



- ✓ **Presentazione delle “idee di ricerca”**
- ✓ **Scelta condivisa (?) delle idee da sviluppare in forma di sinossi di protocollo**
- ✓ **Formazione dei tavoli di lavoro (per affinità di interesse?)**

↓

Svolgimento dei lavori per step successivi (elementi base):

- **presentazione frontale**
- **lavoro di gruppo su template word**
- **presentazione in plenaria**

1



Titolo	F.I.N.E.R. Premessa e stato dell'arte
Fattibilità e Rilevanza	
Obiettivi	Obiettivo Primario Obiettivi Secondari
Disegno di studio	Tipologia di studio (osservazionale / sperimentale: disegni specifici) Procedure di assegnazione del trattamento (randomizzazione, blocking, concealment, mascheramento, stratificazione) Scelta dell'eventuale braccio di controllo
Endpoints	Variable di outcome Indicatore di effetto
Criteri di Selezione	Criteri di inclusione ed esclusione: restrittivi Vs inclusivi
Protocollo diagnostico-terapeutico	Protocollo diagnostico Protocollo terapeutico
Dimensionamento campionario	Indicatore di riferimento (braccio di controllo) per l'outcome primario Delta di interesse clinico (superiorità / non-inferiorità) Dimensione degli errori statistici Procedure di controllo della molteplicità Eventuali misure di imputazione dati mancanti

2



Presentazione delle “idee di ricerca” ...



Formazione dei tavoli di lavoro...

OK, siamo pronti per cominciare!

3



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

4

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.

5

FATTORI DA CONSIDERARE SULL'OPPORTUNITA' DI UNA SPERIMENTAZIONE CLINICA

- Gravità dell'affezione
- Efficacia (
- Tossicità disponibili native.
- Presumibi speriment

Il criterio FINER

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

6

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. 27

7

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

8

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**

9



10

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

11

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

12

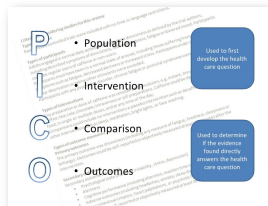


P. Bruzzi, qualche tempo fa...

13

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 (Participants, Interventions, Comparisons and Outcomes) helps to serve as a reminder of these.



14

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



15

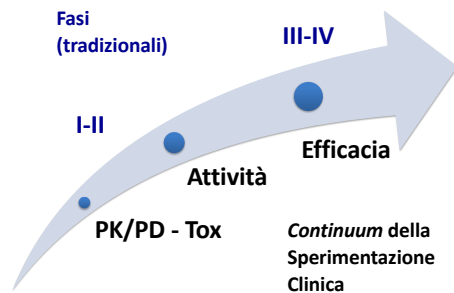


Hypotheses and Objectives

- 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

Joseph F. Collins, Sc.D.

16



17

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

18

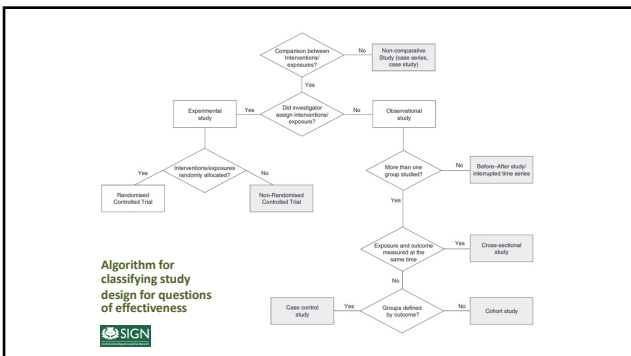


19

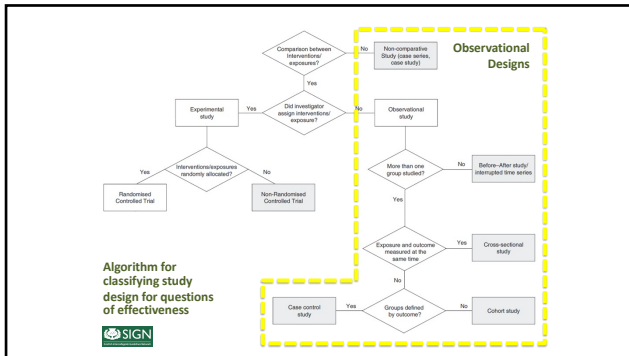


- ✓ Fattibilità e Rilevanza
- ✓ Obiettivi
- ✓ **Disegno di studio**
- ✓ Endpoints
- ✓ Criteri di Selezione
- ✓ Protocollo diagnostico terapeutico
- ✓ Dimensionamento Campionario

20



21



22

[illegible]

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 noninterventional, in the sense that the study protocol does not determine the precise features of any therapy given to the participants in the study.

23

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

24

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

25

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.

26

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

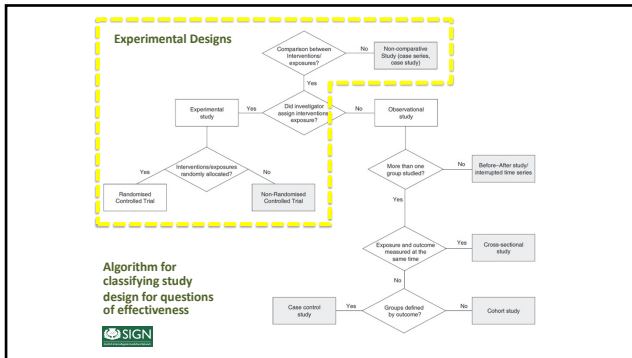
Cohort
Identify the incidence of a particular outcome over time

Non-comparative case series
Report outcomes of patients who received a specific intervention

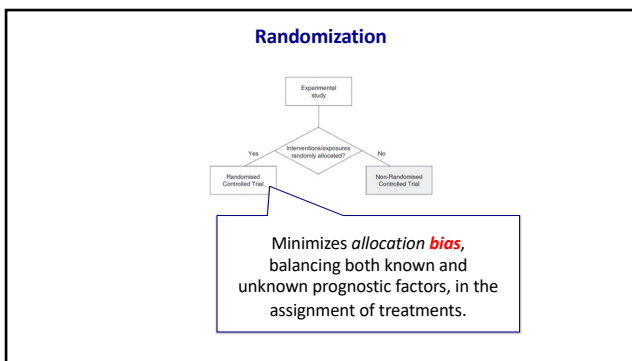
Comparative case series
Compare outcomes between patients who received different interventions

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

27



28



29

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

30

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

31

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

32

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?

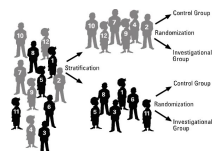
$$\frac{n!}{x! (n-x)!} = \frac{4!}{2! 2!} = 6 \quad \left\{ \begin{array}{l} A A B B \\ A B A B \\ B A B A \\ A B B A \\ B A A B \\ B B A A \end{array} \right.$$

33

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.



34

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

35

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

36

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	studio di non inferiorità
$A > B$ di una quantità Δ di interesse clinico	$A < B$ non oltre una quantità M di rilevanza clinica

37

Superiorità Vs Non-inferiorità

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 margine M

studio di superiorità	studio di non inferiorità
$A > B$ di una quantità Δ di interesse clinico	$A < B$ non oltre una quantità M di rilevanza clinica

38

DISEGNO A BRACCI PARALLELI

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

assegnazione del trattamento

$\swarrow$
 $\searrow$

braccio A

braccio B

—

—

rilevazione esito

39

DISEGNO FATTORIALE

A+B	B	YES B	} Valuta l'efficacia di B
A	nihil	NO B	
YES A	NO A		

(disegno 2x2)

Valuta l'efficacia di A

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

40

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.

41

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.

E-10 Choice of Control Group and Related Issues in Clinical Trials, May 2001

2

42



43

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

44

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.

45



What makes a good endpoint?

Characteristic	Meaning
Relevant	Clinically important/useful
Quantifiable	Measured on an appropriate scale
Valid	Measures the intended effect
Objective	Interpreted the same by all observers
Reliable	Same effect yields consistent measurements
Sensitive	Responds to small changes in the effect
Specific	Unaffected by extraneous influences
Precise	Small variability
Other	Tradition, cost, time, missing data

26-30 September 2014, Madrid, Spain
S. Piantadosi (2005)
esmo.org

46


U.S. Food and Drug Administration
Protecting and Promoting Public Health
www.fda.gov

Regular Approval Basis – "Clinical Benefit"

- **Longer life**
- **Better life**
- **Established Surrogate for one of above**

47

"Surrogate" endpoints

- **Issue:**
 - **Quicker, less expensive, less clinically relevant endpoint or**
 - **More expensive, clinically definitive endpoint?**
- **Hesitate to use the term "surrogate"**
- **Has a specific technical definition**

Mark Conaway, June 2006

48

Validation of Surrogate Endpoints

Property of a Valid Surrogate

Effect of the Intervention on the Clinical Endpoint

is reliably predicted by the

Effect of the Intervention on the Surrogate Endpoint

D. Sargent, ASCO 2011

49



50



STUDI CLINICI: METODOLOGIA
COME SCRIVERE UN PROTOCOLLO DI RICERCA CLINICA
NEGRAR DI VALPOLICELLA
22-23 SETTEMBRE 2021
Corsi formativi
BIOCS Ospedale San Carlo Don Calabria

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

51

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.



52

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.

53

STUDI "ESPLICATIVI" vs "APPLICATIVI"

	<i>studi esplicativi (explanatory)</i>	<i>studi applicativi (pragmatic)</i>
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



54

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



55

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.

56



57



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

58

Protocollo Diagnostico

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

Protocollo Terapeutico

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

59



60



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

61

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

62

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

63

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

64

When Are “Positive” Clinical Trials in Oncology Truly Positive?

Alberto Ocana, Ian F. Tannock
J Natl Cancer Inst 2011;103:16-20

We would define a positive trial as one in which the predefined value of *delta* represents a clinically important difference in an endpoint that directly reflects benefit (mainly OS or quality of life) to patients and for which the results provide a best estimate of the *difference that exceeds that predefined value of delta*. Consistent with a recent commentary suggesting the need to increase the value of *delta* in future clinical trials, we provide an estimate of *delta* that would be generally accepted as representing a minimum clinically important difference in the primary endpoint: approximately 3 months increase in median OS for patients with advanced metastatic solid tumors (usually corresponding to an *hazard ratio of approximately 0.75*).

65

ERRORI STATISTICI

- ✓ **Errore di 1° tipo** (errore *alpha*)
 - 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 *beta*)
 - 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)

66

DIMENSIONAMENTO CAMPIONARIO

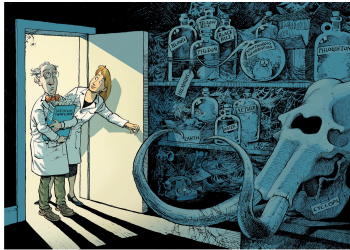
Punti caratterizzanti:

- ✓ 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%

67

P-value is a measure of the probability that an observed difference could have occurred just by random chance.

- Don't base your conclusions solely on whether an association or effect was found to be "statistically significant".
- Don't believe that an association or effect is absent just because it was not statistically significant.
- Don't conclude anything about scientific or practical importance based on statistical significance (or lack thereof).



Retire statistical significance

Valentin Amrhein, Sander Greenland, Blake McShane and more than 800 signatories call for an end to hyped claims and the dismissal of possibly crucial effects.

21 MARCH 2019 | VOL 567 | NATURE | 305

68

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.

69

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)

70

The Minimal Clinical Interesting Difference (M.C.I.D.)

- it's easily understood by clinicians as a key concept in the interpretability of PRO scores;
- will inform judgments about the successfulness of an intervention;
- an individual patient achieving the score equal or greater than the MCID might be considered a beneficiary of the intervention, what would lead to the definition of a **responder**.

71

SINOSI DEI TEST STATISTICI (1)

	variabili quantitative	
	test parametrici	test non parametrici
due gruppi di trattamento composti da differenti individui	t-test per campioni indipendenti	test di <i>Mann-Whitney</i>
tre o più gruppi di trattamento composti da differenti individui	analisi della varianza (covarianza)	test di <i>Kruskal-Wallis</i>
prima e dopo un singolo trattamento negli stessi individui	t-test per dati appaiati	test di <i>Wilcoxon</i> per dati appaiati
trattamenti multipli negli stessi individui	analisi della varianza per misure ripetute	test di <i>Friedman</i>

72

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.

73

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.

74

Risks, Rates and Odds

- **Risk** (proportion of persons with disease = *cumulative incidence*)
 - *Risk Ratio* = ratio of 2 cumulative incidence estimates = *Relative Risk*
- **Rate** (based on events per person-time = *incidence rate*)
 - *Rate Ratio* = ratio of 2 incidence rates = *Relative Rate*
- **Odds** (the number of events divided by the number of non events)
 - *Odds Ratio* = ratio of 2 odds

75

Incidenza Cumulativa

Probabilità (rischio) di sviluppare la malattia in uno specifico periodo di tempo t

- assume follow-up completo
- è una proporzione perciò può assumere valori da 0 ad 1
- deve riferirsi ad uno specifico periodo di tempo

$$IC = \frac{\text{N° di persone che ammalano tra il } t_0 \text{ e } t_1}{\text{N° di persone non malate all'inizio del periodo } t_0}$$

Es. 5 si ammalano / 10 inizialmente non malati = 0.5

76

Risks, Rates and Odds

- **Risk** (proportion of persons with disease = *cumulative incidence*)
 - **Risk Ratio** = ratio of 2 cumulative incidence estimates = *Relative Risk*
- **Rate** (based on events per person-time = *incidence rate*)
 - **Rate Ratio** = ratio of 2 incidence rates = *Relative Rate*
- **Odds** (the number of events divided by the number of non events)
 - **Odds Ratio** = ratio of 2 odds

77

Risks, Rates and Odds

- **Risk** (proportion of persons with disease = *cumulative incidence*)
 - **Risk Ratio** = ratio of 2 cumulative incidence estimates = *Relative Risk*
- **Rate** (based on events per person-time = *incidence rate*)
 - **Rate Ratio** = ratio of 2 incidence rates = *Relative Rate*
- **Odds** (the number of events divided by the number of non events)
 - **Odds Ratio** = ratio of 2 odds

78

Incidence Rate

Incidence rate or person-time rate:

- is a measure of incidence that incorporates time directly into the denominator;
- describes how quickly disease occurs in a population

$$\frac{\text{Number of new cases of disease or injury during specified period}}{\text{Time each person was observed, totaled for all persons}}$$

79

Risks, Rates and Odds

- **Risk** (proportion of persons with disease = *cumulative incidence*)
 - **Risk Ratio** = ratio of 2 cumulative incidence estimates = *Relative Risk*
- **Rate** (based on events per person-time = *incidence rate*)
 - **Rate Ratio** = ratio of 2 incidence rates = *Relative Rate*
- **Odds** (the number of events divided by the number of non events)
 - **Odds Ratio** = ratio of 2 odds

80

Risks, Rates and Odds

- **Risk** (proportion of persons with disease = *cumulative incidence*)
 - **Risk Ratio** = ratio of 2 cumulative incidence estimates = *Relative Risk*
- **Rate** (based on events per person-time = *incidence rate*)
 - **Rate Ratio** = ratio of 2 incidence rates = *Relative Rate*
- **Odds** (the number of events divided by the number of non events)
 - **Odds Ratio** = ratio of 2 odds

81

Risks, Rates and Odds

- Risk (proportion of persons with disease = cumulative incidence)
- Odds Ratios are used to compare the occurrence of the outcome of interest (e.g. disease or unfavourable event), given exposure to the variable of interest (e.g. health characteristic, or intervention). Most commonly used in case-control studies
- Odds (the number of events divided by the number of non events)
 - Odds Ratio = ratio of 2 odds

82

Risks, Rates and Odds

a	b	N1
28	37	65
c	d	N2
7	84	91
N+	N-	N
35	121	156

(28/65)
(7/91)

Risk Ratio

5.6000

CI Start

2.6068

CI End

12.0299

(28/65) - (7/91)

Risk Differ...

0.3538

CI Start

0.2216

CI End

0.4861

(28/37)
(7/84)

Odds Ratio

9.0811

CI Start

3.6405

CI End

22.6524

83

SINOSI DEI TEST STATISTICI (2)

	variabili qualitative	variabili del tipo "tempo a evento"
due gruppi di trattamento composti da differenti individui	test χ^2 per tabella di contingenza	log-rank test
tre o più gruppi di trattamento composti da differenti individui	test χ^2 per tabella di contingenza	log-rank test
prima e dopo un singolo trattamento negli stessi individui	test di McNemar	-
trattamenti multipli negli stessi individui	-	-

84

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.

85

	Malati		
	Si	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

Sensibilità=VP/Malati

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

Specificità=VN/Sani

86

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

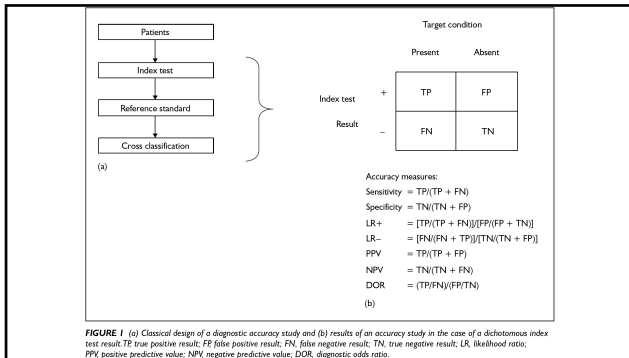
Valore predittivo positivo=% di malati tra i pazienti positivi

VP+=VP/Positivi

Valore predittivo negativo=% di sani tra i pazienti negativi

VP-=VN/Negativi

87



88

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)

89

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.

90

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.

91

STIMA DELLA FUNZIONE DI SOPRAVVIVENZA (Kaplan-Meier)

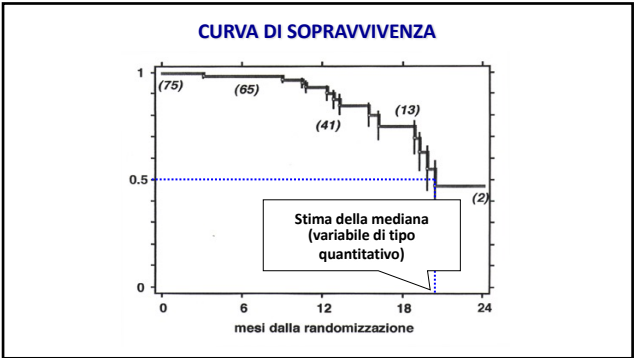
- Stima della probabilità di sopravvivere in corrispondenza di ciascuno dei tempi in cui si verifica almeno un evento.
- Si assume che l'esperienza dei pazienti “usciti vivi” o “persi di vista” sia simile a quella dei pazienti rimasti sotto osservazione sino al momento in cui si è verificato l'evento.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

92

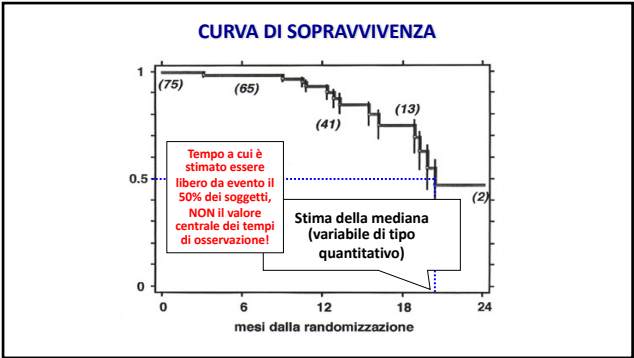
Tempi di risposta	Tempi troncati	N° soggetti esposti a rischio	N° eventi terminali	Rischio istantaneo di "morte"	Probabilità cumulativa di sopravvivere
t_0	t^*	n_i	d_i	$l(t_i)$	P_i
9		20	1	$1/20 = .050$	$(1 - 1/20) \times 1 = .950$
13		19	1	$1/19 = .053$	$(1 - 1/19) \times .950 = .896$
20		18	1	$1/18 = .055$	$(1 - 1/18) \times .896 = .850$
26		17	1	$1/17 = .059$	$(1 - 1/17) \times .850 = .799$
27		16	1	$1/16 = .062$	$(1 - 1/16) \times .799 = .750$
28		15	1	$1/15 = .067$	$(1 - 1/15) \times .750 = .700$
30		14	1	$1/14 = .071$	$(1 - 1/14) \times .700 = .650$
32		13	2	$2/13 = .154$	$(1 - 2/13) \times .650 = .550$
75		11	1	$1/11 = .091$	$(1 - 1/11) \times .550 = .500$
79		10	1	$1/10 = .100$	$(1 - 1/10) \times .500 = .450$
91		9	1	$1/9 = .111$	$(1 - 1/9) \times .450 = .400$
130	177*	8	0	$0/8 = .0$	$(1 - 0/8) \times .400 = .400$
541		7	1	$1/7 = .143$	$(1 - 1/7) \times .400 = .343$
1129		6	1	$1/6 = .167$	$(1 - 1/6) \times .343 = .286$
1585	1499*	5	1	$1/5 = .200$	$(1 - 1/5) \times .286 = .229$
		4	0	$0/4 = .0$	$(1 - 0/4) \times .229 = .229$
		3	1	$1/3 = .333$	$(1 - 1/3) \times .229 = .153$

TABELLA 10.
Calcolo secondo Kaplan e Meier della curva di sopravvivenza dei rene trapiantati nei pazienti di tabella 5.

93



94



95

Analisi della sopravvivenza in sperimentazioni cliniche controllate e nelle osservazioni prospettive

Tempi di risposta	Tempi troncati*	N° soggetti esposti a rischio	N° eventi terminali	Rischio istantaneo di "morire"	Probabilità cumulativa di sopravvivere l_{ij}
t_{ij}	t^*	n_i	d_i	$\lambda(t_{ij})$	P_i
9		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	$1/18 = .056$	$(1 - 1/18) \times .8996 = .8501$
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		9	1	$1/9 = .111$	$(1 - 1/9) \times .4501 = .4001$
	177*	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.

96

Indicatori riassuntivi di effetto di variabili tempo-a-evento

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

97

SINOSSI DEI TEST STATISTICI (3)

	variabili qualitative	variabili del tipo "tempo a evento"
due gruppi di trattamento composti da differenti individui	test χ^2 per tabella di contingenza	log-rank test
tre o più gruppi di trattamento composti da differenti individui	test χ^2 per tabella di contingenza	log-rank test
prima e dopo un singolo trattamento negli stessi individui	test di McNemar	-
trattamenti multipli negli stessi individui	-	-

98

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.

99

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

100

Statistical Issues and Recommendations for Noninferiority Trials in Oncology: A Systematic Review
Shiro Tanaka¹, Yousuke Kirjo², Yoshiki Kataoka², Kanichi Yoshimura¹, and Satoshi Teramukai¹
Clin Cancer Res. 18(7): 1837-47. ©2012 AACR.

Fixed Margin Method

101

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

London, 27 July 2005
Doc. Ref. EMEA/CPMP/EPW/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.

102

Good Enough: A Primer on the Analysis and Interpretation of Noninferiority Trials

Sanjay Kaul, MD, and George A. Diamond, MD
Ann Intern Med. 2006;145:62-69

Active-control noninferiority trials are being performed with increasing frequency, especially in cardiovascular and oncologic applications when placebo-controlled trials are considered unethical.

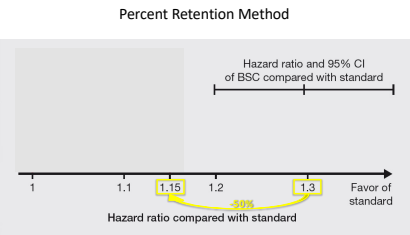
Because of this uncertainty, the noninferiority margin is typically defined in terms of some fraction (f) of the standard treatment effect to be preserved (9, 11).

In the context of oncologic and thrombolytic trials, when mortality is evaluated, the U.S. Food and Drug Administration has suggested an f value of 0.5 (9, 11, 16).

103

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

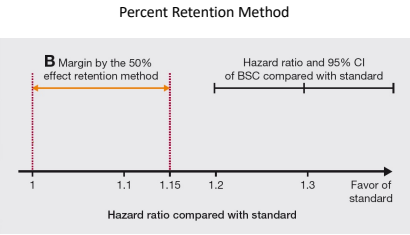
Shiro Tanaka¹, Yousuke Kirjo², Yoshiki Kataoka², Kenichi Yoshimura¹, and Satoshi Teramukai¹
Clin Cancer Res. 18(7): 1837-47. ©2012 AACR.



104

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

Shiro Tanaka¹, Yousuke Kirjo², Yoshiki Kataoka², Kenichi Yoshimura¹, and Satoshi Teramukai¹
Clin Cancer Res. 18(7): 1837-47. ©2012 AACR.



105

Studi di non-inferiorità: Analisi ITT Vs analisi PP



The European Agency for the Evaluation of Medicinal Products
Evaluation of Medicines for Human Use
London, 27 July 2000 CPM/3/WP/482/99

IV.2.3 Choice of analysis set

In a superiority trial the full analysis set, based on the ITT (intention-to-treat) principle, is the analysis set of choice, with appropriate support provided by the PP (per protocol) analysis set. In a non-inferiority trial the full analysis set and the PP analysis set have equal importance and their use should lead to similar conclusions for a robust interpretation.

Guidance for Industry Non-Inferiority Clinical Trials

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

Although an "as-treated" analysis is therefore often suggested as the primary analysis for NI studies, there are also significant concerns with the possibility of informative censoring in an as-treated analysis. It is therefore important to conduct both ITT and as-treated analyses in NI studies. Differences in results using the two analyses will need close examination.

106

Multiplicity



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

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

modified from Schulz & Grimes; Lancet 2005; 365: 1657-61

107

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

Misure per contenere la Multiplicity

- **Alpha-split, alpha-spending**
ripartire (tra confronto primario e sotto-gruppi di interesse, oppure tra diversi endpoints) la probabilità convenzionale (5%) di risultato falsamente positivo

- **Hierarchical approach**

- stabilito preventivamente l'ordine in cui saranno effettuati i test statistici;
- la sequenza di test si interromperà al primo risultato statisticamente NON significativo.

La non indipendenza e la non simultaneità dei test permetterà di contenere il rischio di risultato falsamente positivo entro il convenzionale 5%.

108

Multiplicity in randomised trials II: subgroup and interim analyses
Altman DG, Schulz DF, Moher D, et al. 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	0.005
	2	0.022	0.001	0.014
	3 (final)	0.022	0.05	0.045
4	1	0.018	0.001	0.001
	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	0.0001
	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^{a,b}

109

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.

110

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

Misure per contenere la Multiplicity

- Alpha-split, alpha-spending**
ripartire (tra confronto primario e sotto-gruppi di interesse, oppure tra diversi endpoints) la probabilità convenzionale (5%) di risultato falsamente positivo
- Hierarchical approach**
- stabilito preventivamente l'ordine in cui saranno effettuati i test statistici;
- la sequenza di test si interromperà al primo risultato statisticamente NON significativo.

La non indipendenza e la non simultaneità dei test permetterà di contenere il rischio di risultato falsamente positivo entro il convenzionale 5%.

111

Statistical Analysis Plan
(to implement a 2-sided alpha level of 0.05)

- Overall Survival (Interim)
- Overall Survival (final)
- Objective Response Rate
- Progression Free Survival

alpha spending (bracketed next to Overall Survival outcomes)

hierarchical procedure (bracketed next to Objective Response Rate and Progression Free Survival)

112

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.

113

114