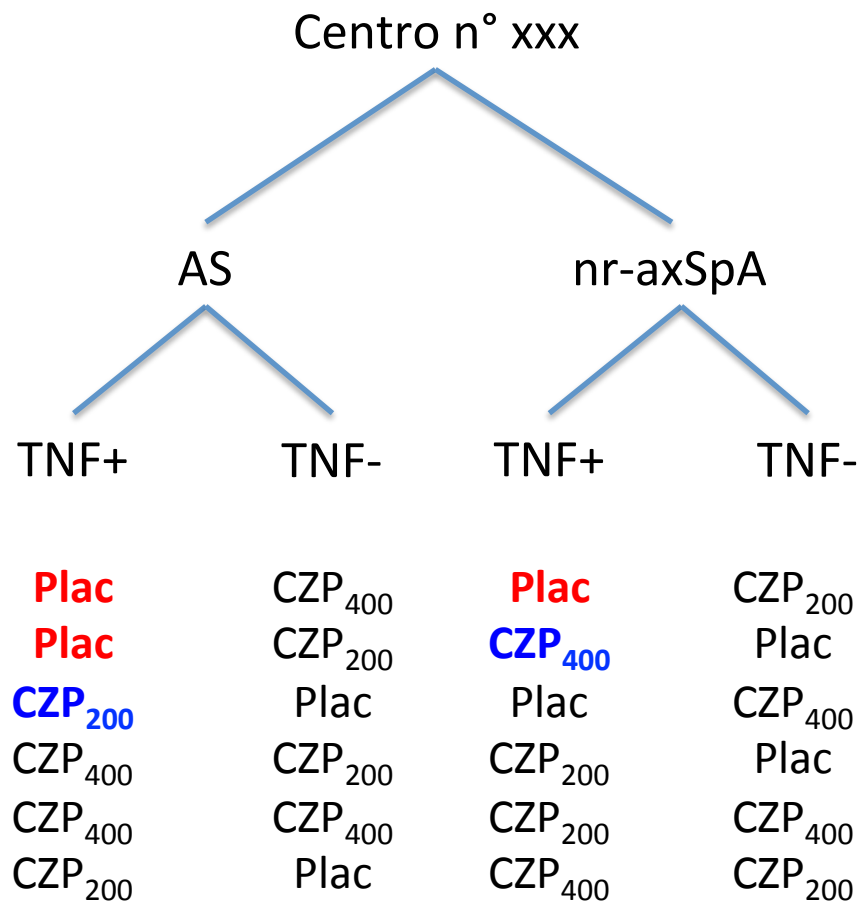
- ✓ Site * 92
- ✓ Prior TNF inhibitor use * 2

184 Liste di Randomizzazione
(409 pazienti)

Attenzione alla *overstratification* !

OVERSTRATIFICATION IN RAPID axSpA?

Possibile Scenario (potrebbe valere anche l'ipotesi opposta)



(alcune delle possibili permutazioni del blocco di 6)

Nel Centro n° xxx sono stati arruolati 5 pazienti con precedente esposizione a TNFi:

- ✓ di 3 pazienti con AS, 2 sono stati assegnati a Plac e 1 a CZP₂₀₀
- ✓ di 2 pazienti con nr-axSpA, 1 è stato assegnato a Plac e 1 a CZP₄₀₀

Gli strati TNF+ hanno un arruolamento non sufficiente a garantire il riempimento del blocco di 6 pazienti

MASCHERAMENTO

- Insieme delle procedure atte a prevenire distorsioni dovute al fatto che il Paziente, il Medico o il Valutatore sono a conoscenza del trattamento ricevuto dal Paziente:
 - *Paziente = singolo cieco*
 - *Paziente + Medico = doppio cieco*
 - *Paziente + Medico + Valutatore = triplo cieco*
- Necessità connessa agli obiettivi dello studio e al tipo di **endpoint** utilizzato

OPPORTUNITA' DEL MASCHERAMENTO

	<i>paziente</i>	<i>medico</i>	<i>valutatore</i>
Decesso (per ogni causa)	no	no	no
Decesso per causa specifica	no	no	si
Recidiva, progressione	no	no(?)	si
Risposta clinica	no (?)	no(?)	si
Risposta soggettiva	si	si	si
Dolore	si	si	si
Stato psichico	si	si	si

WARNING!

Allocation Concealment

≠

Blinding

- It prevents **selection bias** in intervention assignment by protecting the allocation sequence **before and until** assignment
 - It can always be successfully implemented regardless of the study topic
- It seeks to prevent **performance and detection bias** by protecting the sequence **after** assignment
 - Not always feasible – for example, in trials comparing surgical with medical interventions

Superiorità Vs Non-inferiorità

Si ritiene che il trattamento in esame
“A” abbia le potenzialità per
migliorare il trattamento standard
“B” almeno di una **quantità Δ**

**studio di
superiorità**

**A > B di una
quantità Δ
di interesse
clinico**

**studio di
non inferiorità**

**A < B non oltre
una quantità **M**
di rilevanza
clinica**

riorità

Vista la **migliore tollerabilità** del trattamento in esame “A”, si è disposti ad accettarne una eventuale minore efficacia rispetto al trattamento standard “B” purché questa non vada oltre un **margin** **M**

**studio di
superiorità**

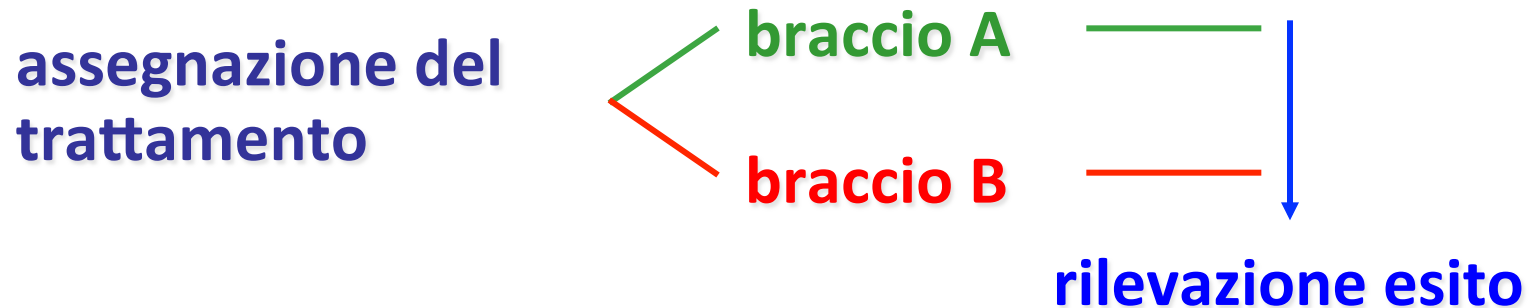
**A > B di una
quantità Δ
di interesse
clinico**

**studio di
non inferiorità**

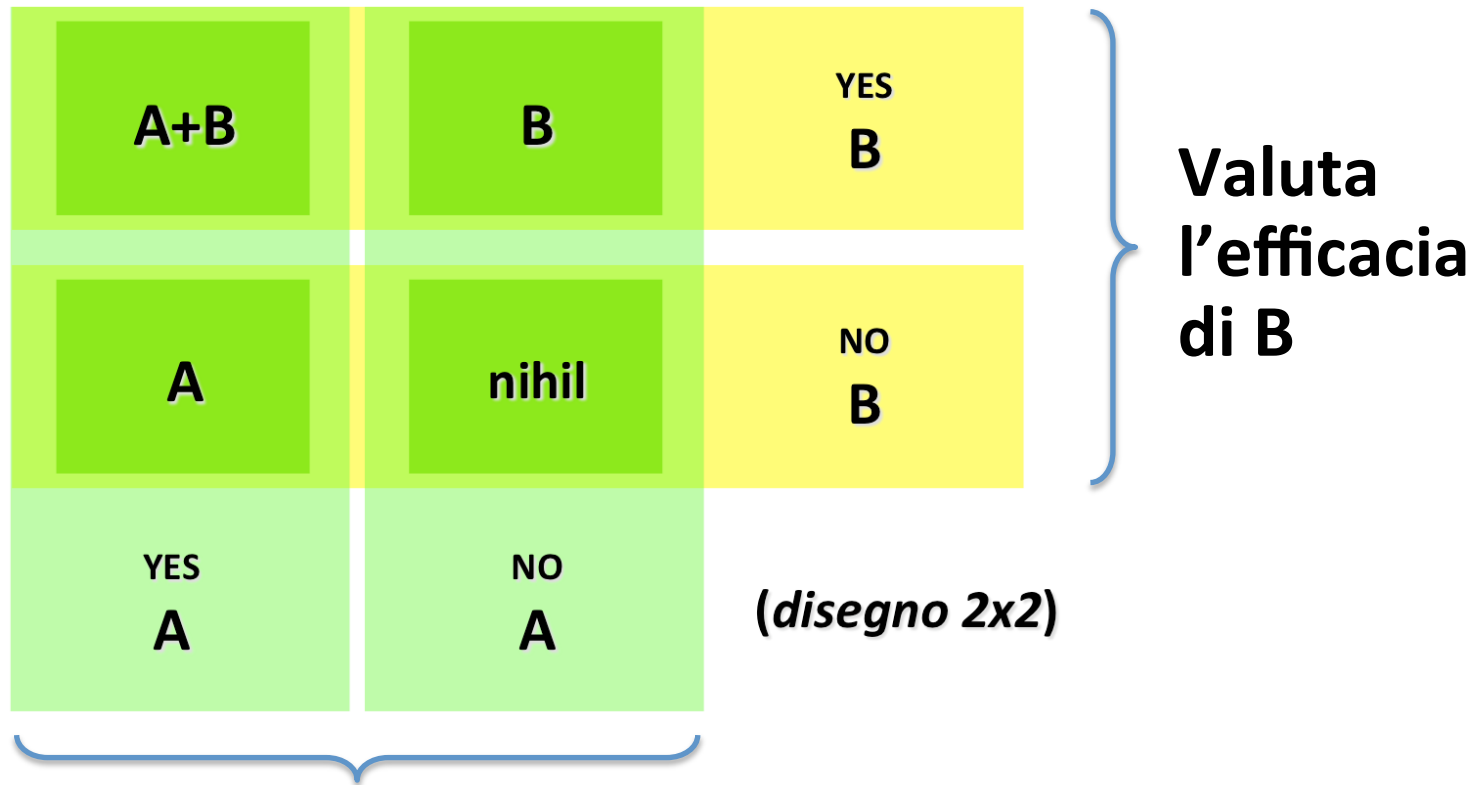
**A < B non oltre
una quantità **M**
di rilevanza
clinica**

DISEGNO A BRACCI PARALLELI

Assegnazione del paziente a un gruppo di trattamento, al quale si appartiene per l'intera durata dello studio:



DISEGNO FATTORIALE

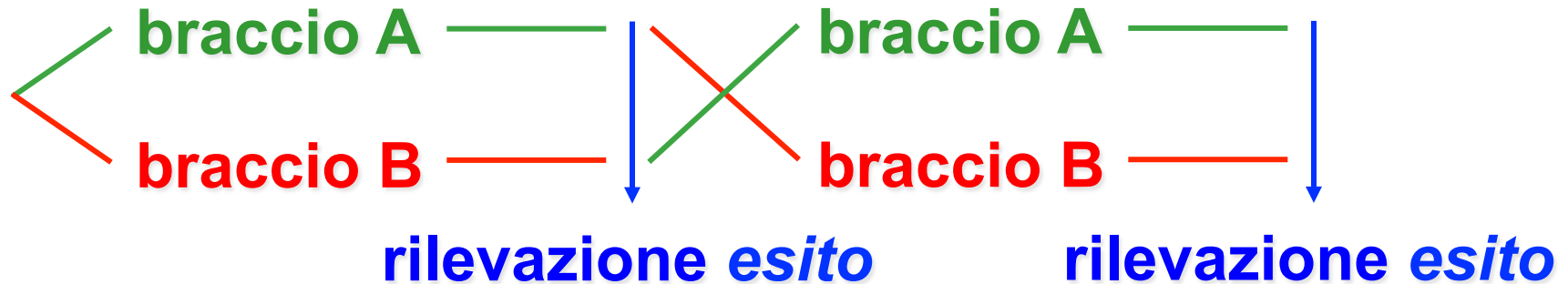


Valuta l'efficacia di A

Prerequisito: non interazione
tra gli effetti degli interventi
("righe Vs colonne")

DISEGNO CROSSOVER

Ciascun paziente riceve entrambi i trattamenti
oggetto di sperimentazione clinica (*within patient*
vs between patient):

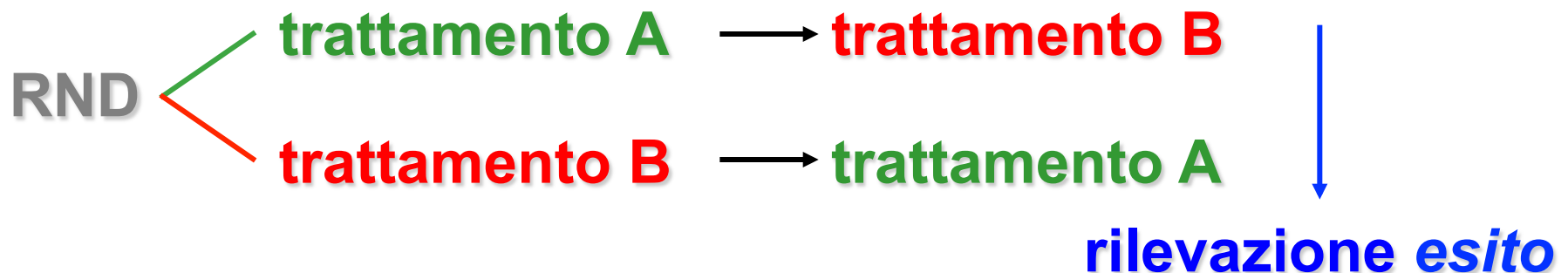


Viene “sottratta” dal confronto dei trattamenti
l’influenza delle caratteristiche del paziente, le quali
possono influire sulla misura di *outcome* →
dimensione campionaria minore rispetto a uno studio
a bracci paralleli.

ATTENZIONE A NON CONFONDERE UN DISEGNO CROSSOVER...



...CON UN DISEGNO A BRACCI PARALLELI DI TIPO SEQUENZIALE



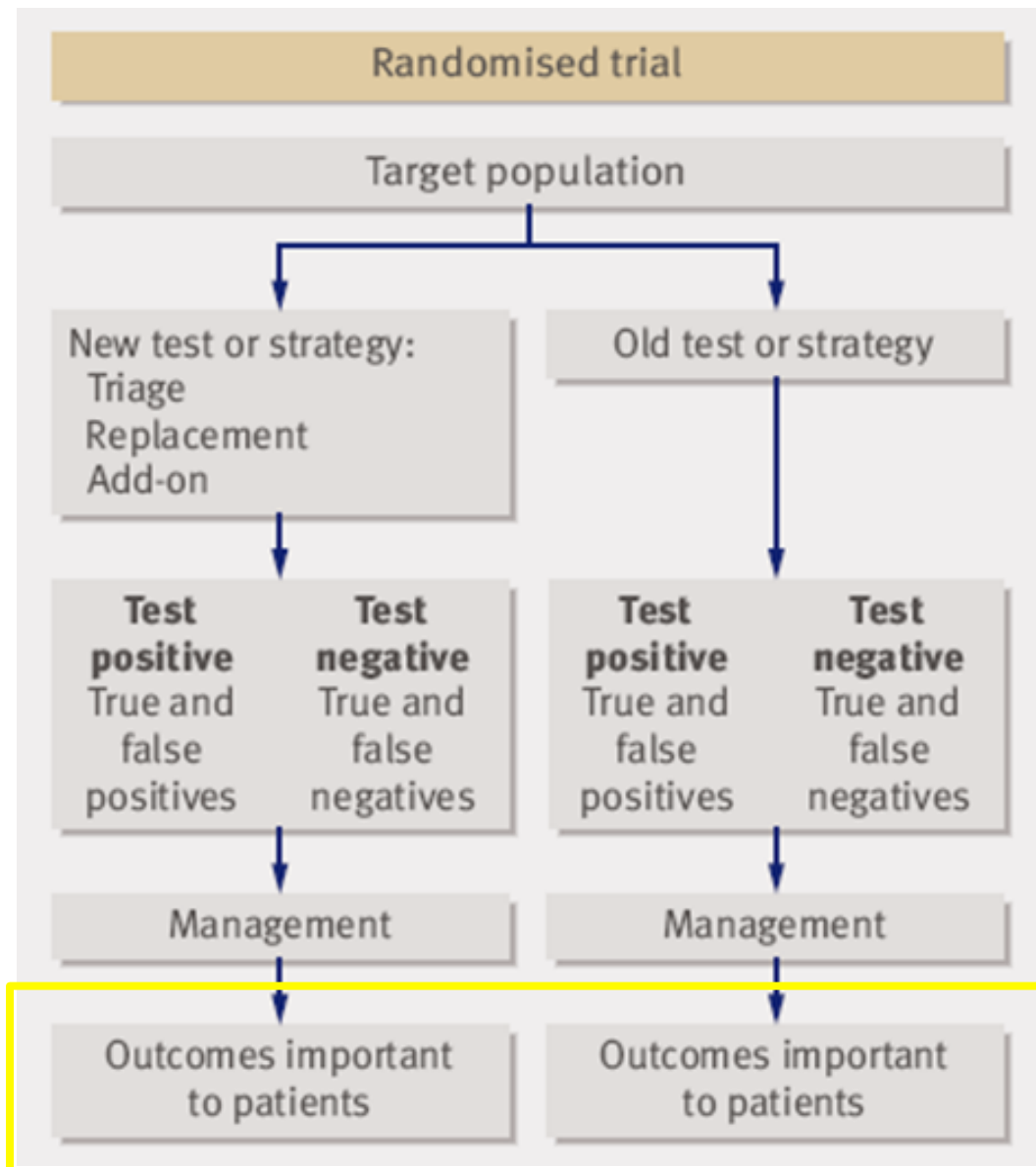
Choice of Control Group

- The selection of an appropriate control group is a critical decision which **impacts on the scientific validity and ethical acceptability** of a clinical investigation.
- The proper control group allows for discrimination between patient outcomes caused by the test treatment, and outcomes caused by other factors such as the natural progression of the disease, observer or patient expectations, or other treatments.



GRADE: grading quality of evidence and strength of recommendations for diagnostic tests and strategies

The best way to assess any diagnostic strategy—but in particular new strategies with putative superior accuracy—is a randomised controlled trial in which investigators randomise patients to experimental or control diagnostic approaches and measure mortality, morbidity, symptoms, and quality of life.





Metodologia della Ricerca Clinica

**Corso teorico pratico:
“Come si scrive un protocollo
di ricerca clinica”**

Coordinatore:
Dr.ssa Stefania Gori

*Evento ECM MODULO 4
(formazione avanzata)*



**NEGRAR
24-25 Maggio 2016**

Centro Formazione
Ospedale Sacro Cuore
Don Calabria

- ✓ Fattibilità e Rilevanza
- ✓ Obiettivi
- ✓ Disegno di studio
- ✓ **Endpoints**
- ✓ Criteri di Selezione
- ✓ Protocollo diagnostico
terapeutico
- ✓ Dimensionamento
Campionario

ENDPOINT

- **Variabile la cui misura viene utilizzata per valutare gli effetti del trattamento.**
- **Rilevato attraverso uno strumento di rilevazione (oggettivo? standardizzato?)**
- **Identificato in base a:**
 - **scopo specifico dello studio;**
 - **storia naturale della malattia;**
 - **effetti prevedibili dei trattamenti.**

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 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 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 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 “TEMPO A EVENTO”

- **Apparentemente assimilabile a una variabile di tipo quantitativo (intervallare).**
- **Ma il verificarsi o meno di un evento la rende assimilabile a una variabile di tipo qualitativo (nominale)**
- **In alcuni soggetti inoltre l'evento di interesse potrebbe non essersi ancora verificato al momento della analisi**
- **Tali risultati vengono quindi meglio rappresentati come stima della funzione di sopravvivenza:**
 - **probabilità di sopravvivere oltre un determinato tempo, misurato dalla data di inizio dell'osservazione.**

Effect measures for continuous outcomes

The **mean difference** (more correctly, 'difference in means') measures the absolute difference between the mean value in two groups in a clinical trial.

- ✓ It estimates the amount by which the experimental intervention changes the outcome on average compared with the control.

The **standardized mean difference** is used as a summary statistic in meta-analysis when the studies all assess the same outcome but measure it in a variety of ways.

- ✓ The standardized mean difference expresses the size of the intervention effect in each study relative to the variability observed in that study.

Selection of summary statistics for continuous data is principally determined by whether studies all report the outcome using the same scale (when the mean difference can be used) or using different scales (when the standardized mean difference has to be used)

Tutte le misure di rischio e beneficio derivano dalla tabella 2x2, da cui si ricavano incidenze* e odds

	Evento	NON Evento
Trattamento sperimentale	a	b
Gruppo di controllo	c	d

- $(a/a+b)$ = **incidenza** dell'evento nel gruppo sperimentale
- $(c/c+d)$ = **incidenza** dell'evento nel gruppo di controllo

* rischi

Rischio Relativo

$\text{incidenza}^{\text{sperim}}$

$\text{incidenza}^{\text{contr}}$

RRR (RRI) e RR sono misure complementari:

RR 0.40 → RRR 60%

RR 1.20 → RRI 20%

MISURE ASSOLUTE DI RISCHIO e BENEFICIO (1)

- **Riduzione Assoluta del Rischio (ARR)**
- **Incremento Assoluto del Rischio (ARI)**
 - Differenza aritmetica tra $\text{incid}^{\text{sperim}}$ e $\text{incid}^{\text{control}}$ (“Risk Difference”)
 - Riduzione / Incremento assoluto di evento nel gruppo dei pazienti
 - nei quali è presente il fattore di rischio
 - trattati con il trattamento sperimentale
 - **Tendono a sottostimare l’entità del rischio o del beneficio (in caso di bassa $\text{incid}^{\text{control}}$)**

MISURE ASSOLUTE DI RISCHIO e BENEFICIO (2)

- **Number Needed to Treat (NNT) = $1/ARR$**
 - Numero di pazienti da trattare per ottenere 1 beneficio terapeutico*
 - Al diminuire del NNT aumenta l'efficacia del trattamento
- **Number Needed to Harm (NNH) = $1/ARI$**
 - Numero di pazienti da trattare per osservare 1 effetto avverso del trattamento*
 - All'aumentare del NNH e aumenta la sicurezza del trattamento

** rispetto al braccio di controllo*

Tutte le misure di rischio e beneficio derivano dalla tabella 2x2, da cui si ricavano incidenze* e odds

	Evento	NON Evento
Trattamento sperimentale	a	b
Gruppo di controllo	c	d

- (a/b) = **odds** dell'evento nel gruppo sperimentale
- (c/d) = **odds** dell'evento nel gruppo di controllo

* rischi

Rischio Relativo

$$\frac{\text{incidenza}^{\text{sperim}}}{\text{incidenza}^{\text{contr}}}$$

Odds Ratio

$$\frac{\text{odds}^{\text{sperim}}}{\text{odds}^{\text{contr}}}$$

	Malati		
	Sì	No	
Test positivo	VP	FP	Positivi
Test negativo	FN	VN	Negativi
	Malati	Sani	Totale

Sensibilità = Capacità del test di individuare i pazienti con malattia; % persone malate positive al test

$$\text{Sensibilità} = \text{VP} / \text{Malati}$$

Specificità = Capacità del test di individuare i pazienti senza malattia; % delle persone sane negative al test

$$\text{Specificità} = \text{VN} / \text{Sani}$$

	Malati		
	Sì	No	
Test positivo	VP	FP	Positivi
Test negativo	FN	VN	Negativi
	Malati	Sani	Totale

Valore predittivo positivo=% di malati tra i pazienti positivi

$$VP+ = VP / \text{Positivi}$$

Valore predittivo negativo=% di sani tra i pazienti negativi

$$VP- = VN / \text{Negativi}$$

Quale valore predittivo è da preferire ?

Se il fine è individuare il maggior numero di malati, il test migliore è quello a **sensibilità maggiore**

Ciò comporta:

un miglior valore predittivo dell'esito negativo (un esito negativo indica quasi certamente un soggetto sano),

un minor valore predittivo dell'esito positivo (in molti casi, ad un esito positivo può corrispondere un soggetto sano)

Se il fine è individuare i soggetti sicuramente malati, il test migliore ha **specificità maggiore**

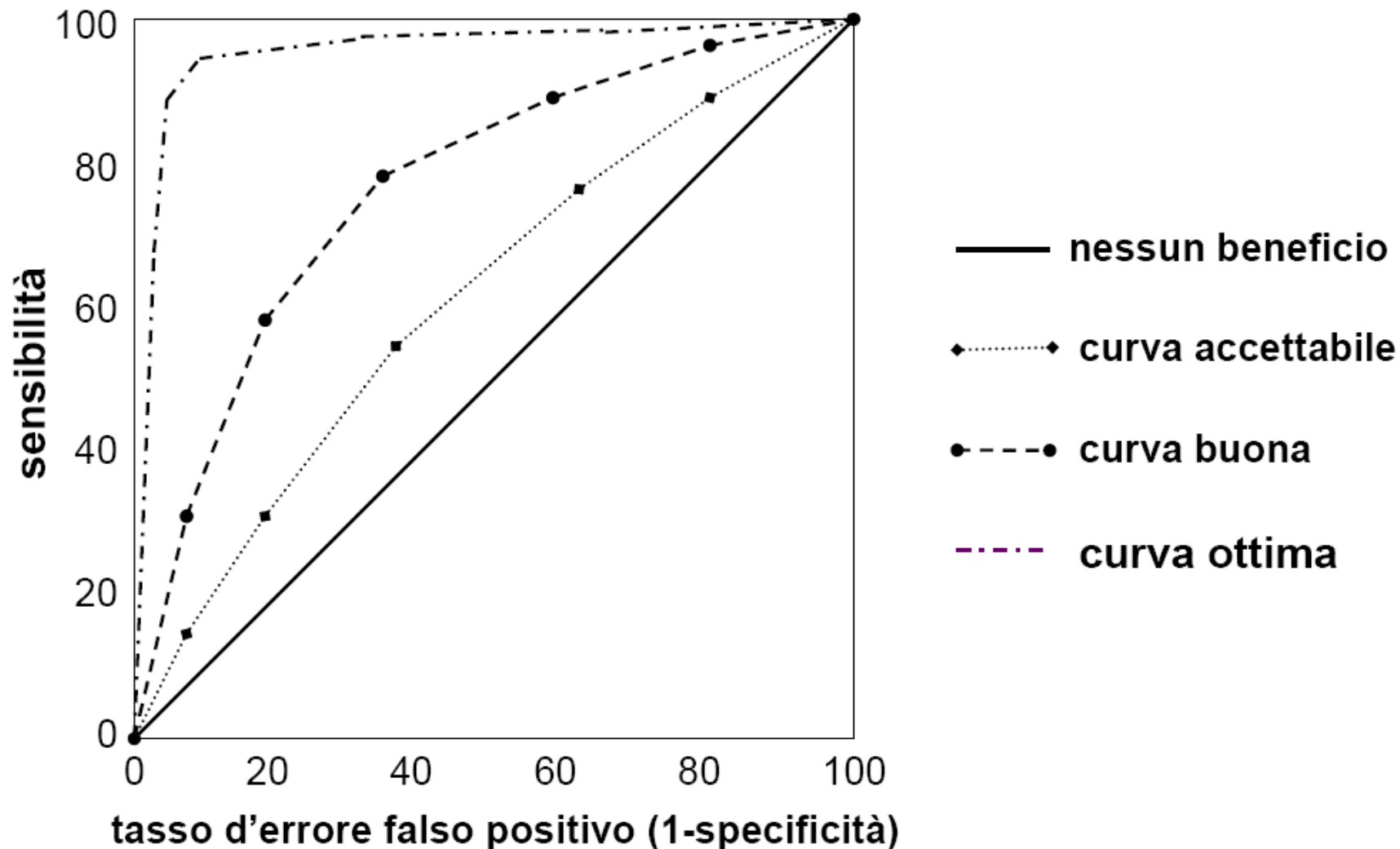
Essa comporta:

un miglior valore predittivo dell'esito positivo (un esito positivo indica quasi certamente un soggetto malato),

un minor valore predittivo dell'esito negativo (in molti casi, ad un esito negativo può corrispondere un soggetto malato)

Curve ROC

(receiver operating characteristics)



Indicatori riassuntivi di effetto di variabili tempo-a-evento

- Differenza tra stime della mediana di sopravvivenza (KM)
- Differenza media di sopravvivenza (*restricted means*)
- Differenza tra stime di sopravvivenza (KM) al tempo x
- Hazard Ratio (KM+Cox)



Metodologia della Ricerca Clinica
Corso teorico pratico:
**“Come si scrive un protocollo
di ricerca clinica”**

Coordinatore:
Dr.ssa Stefania Gori

*Evento ECM MODULO 4
(formazione avanzata)*



NEGRAR
24-25 Maggio 2016

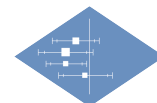
Centro Formazione
Ospedale Sacro Cuore
Don Calabria

- ✓ Fattibilità e Rilevanza
- ✓ Obiettivi
- ✓ Disegno di studio
- ✓ Endpoints
- ✓ **Criteri di Selezione**
- ✓ Protocollo diagnostico
terapeutico
- ✓ Dimensionamento
Campionario

CRITERI DI SELEZIONE DEI PAZIENTI

Caratteristiche che identificano, con precisione e senza possibilità di dubbio, i pazienti che possono essere inseriti nella sperimentazione:

- ✓ **devono comprendere informazioni disponibili con una certa facilità al momento dell' arruolamento;**
- ✓ **devono essere comprensibili anche a persone che non hanno partecipato alla stesura del protocollo.**



Criteri Restrittivi vs Inclusivi

- Criteri di Selezione **RESTRITTIVI**:
 - ✓ difficoltà a reclutare il numero di pazienti necessario e minore generalizzabilità dei risultati;
 - ✓ maggiore precisione delle valutazioni.
- Criteri di Selezione **INCLUSIVI**:
 - ✓ reclutamento più facile;
 - ✓ possibilità che l'efficacia della terapia venga oscurata da un'eccessiva variabilità della variabile di risposta.

STUDI “ESPLICATIVI” vs “APPLICATIVI”

*studi esplicativi
(explanatory)*

*studi applicativi
(pragmatic)*

outcome

decorso clinico

esito

selezione

restrittiva

inclusiva

accertamenti

**frequenti e
specifici**

**come da
pratica clinica**

informazioni

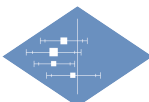
**molti dati
da raccogliere**

**pochi dati
da raccogliere**

centri

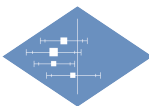
uno / pochi

molti



ELEMENTI DA CONSIDERARE NELLA SCELTA DI UNO STUDIO MONO Vs MULTICENTRICO

- **Patologia in esame**
- **Scopo specifico dello studio**
- **Numero di pazienti necessari**
- **Tempo ragionevole in cui completare l'arruolamento**
- **Tipo di trattamento**
- **Procedure diagnostiche necessarie**
- **Strategia di selezione dei pazienti**



Fattibilità vs Precisione

- **Ottenimento di una popolazione il più possibile omogenea:**
 - ✓ risultati della sperimentazione solo in parte influenzati dalla variabilità della popolazione in studio.
- **Riduzione delle restrizioni per l'inserimento dei pazienti in studio:**
 - ✓ poste le basi per la generalizzazione dei risultati.



Metodologia della Ricerca Clinica
Corso teorico pratico:
**“Come si scrive un protocollo
di ricerca clinica”**

Coordinatore:
Dr.ssa Stefania Gori

*Evento ECM MODULO 4
(formazione avanzata)*



NEGRAR
24-25 Maggio 2016

Centro Formazione
Ospedale Sacro Cuore
Don Calabria

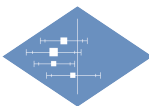
- ✓ Fattibilità e Rilevanza
- ✓ Obiettivi
- ✓ Disegno di studio
- ✓ Endpoints
- ✓ Criteri di Selezione
- ✓ **Protocollo diagnostico
terapeutico**
- ✓ Dimensionamento
Campionario

Protocollo Diagnostico

- **Procedure di screening**
- **Determinazioni basali**
- **Procedure di follow-up**

Protocollo Terapeutico

- **Dettaglio dei trattamenti/procedure previsti**
- **Eventuali modificazioni di dose/schedule**





Metodologia della Ricerca Clinica

**Corso teorico pratico:
“Come si scrive un protocollo
di ricerca clinica”**

Coordinatore:
Dr.ssa Stefania Gori

*Evento ECM MODULO 4
(formazione avanzata)*



**NEGRAR
24-25 Maggio 2016**

Centro Formazione
Ospedale Sacro Cuore
Don Calabria

- ✓ Fattibilità e Rilevanza
- ✓ Obiettivi
- ✓ Disegno di studio
- ✓ Endpoints
- ✓ Criteri di Selezione
- ✓ Protocollo diagnostico terapeutico
- ✓ **Dimensionamento
Campionario**

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)

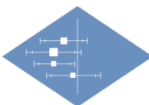
DIMENSIONAMENTO CAMPIONARIO

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%

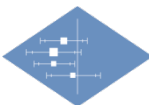
DIMENSIONAMENTO CAMPIONARIO PER VARIABILI DI TIPO QUANTITATIVO:

- 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.



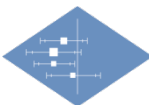
DIMENSIONAMENTO CAMPIONARIO PER VARIABILI DI TIPO QUALITATIVO:

- 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.



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.



Superiorità Vs Non-inferiorità

La dimensione di Δ dipende da:

- ✓ **tipo e fase della malattia in atto**
- ✓ **efficacia dei trattamenti disponibili**
- ✓ **caratteristiche del trattamento sperimentale**

studio di
superiorità

$A > B$ di una
quantità Δ
di interesse
clinico

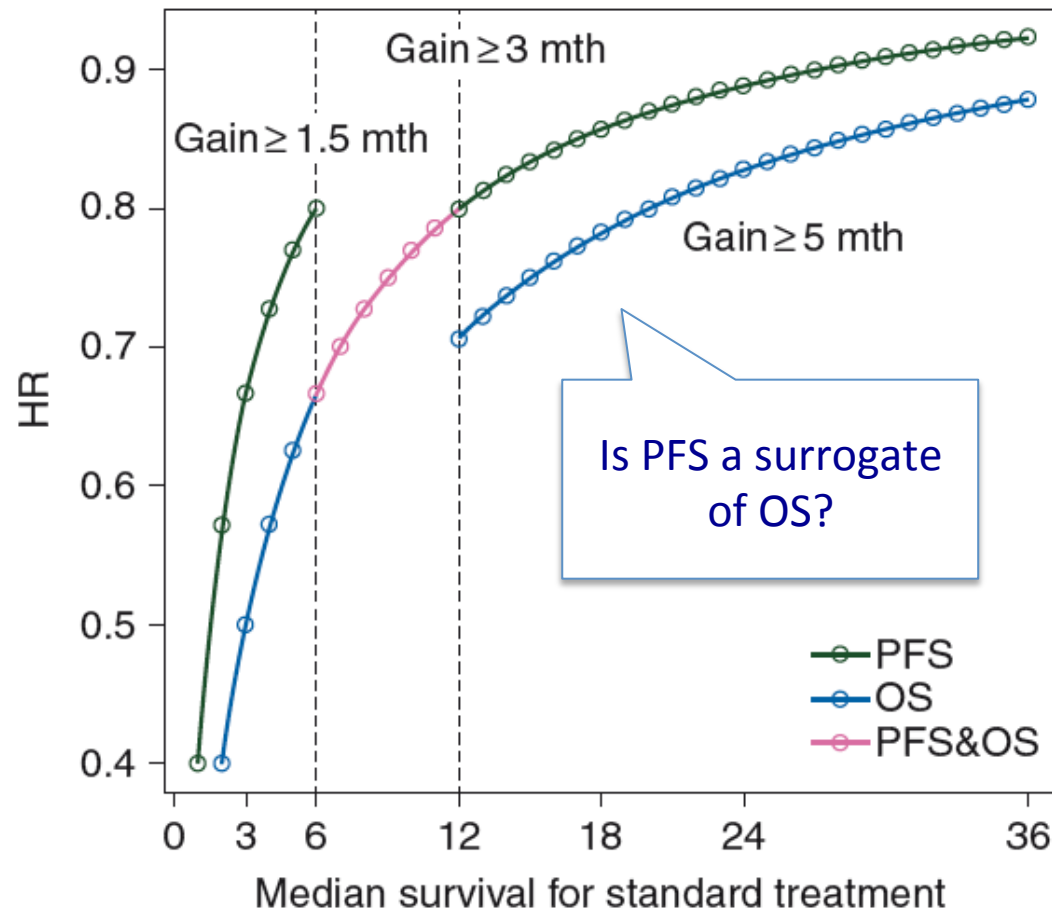
studio di
non inferiorità

$A < B$ non oltre
una quantità M
di rilevanza
clinica

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 Scale (ESMO-MCBS)

N. I. Cherny^{1*}, R. Sullivan², U. Dafni³, J. M. Kerst⁴, A. Sobrero⁵, C. Zielinski⁶, E. G. E. de Vries⁷ & M. J. Piccart^{8,9}

Annals of Oncology 26: 1547–1573, 2015

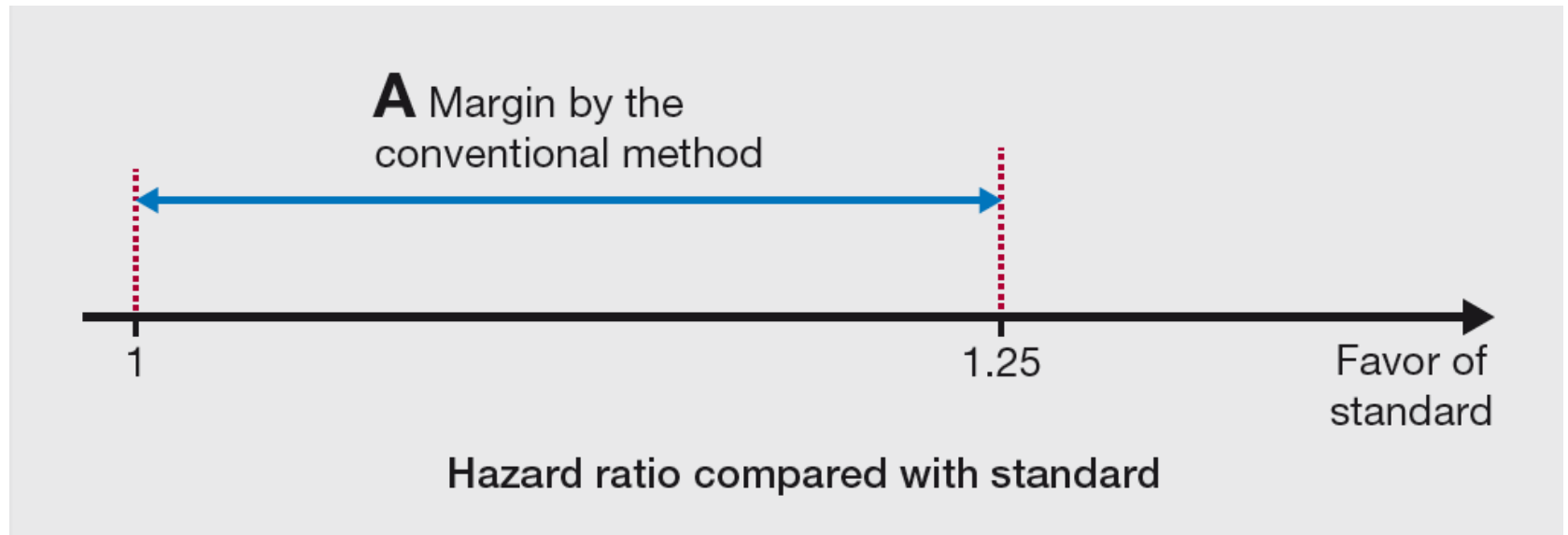


The correspondence between an HR value and the minimum absolute gain in months considered as beneficial according to the ESMO-MCBS by median survival (OS or PFS) for control

Statistical Issues and Recommendations for Noninferiority Trials in Oncology: A Systematic Review

Shiro Tanaka¹, Yousuke Kinjo², Yoshiki Kataoka², Kenichi Yoshimura¹, and Satoshi Teramukai¹
Clin Cancer Res; 18(7); 1837–47. ©2012 AACR.

Fixed Margin Method





European Medicines Agency
Pre-authorisation Evaluation of Medicines for Human Use

London, 27 July 2005

Doc. Ref. EMEA/CPMP/EWP/2158/99

**COMMITTEE FOR MEDICINAL PRODUCTS FOR HUMAN USE
(CHMP)**

GUIDELINE ON THE CHOICE OF THE NON-INFERIORITY MARGIN

In order to demonstrate non-inferiority, the recommended approach is to pre-specify a margin of non-inferiority in the protocol.

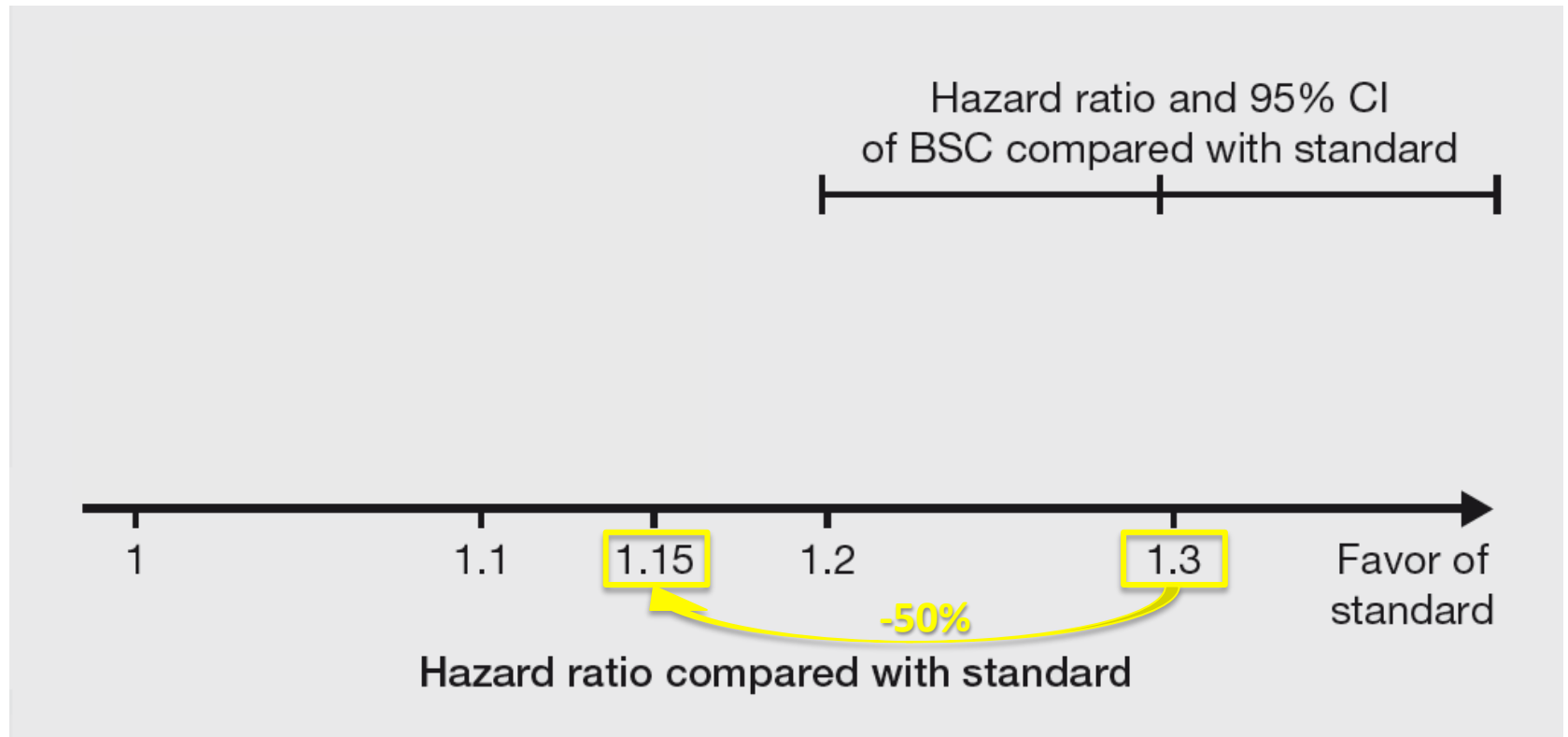
After study completion, a two-sided 95% confidence interval (or one-sided 97.5% interval) for the true difference between the two agents will be constructed. This interval should lie entirely on the positive side of the non-inferiority margin.

The choice of delta 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.

Statistical Issues and Recommendations for Noninferiority Trials in Oncology: A Systematic Review

Shiro Tanaka¹, Yousuke Kinjo², Yoshiki Kataoka², Kenichi Yoshimura¹, and Satoshi Teramukai¹
Clin Cancer Res; 18(7); 1837–47. ©2012 AACR.

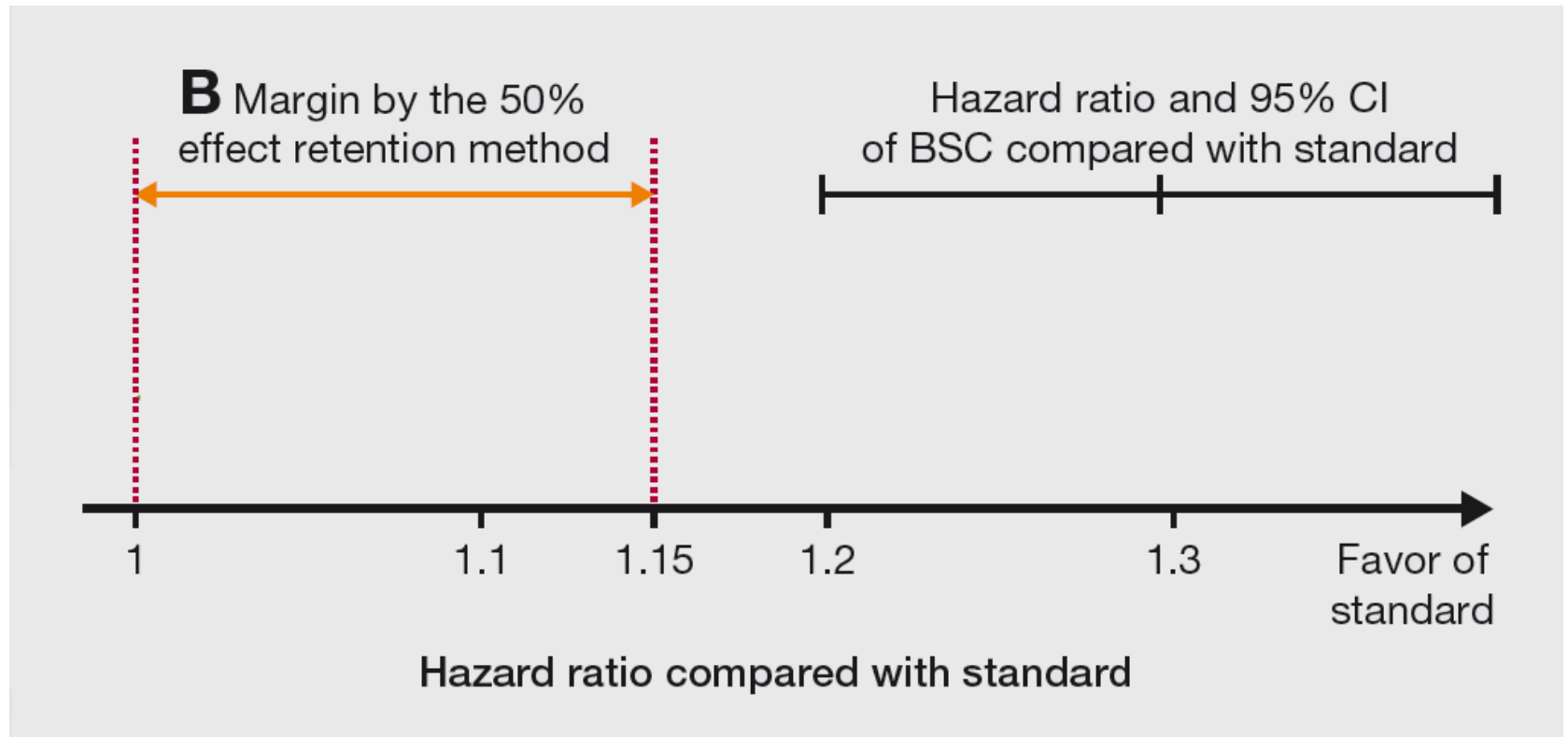
Percent Retention Method



Statistical Issues and Recommendations for Noninferiority Trials in Oncology: A Systematic Review

Shiro Tanaka¹, Yousuke Kinjo², Yoshiki Kataoka², Kenichi Yoshimura¹, and Satoshi Teramukai¹
Clin Cancer Res; 18(7); 1837–47. ©2012 AACR.

Percent Retention Method



Multiplicity

Multiple comparisons, **multiplicity** or multiple testing problem occurs when one considers a set of statistical inferences simultaneously or infers a subset of parameters selected based on the observed values

Probability of at least one false significant result

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

Several statistical techniques have been developed to prevent this from happening, allowing significance levels for single and multiple comparisons to be directly compared. These techniques generally require a higher significance threshold for individual comparisons, so as to compensate for the number of inferences being made.

Multiplicity is everywhere...

- In **subgroup** analyses (also when pre-specified)
- In **multiple endpoints** (this is why the primary endpoint must be pre-specified)
- In **interim** analyses
- In **reanalysis** of the same study (spanish intermittent ADT study)
- In **model building** (prognostic models, gene-signatures ...)

Multiplicity

Multiple comparisons, **multiplicity** or multiple testing problem occurs when one considers a set of statistical inferences simultaneously or infers a subset of parameters selected based on the observed values

Probability of at least one false significant result

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

Several statistical techniques have been developed to prevent this from happening, allowing significance levels for single and multiple comparisons to be directly compared. These techniques generally require a higher significance threshold for individual comparisons, so as to compensate for the number of inferences being made.

Multiplicity in randomised trials II: subgroup and interim analyses

Kenneth F Schulz, David A Grimes

Lancet 2005; 365: 1657–61

Number of planned interim analyses	Interim analysis	Pocock	Peto	O'Brien-Fleming
2	1	0.029	0.001	0.005
	2 (final)	0.029	0.05	0.048
3	1	0.022	0.001	0.0005
	2	0.022	0.001	0.014
	3 (final)	0.022	0.05	0.045
4	1	0.018	0.001	0.0001
	2	0.018	0.001	0.004
	3	0.018	0.001	0.019
	4 (final)	0.018	0.05	0.043
5	1	0.016	0.001	0.00001
	2	0.016	0.001	0.0013
	3	0.016	0.001	0.008
	4	0.016	0.001	0.023
	5 (final)	0.016	0.05	0.041

Overall $\alpha=0.05$.

Table 2: Interim stopping levels (p values) for different numbers of planned interim analyses by group sequential design^{14,15}

**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.

IMPUTAZIONE DEI DATI MANCANTI

Sostituzione dei valori mancanti o errati di un dato record con alternative coerenti e plausibili ottenute dai dati stessi, da fonti esterne all'indagine effettuata o dalla combinazione di entrambi, in conformità a regole e metodi prestabiliti.

L'obiettivo è di ridurre le distorsioni introdotte dalla presenza di dati mancanti, ma in taluni casi vengono amplificate le distorsioni esistenti.

Deal with missing data

- ▶ Use what you know about
 - ▶ Why data is missing
 - ▶ Distribution of missing data
- ▶ Decide on the best analysis strategy to yield the least biased estimates
 - ▶ Deletion Methods
 - ▶ Listwise deletion, pairwise deletion
 - ▶ Single Imputation Methods
 - ▶ Mean/mode substitution, dummy variable method, single regression
 - ▶ Model-Based Methods
 - ▶ Maximum Likelihood, Multiple imputation

The Prevention and Treatment of Missing Data in Clinical Trials

Roderick J. Little, Ph.D., Ralph D'Agostino, Ph.D., Michael L. Cohen, Ph.D., Kay Dickersin, Ph.D.,
Scott S. Emerson, M.D., Ph.D., John T. Farrar, M.D., Ph.D., Constantine Frangakis, Ph.D.,
Joseph W. Hogan, Sc.D., Geert Molenberghs, Ph.D., Susan A. Murphy, Ph.D.,
James D. Neaton, Ph.D., Andrea Rotnitzky, Ph.D., Daniel Scharfstein, Sc.D.,
Weichung J. Shih, Ph.D., Jay P. Siegel, M.D., and Hal Stern, Ph.D.

N ENGL J MED 367;14 NEJM.ORG OCTOBER 4, 2012

No matter what approach is taken, there is no way to adequately test the robustness of the assumptions about missing data required by the analysis. This need to rely on untestable assumptions regarding missing data reinforces the importance of preventing missing data in the first place. The key is to design and carry out the trial in a way that limits the problem of missing data.