



Introduzione alla valutazione economica per il mini-HTA

Le evidenze di efficacia e le sintesi, per la valutazione economica

Chiara Alberti - Anna Fratucello

UVEF – Servizio di Farmacia AOUI Verona

**20-21-22
27-28 aprile 2010**

**Corso itinerante di formazione Regionale in tema di HTA
2[^] giornata – sessione I**

STUDI ESPLICATIVI

STUDI DESCRITTIVI

Esaminano l'eziologia,
le cause, l'efficacia,
utilizzando la strategia
dei confronti

Documentano e comunicano esperienze:
idee, programmi, trattamenti, eventi
inusuali, osservazioni.
Danno il via alla ricerca esplicativa

Studi sperimentali

Studi osservazionali

Esempi:
Case report
Descrizione di serie di casi
Studi su popolazioni

Valutano l'efficacia di
interventi terapeutici,
educativi, amministrativi


Ricercano cause,
eziologie, predittori

Esempi:
Sperimentazioni
cliniche controllate

Studi caso-controllo
Studi di coorte
Studi trasversali

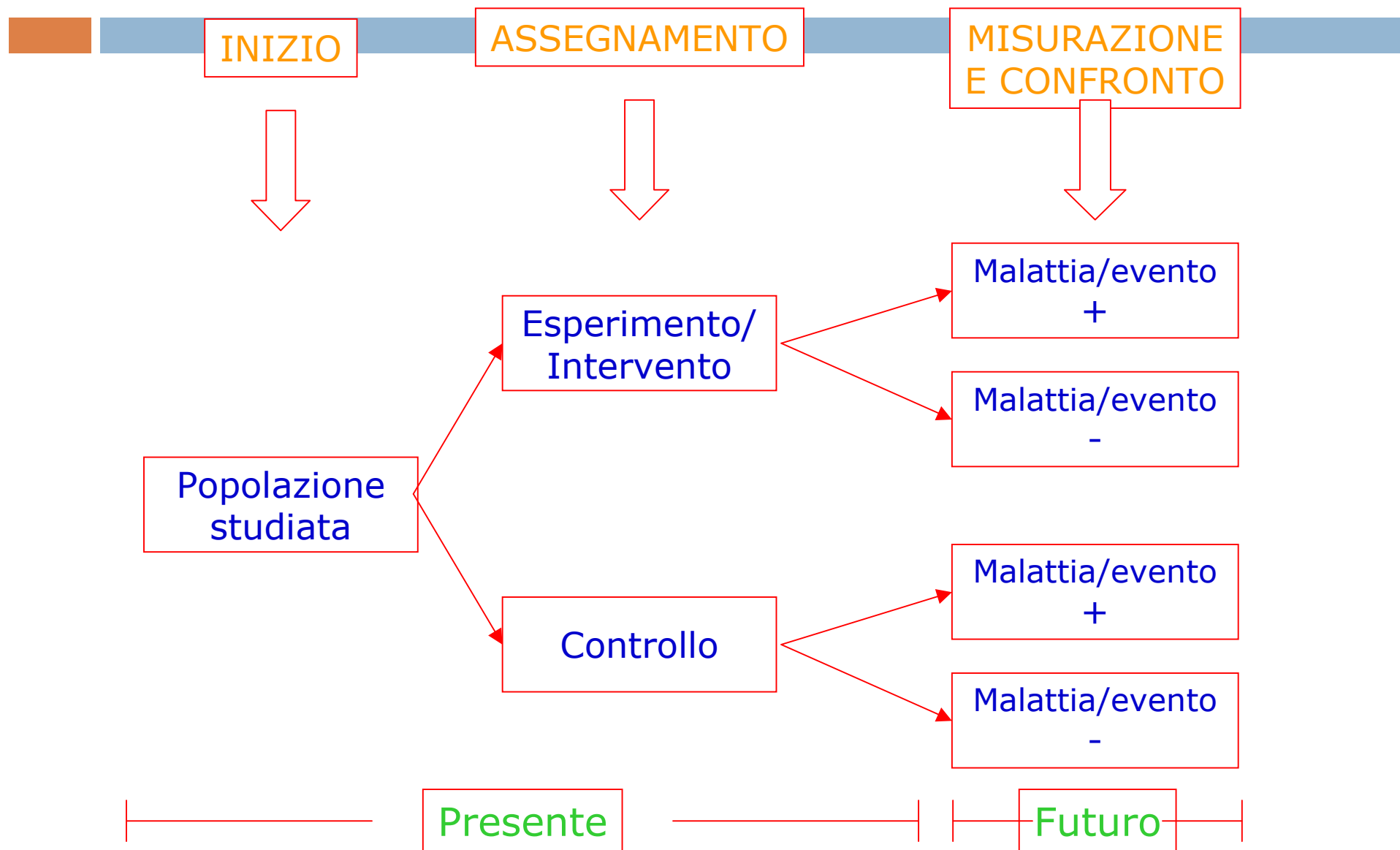
La forza delle evidenze





**Il gold-standard per gli studi di efficacia
è rappresentato dagli
studi clinici controllati
randomizzati in doppio cieco (RCT)**

RCT: Disegno



L'importanza della definizione degli end-point

A seconda delle fasi di sviluppo di una tecnologia (sia essa un farmaco, dispositivo medico o un altro tipo di intervento), gli end-point devono essere in grado di misurare:

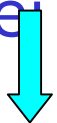
- Tossicità/tollerabilità
- Effetti a breve termine
- Effetti a lungo termine
- Combinazione degli stessi

End points surrogati

“Un end point surrogato è una misura di laboratorio o un segno obiettivo usato come sostituto di un end point clinicamente significativo, che misura direttamente le misure soggettive, le funzioni, o la sopravvivenza di un paziente”

Food and
Drug Administration

Un effetto su un end point surrogato non è di per sé di nessun valore per il paziente. Lo diventa solo se esso causa o predice un miglioramento nei risultati (es. minori infarti, ictus, o decessi).



colesterolo



Eventi cardiovascolari

END POINT INTERMEDI E COMPOSITI

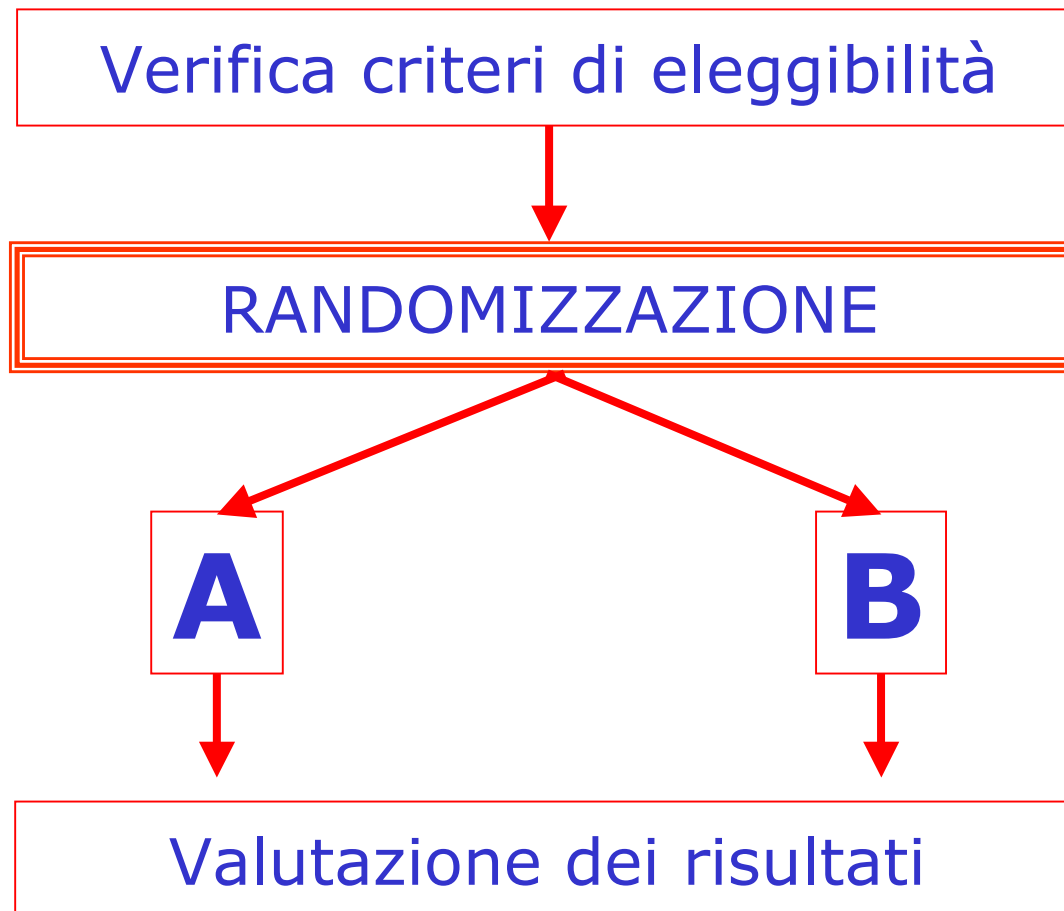
Gli end point surrogati vengono spesso confusi con gli **end point intermedi**. Questi ultimi rappresentano reali end point clinici (misure di sintomi o funzioni come ad es. sintomi di ipoglicemia, frequenza di episodi di angina, test da sforzo), ma non costituiscono l'end point finale. Un miglioramento in un end point intermedio è importante per il paziente, anche se esso non determina una riduzione della mortalità o della morbilità.

Tuttavia, anche l'uso di end point intermedi può portare a conclusioni errate.

Gli **end-point compositi** sono costituiti da molteplici misure di efficacia di differente rilevanza clinica che vengono valutate in modo combinato. L'uso degli end-point compositi riduce la dimensione campionaria necessaria per ottenere la potenza statistica adeguata. Rischiano, però, di produrre risultati fuorvianti: non è raro, infatti, che i dati legati all'obiettivo multifattoriale divergano dall'end-point più significativo ossia la mortalità da ogni causa.

I diversi tipi di disegno

Disegno comparativo semplice a 2 bracci



Molina J.M. et al. Lancet 2008; 372: 646-55.

**883 pz con HIV drugs-naive trattati
Con 2 NRTI**

RANDOMIZZAZIONE

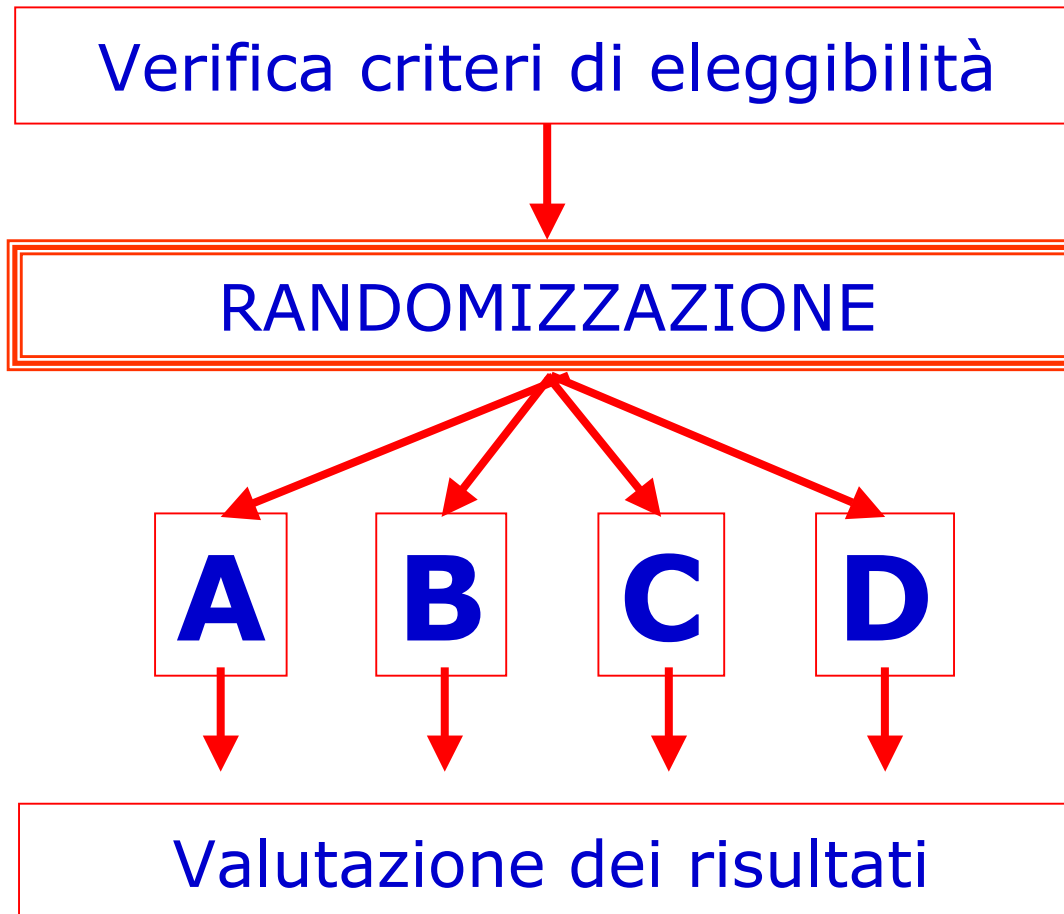
**ATAZANAVIR+
RITONAVIR**

**LOPINAVIR+
RITONAVIR**

ridurre la carica di HIV RNA al di sotto
delle 50 copie/mL alla settimana 48

I diversi tipi di disegno

Disegno comparativo a più bracci



Griffiths C et al. NEJM 2009; 362: 118-128

903 pz con psoriasi mod-grave precedentemente trattati

RANDOMIZZAZIONE

Ustekinumab 45 mg

Ustekinumab 90 mg

Etanercept

PASI 75 alla 12 settimana

PASI: Psoriasis Area and Severity Index La scala di punteggi è la seguente: 0% = nessun miglioramento rispetto al basale; <25% = miglioramento minimo; 25%-49% = miglioramento moderato; 50%-74% = buon miglioramento; 75%-99% = miglioramento eccellente; 100% = scomparsa completa.

Monk B.J. et al, Journal of Clinical Oncology 2009

472 pz ≥ 18 anni, carcinoma recidivante della cervice
uterina, avanzato (stadio IVB)

RANDOMIZZAZIONE

PACLITAXEL+
CISPLATINO

VINORELBINA+
CISPLATINO

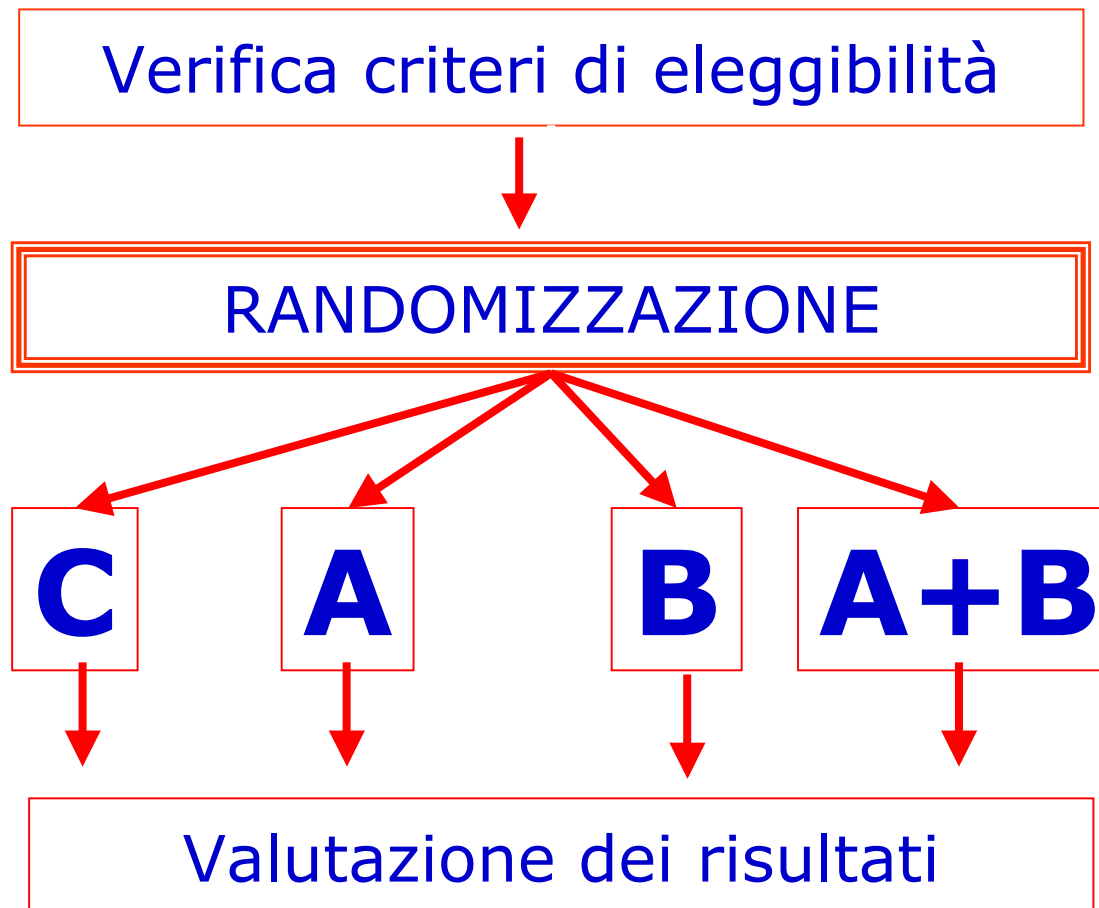
GEMCITABINA+
CISPLATINO

TOPOTECAN+
CISPLATINO

SOPRAVVIVENZA COMPLESSIVA

I diversi tipi di disegno

Disegno fattoriale



I diversi tipi di disegno

Disegno fattoriale

Vantaggi

È uno studio molto efficiente, poiché permette di ottenere due risposte con un singolo studio delle stesse dimensioni occorrenti per saggiare una sola delle ipotesi.

Consente di testare **l'effetto congiunto** dei due trattamenti in esame.

Svantaggi

Può essere utilizzato solo quando:

- Si è realmente interessati a testare l'efficacia di più trattamenti.
- I trattamenti messi a confronto possono essere somministrati contemporaneamente, non interferiscono cioè fra di loro e la tossicità additiva non è proibitiva.
- I meccanismi d'azione dei trattamenti sono diversi

La conduzione dello studio e l'analisi dei dati sono più complessi.

J Cassidy et al, J Clin Oncol 2008

Randomized Phase III Study of Capecitabine Plus Oxaliplatin Compared With Fluorouracil/Folinic Acid Plus Oxaliplatin As First-Line Therapy for Metastatic Colorectal Cancer

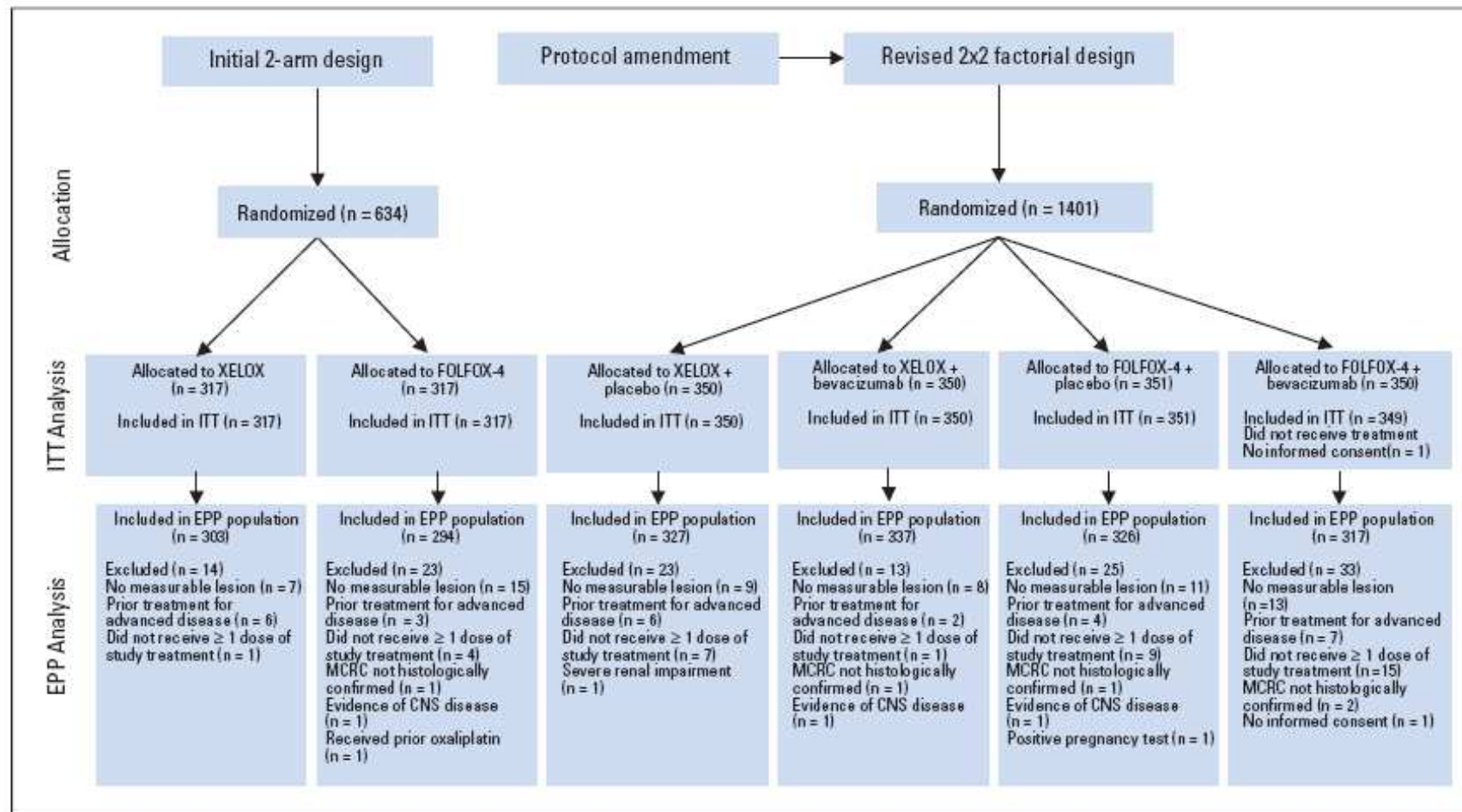
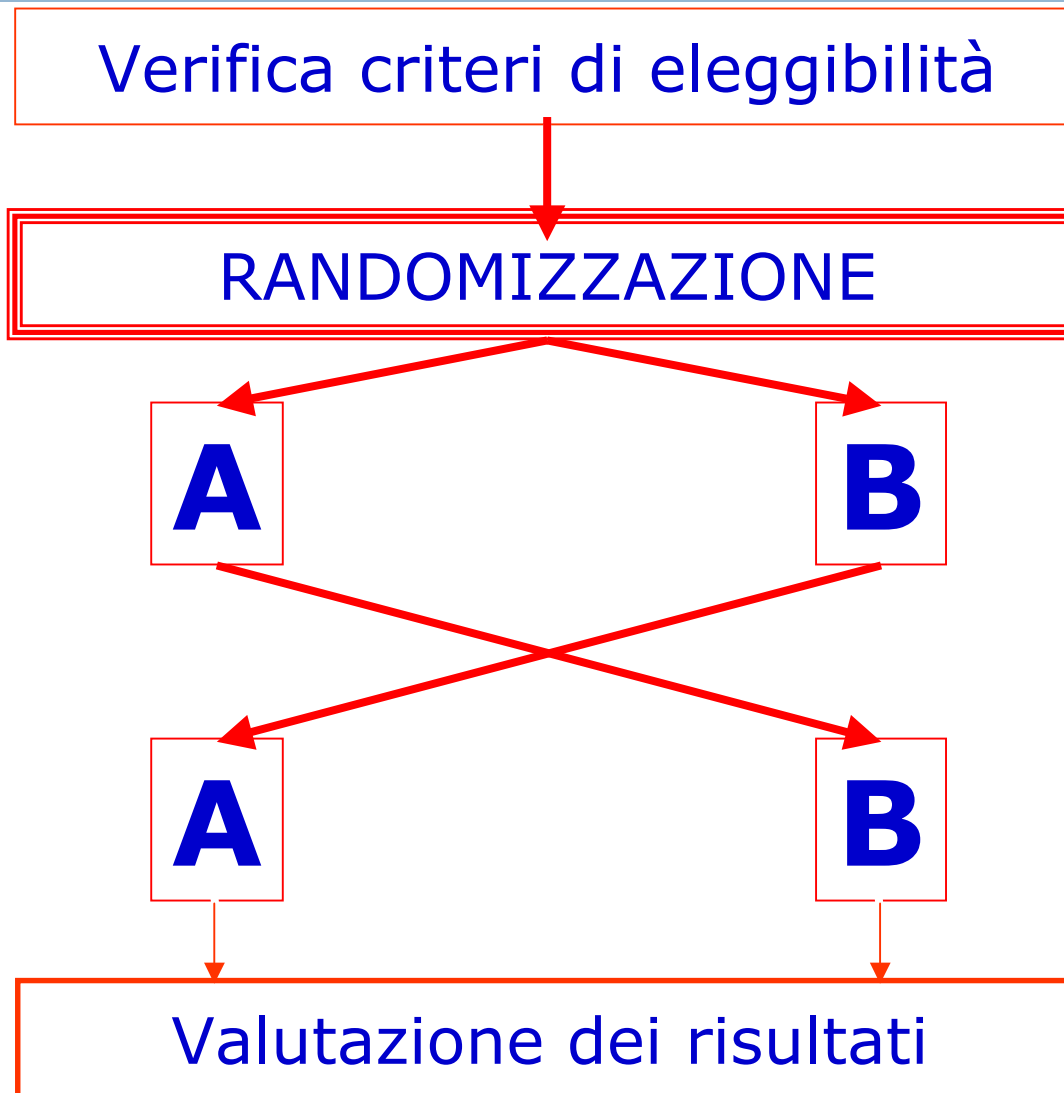


Fig 1. CONSORT diagram. XELOX, capecitabine and oxaliplatin; FOLFOX-4, infused fluorouracil, folinic acid, and oxaliplatin; ITT, intent-to-treat; EPP, eligible patient population; MCRC, metastatic colorectal cancer.

I diversi tipi di disegno

Disegno crossover



I diversi tipi di disegno

Disegno crossover

Vantaggi

Ogni soggetto funge da “controllo” di se stesso, rendendo lo studio più efficiente (si elimina il problema della variabilità fra individui) e consentendo una stima più precisa dell'effetto del trattamento con un numero minore di pazienti.

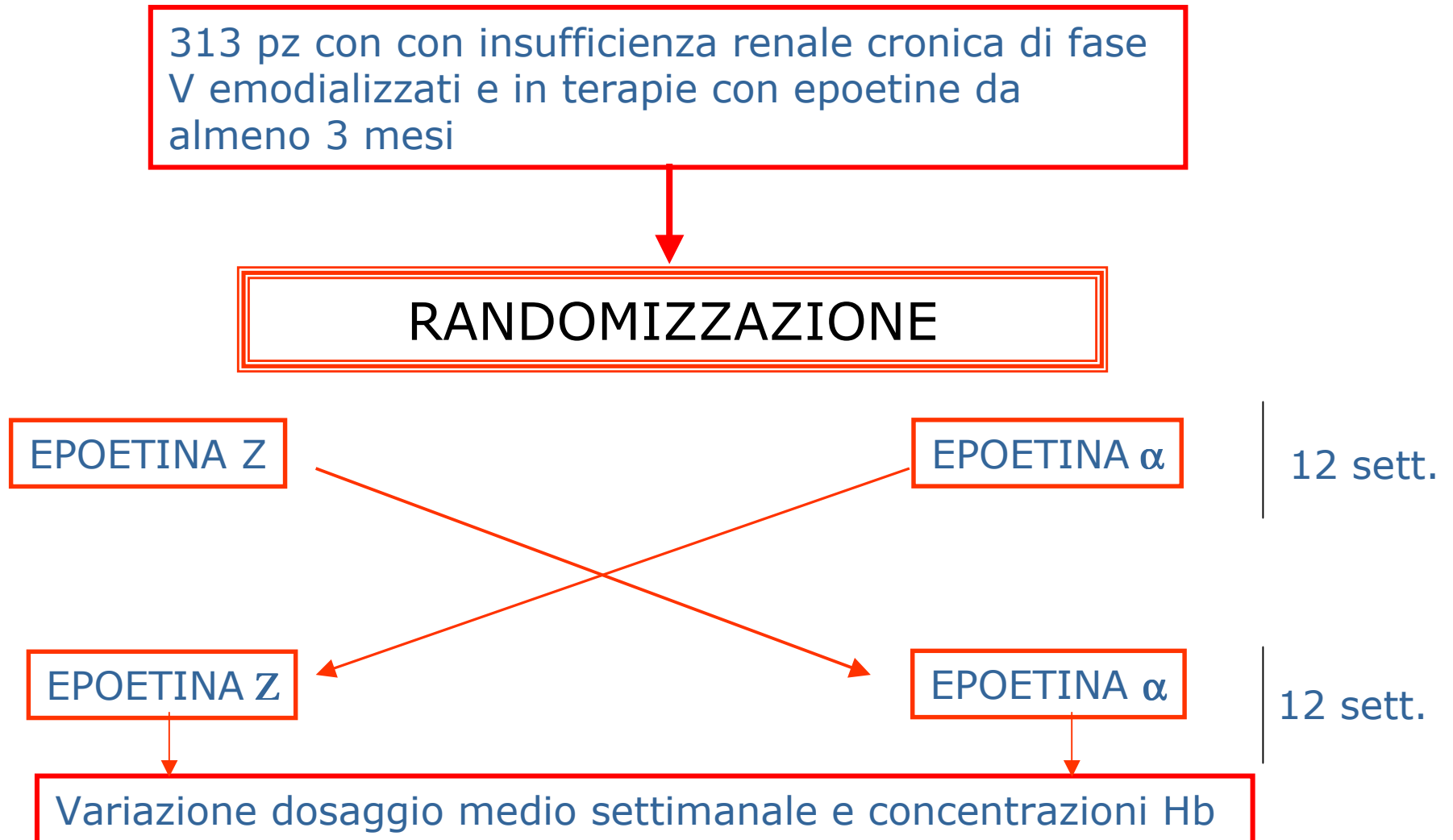
Svantaggi

Può essere utilizzato solo per studiare la risposta del paziente a terapie relativamente brevi, generalmente per patologie croniche in cui si è interessati a valutare il miglioramento a breve termine di segni e sintomi.

L'effetto residuo del primo trattamento (carry over) spesso richiede un periodo di washout (nessun trattamento).

La sequenza con cui i due trattamenti vengono somministrati può influenzare il risultato (effetto periodo).

Il ritiro dei pazienti dallo studio può influire seriamente sull'analisi.



Requisiti necessari di una sperimentazione clinica



Validità scientifica:
se
esiste

**Capacità di evidenziare un effetto
esso**

Generalizzabilità:
maggior
pazienti

**Trasferibilità dei risultati alla
parte dei**

Aspetti da valutare per verificare la validità scientifica



- **Procedura di randomizzazione**
- **Cecità dei pazienti / medici / esaminatori**
- **Stima delle dimensioni del campione e modalità di analisi statistica**
- **Modalità e durata del follow-up e informazioni sui pazienti persi al follow-up**

Aspetti da valutare per verificare la generalizzabilità



- **Descrizione dei criteri di inclusione/esclusione dei pazienti**
- **Descrizione dei regimi terapeutici**
- **Presentazione degli end-points principali**
- **Presentazione di valutazioni retrospettive**
- **Discussione degli effetti collaterali dei trattamenti**

Randomizzazione

Tecnica che garantisce che ogni soggetto abbia una probabilità prestabilita (abituamente la stessa probabilità) di essere assegnato ad ognuno dei trattamenti in studio, senza tuttavia che sia possibile prevedere quale di essi riceverà.

La verifica dell'eleggibilità del paziente e la decisione di inserirlo nella sperimentazione devono precedere l'arruolamento e l'assegnazione ad uno dei bracci dello studio.

Da questo punto di vista, l'assegnazione alternata o basata sulla data di nascita (pari o dispari) o altri metodi simili non possono essere considerati adeguati.

I metodi più utilizzati sono le buste chiuse e la **randomizzazione telefonica centralizzata**.

Tuttavia solo quest'ultima fornisce la garanzia che il clinico non possa conoscere a priori a quale braccio il paziente verrà assegnato.

DOPPIO CIECO

Si definiscono in doppio cieco gli RCT in cui né il paziente, né il medico o il valutatore sono a conoscenza di quale sia il trattamento somministrato.

Il paziente

Se il paziente sa di ricevere un nuovo trattamento, ciò può essere positivo psicologicamente. Al contrario, il paziente che sa di essere sottoposto al trattamento standard (o a nessun trattamento), può reagire negativamente, soprattutto se è a conoscenza che altri pazienti "privilegiati" ricevono un trattamento innovativo.

Per contro, altri pazienti potrebbero sentirsi più rassicurati se sottoposti al trattamento "standard".

Tutto ciò potrebbe determinare gradi diversi di adesione e di abbandono dello studio.

DOPPIO CIECO

Il medico

Se il medico sa che un paziente sta ricevendo un nuovo trattamento, potrebbe seguirne l'andamento con più attenzione di quanto farebbe con la terapia standard.

Se il medico sa che un paziente sta ricevendo il trattamento standard o nessun trattamento, potrebbe essere portato ad utilizzare cure supplementari.

L'entusiasmo (o la sfiducia) per un nuovo trattamento potrebbe essere trasmesso al paziente, influenzandone l'atteggiamento

Il valutatore

È di fondamentale importanza che chi valuta la risposta del paziente sia obiettivo al massimo.

Di conseguenza, la conoscenza del trattamento ricevuto da ogni singolo paziente può creare grossi problemi legati alle aspettative di beneficio di un nuovo trattamento.

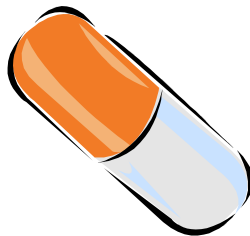
DOUBLE DUMMY

Negli studi in cui si confrontano due farmaci attivi di colori e/o forme, vie di somministrazione, regimi terapeutici diversi, per mantenere la condizione di cecità è necessario ricorrere all'uso di due placebo.

Farmaco
A



Farmaco
B



Placebo
A



Placebo
B



IL PLACEBO

- Il fatto di assumere un trattamento può in sé produrre dei benefici al paziente, quindi l'uso del placebo è essenziale per valutare se un intervento ha degli effetti che vanno al di là dell'**effetto placebo**.
- L'effetto placebo può variare enormemente in relazione alle caratteristiche del paziente, del farmaco, del medico, della precedente esposizione ad altri trattamenti .
- L'uso del placebo non è etico in presenza di altri trattamenti di comprovata efficacia.
- Gli studi vs placebo hanno scarsa generalizzabilità, in quanto nella pratica clinica il medico non è chiamato a scegliere tra un trattamento e il placebo.

FOLLOW UP



E' importante valutare se i pazienti randomizzati sono usciti anticipatamente dallo studio e per quali motivi:

- Decesso.
- Rifiuto del paziente o del medico.
- Eventi clinici che hanno portato il paziente a non ricevere il trattamento.
- **Effetti collaterali.**
- **Mancanza di efficacia clinica.**

Analisi Intention To Treat (ITT)

Si tratta di una strategia applicata nell'analisi di uno studio che confronta i pazienti mantenendoli all'interno del gruppo a cui erano stati randomizzati, a prescindere da eventuali violazioni del protocollo, dal livello di compliance o dall'abbandono prematuro dello studio.

Qualsiasi esclusione successiva alla randomizzazione potrebbe infatti rendere i gruppi messi a confronto non più comparabili.

L'intention to treat analysis fornisce inoltre stime più realistiche del beneficio di un trattamento, perché tiene conto della non compliance e delle deviazioni dalle pratiche standard.

Analisi Per Protocol (PP)



Strategia di analisi che considera i dati esclusivamente dei pazienti aderenti al protocollo al momento dell'analisi.

Non considera i pazienti randomizzati ma che non hanno ricevuto il trattamento o hanno interrotto lo studio precocemente.

Può presentare bias in particolare se il trattamento presenta effetti avversi o è inefficace; con tale analisi viene sovrastimato il beneficio e vengono sottostimati gli eventi avversi del trattamento in studio.

Valutazione secondo la scala JADAD

1. Lo studio è descritto come randomizzato?	μ Sì =+1	μ No =0
2. Se è descritto il metodo di randomizzazione, i dettagli forniti sul metodo stesso sono sufficienti a stabilire che l'attribuzione casuale ad uno dei trattamenti sia stata condotta in modo corretto per tutti i pazienti?	μ Sì =+1	μ No =-1
3. Lo studio è descritto come in doppio cieco?	μ Sì =+1	μ No =0
4. Se descritto, il metodo utilizzato per garantire la doppia cecità risulta adeguato allo scopo?	μ Sì =+1	μ No =-1
5. E' indicato il numero dei pazienti che non sono stati aderenti al trattamento, che sono stati trasferiti dal trattamento, o che sono stati sottratti al follow-up, e sono indicate le ragioni di questi fenomeni?	μ Sì =+1	μ No =0
PUNTEGGIO TOTALE		

(*) Sì = 1, No = 0

(**) Sì = 1, No = -1, mancata descrizione = 0


Se il punteggio è < 3, lo studio è di scarsa qualità

Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds JM, Gavaghan DJ, McQuay HJ. Assessing the quality of reports of randomized clinical trials: is blinding necessary?. *Controlled Clinical Trials*. 1996; 17: 1-12.

La stima delle dimensioni del campione

Sensibilità Capacità di evidenziare una differenza fra i trattamenti, se essa esiste

Specificità Capacità di escludere una differenza fra i trattamenti, se essa non esiste

Falso positivo  Rigetto ipotesi nulla quando è vera
ossia si evidenzia una differenza quando non c'è

Falso negativo  Accetto l'ipotesi nulla quando è falsa
ossia non si evidenzia

una differenza quando esiste
In uno studio clinico viene decisa la probabilità di risultato falso negativo (β) o di risultato falso positivo (α) che accettiamo di avere.

In generale, α è fissato ad un livello ≤ 0.05 , mentre β non deve mai superare 0.20, ma valori più bassi sono preferibili (0.05-0.10).

La stima delle dimensioni del campione

$1-\beta$ rappresenta la **potenza statistica** di uno studio, ossia la probabilità di cogliere un determinato effetto del trattamento se questo, in realtà, esiste.

La potenza statistica è determinante per interpretare il risultato: molti studi sono numericamente inadeguati per cogliere effetti realistici, ma la loro incapacità a dimostrare l'esistenza di un beneficio **NON indica in alcun modo che tale beneficio sia assente.**

Gli studi di non inferiorità

- Dal punto di vista statistico necessitano di una numerosità campionaria più bassa in quanto l'analisi dei risultati viene fatta con test ad una coda.
- I risultati non permettono di rilevare alcuna differenza di efficacia a vantaggio del trattamento sperimentale ma solo di rilevare la non inferiorità rispetto allo standard, a meno che non sia pianificata a priori anche l'analisi di superiorità.
- E' necessario valutare attentamente dal punto di vista clinico i limiti di non inferiorità definiti nell'analisi statistica.

ANALISI "POST HOC": L'ESEMPIO DEL CETUXIMAB



L'ipotesi dello studio e le eventuali analisi per sottogruppi dovrebbero essere fissate prima dell'inizio della raccolta dei dati: l'analisi a posteriori di tali dati produce evidenze di basso livello perché potrebbero non avere la potenza statistica adeguata per rilevare una differenza significativa, in quanto non pianificate a priori nell'analisi statistica.

e) ANALISI DEGLI STUDI CLINICI

Ref.	Disegno/ fase	Pazienti e trattamento	Misure di esito	Risultati
Cristal NEJM 2009	RCT/III Analisi retrospettiva	540 pz K colonretto metast, 1° linea con valutazione Kras Cmab+FOLFIRI vs FOLFIRI	1°= PFS 2°= OS	9,9 vs 8,7 (p=0.017) 24.9 vs 21 (p=ns)
Opus JCO 2009	RCT/II Analisi retrospettiva	233 pz K colonretto metast, 1° linea con valutazione Kras Cmab+FOLFOX4 vs FOLFOX4	1°= OR (CR+PR) 2°= PFS	61% vs 37% (p=0.011)
EPIC JCO 2008	RCT/III aperto	1298 pz K colonretto metast, EGFR+, 2° linea Cmab+IRINO vs IRINO	1°= OS 2°= PFS, CR	10,7 vs 10 (p=ns) 4 vs 2,6 (p<0.0001) 16,4% vs 4,2%
EPAR	Analisi retrospettiva EPIC	300 pz con valutazione kras	1°= OS 2°= PFS	10,9 vs 11,6 (p=ns) 4,0 vs 2,8 (p=ns)
NEJM 2007	RCT/III aperto	572 pz K colonretto metast, 3° linea, EGFR+ Cmab+ BSC vs BSC	OS	6,1 vs 4,6 (p=0.005)
NEJM 2008	Analisi retrospettiva del precedente	394 pz con valutazione kras	OS	9,5 mesi vs 4,8 mesi OS% 1 anno: 28,3% vs 20,1% (p<0.001)

OS: overall survival, PFS: progression free survival, OR: overall response, Cmab: cetuximab, BSC: best supportive care

Valore di p



P<0.05 Se si rifiuta l'ipotesi nulla (ossia si dichiara che un trattamento è più efficace dell'altro), c'è una probabilità inferiore al 5% di affermare il falso

Limiti

- Valore soglia arbitrario
- La probabilità di trovare un $p < 0.05$ aumenta con il numero di tests eseguiti
- Il valore di p non fornisce indicazioni sulla rilevanza clinica della differenza trovata

Intervalli di Confidenza

In qualsiasi studio, viene valutato un **campione casuale** della popolazione costituita da tutti i pazienti affetti da quella specifica patologia e si fa l'assunzione che tale campione sia rappresentativo dell'intera popolazione.

In realtà, se lo studio fosse ripetuto tante volte, non otterremmo sempre un risultato identico; è quindi importante cercare di stimare quale sia **l'effetto vero**, che si otterrebbe studiando l'intera popolazione, e non un campione di essa.

Gli intervalli di confidenza al 95% indicano che c'è la probabilità del 95% che l'effetto vero cada all'interno di un certo intervallo.

Misure di efficacia di un trattamento

Gruppo	Evento		Totale
	Si	No	
Controllo	a	b	a+b
Sperimentale	c	d	c+d

Tasso (rischio) di eventi nel gruppo di controllo = CER = $a/a+b$

Tasso (rischio) di eventi nel gruppo sperimentale = EER = $c/c+d$

Rischio relativo = RR = CER/EER

Riduzione relativa del rischio = RRR = $(CER-EER)/CER$

Riduzione assoluta del rischio = ARR = CER - EER

Number needed to treat = NNT = $1/ARR = 1/(CER - EER)$

LA PRESENTAZIONE DEI RISULTATI CLINICI: L'ESEMPIO DEL RIVAROXABAN

R. Lassen et al, N Engl J Med 2008; 358: 277 - 86

RIVAROXABAN VS ENOXAPARIN FOR THROMBOPROPHYLAXIS AFTER
TOTAL KNEE ARTHROPLASTY

End point 1°composito (studio di non inferiorità e, se dimostrata, di superiorità): TVP, embolia polmonare non fatale, morte per qualsiasi causa dai 13 ai 17 giorni dopo intervento

Rivaroxaban: 9,6%	➔	ARR 9,2% (CI 95% 5.9 – 12.4; p<0.001)
Enoxaparina: 18,9%		RRR 49% (CI 35 – 61; p<0.001)
		NNT=11

I principali bias degli studi clinici

- Mancanza di randomizzazione e condizioni di cecità
- Calcolo della dimensione del campione
- Scelta degli end- point (compositi, surrogati)
- Scelta dell'analisi statistica (PP vs ITT)
- Analisi per sottogruppi non pianificate a priori
- Modalità e durata del follow up
- Conflitto di interesse

Bias di pubblicazione

E' stato stimato che circa l'85% degli studi pubblicati riporta risultati positivi. E' verosimile, quindi, che gli studi negativi non siano stati sottoposti ad alcuna rivista scientifica.

Gli studi pubblicati rispetto ai non pubblicati presentano in percentuale maggiore risultati statisticamente positivi



REGISTRI DEGLI STUDI CLINICI

Impact factor

Strumento quantitativo, proprietà di Thomson Reuters, per categorizzare, valutare, confrontare, ordinare le riviste scientifiche.

Calcolo:

$$\frac{\text{n. citazioni in un particolare anno di articoli pubblicati nei 2 anni precedenti}}{\text{n. articoli pubblicati sulla rivista nei 2 anni precedenti.}}$$

Caratteristiche principali che consentono ad una rivista di essere presa in considerazione per la valutazione dell'IF:

- Puntualità pubblicazione;
- Processo di valutazione editoriale basato sul peer review;
- Presenza di abstract e informazioni bibliografiche in inglese;
- Internazionalità degli autori;
- Interesse del contenuto scientifico in relazione all'attuale coperture della specifica tematica;
- Presenza di citazioni nelle riviste già censite da Thomson Reuters.

NB: la bontà o rilevanza di uno studio clinico non è necessariamente correlata con il prestigio della rivista ove è pubblicato.

Sintetizzare le evidenze

La quantità delle informazioni cui si trova di fronte l'operatore sanitario chiamato ad affrontare un quesito clinico assistenziale è enorme e in continua crescita. Oltre alla quantità delle informazioni esiste il problema della loro qualità.

I risultati degli studi primari sono talvolta ambigui o addirittura discordanti tra loro. Basare dunque le proprie scelte sulla lettura di un singolo studio o dare eccessivo credito a notizie che riguardano l'ultima scoperta farmacologica spesso può rivelarsi un errore.

In campo medico le prime revisioni sistematiche e metanalisi sono state pubblicate su importanti riviste a partire dalla fine degli anni settanta. Da allora la quantità di questa tipologia di articolo all'interno dei database internazionali sta crescendo enormemente, a riflettere il suo passaggio definitivo da uno status emergente ad uno status di riconosciuto valore accademico e scientifico.

Poter accedere a revisioni affidabili e di buona qualità rappresenta un vantaggio enorme in termini di tempo e orientamento.



**THE COCHRANE
COLLABORATION®**

E' un'organizzazione internazionale che prepara, aggiorna e diffonde revisioni sistematiche sugli effetti degli interventi e delle procedure medico-sanitarie

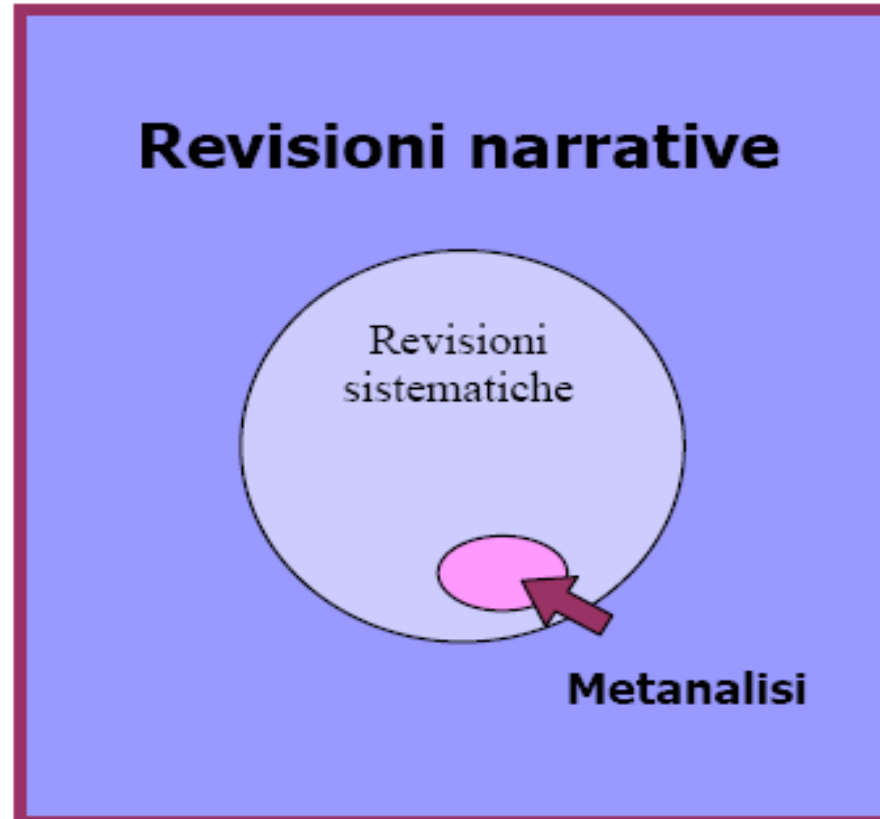
La CC è responsabile della Cochrane library, database che raccoglie tutte le RS e MA prodotte ed i protocolli in corso

La Cochrane library è ormai considerata la miglior fonte indipendente di evidenza affidabile circa gli effetti degli interventi sanitari.

Raccoglie la sfida del medico inglese cui deve il nome, Archibald Cochrane:

garantire che le decisioni in ambito sanitario, relative sia alle terapie sia alle future linee di ricerca, si basino su una sintesi accurata e su una valutazione critica di tutte le prove ed evidenze disponibili in letteratura.

Le revisioni



Le revisioni narrative danno una visione panoramica di un determinato argomento, di cui generalmente affrontano ogni aspetto. Es: capitoli di un libro di testo. L'esperto poi seleziona gli studi spesso in base a criteri soggettivi, e ne dà una descrizione solamente qualitativa.

Revisioni sistematiche

Le RS sono veri e propri progetti di ricerca che sintetizzano e valutano criticamente in un unico documento gli esiti di tutti gli studi sperimentali condotti riguardo ad un determinato e ben definito quesito clinico o intervento sanitario.

Per ridurre al minimo i rischi di distorsione i revisori si avvalgono, in ogni fase del processo di elaborazione, di una metodologia scientifica standardizzata.

Dal momento che le RS sono per natura retrospettive, esse devono essere basate su un dettagliato ed esaustivo protocollo di ricerca.

Tappe fondamentali da esplicitare nel protocollo:

- ✓ Definizione del quesito
- ✓ Ricerca sistematica delle fonti
- ✓ Valutazione della qualità metodologica degli studi eleggibili
- ✓ Discussione delle ragioni di concordanza e discordanza tra i risultati dei diversi studi.

Revisioni sistematiche / metanalisi

Nell'ambito di una RS può (se sussistono le condizioni di similarità tra tipi di pazienti e di trattamenti indagati) venire realizzata una meta-analisi, ovvero una combinazione quantitativa pesata dei risultati dei singoli studi.

Una RS non si conclude quindi forzatamente con una meta-analisi.

Là dove è prevalente un alto grado di eterogeneità degli studi può essere fuorviante sviluppare una meta-analisi. In questi casi la RS dovrebbe privilegiare un approccio qualitativo.

NB: Principale critica alle RS/MA è di fare "un'insalata di risultati" mettendo insieme dati di studi fra loro non comparabili per quesito, qualità metodologica, tipologia di pazienti ecc....

I risultati di una RS possono essere presentati in diverse forme: commenti scritti, espressioni numeriche, grafici o tabelle.

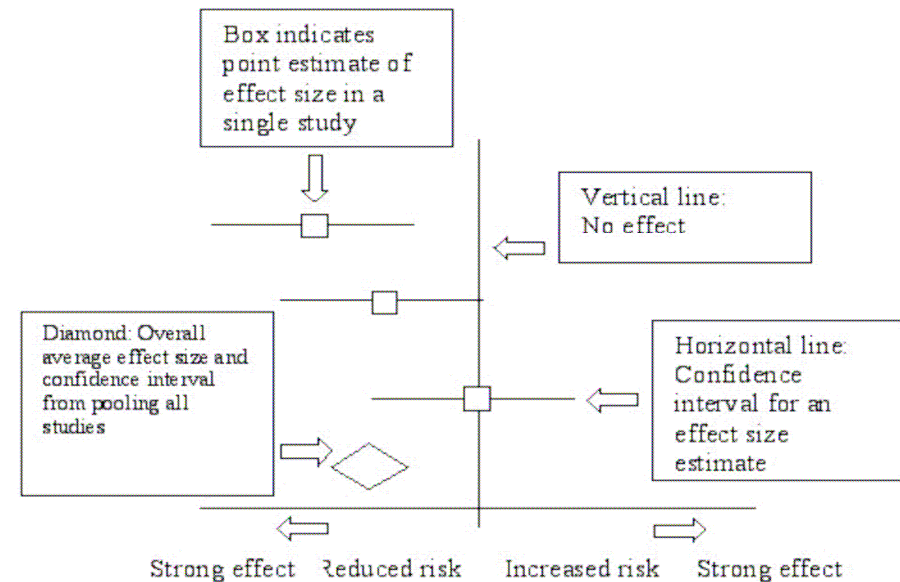
In qualunque caso, i risultati di una revisione fanno sempre riferimento a precise misure di effetto espresse come rischio relativo, o differenza di rischio.

Meta-analisi

Il **Forest Plot** rappresenta la modalità di presentazione delle componenti di una meta-analisi.

Permette di visualizzare la "foresta" dell'evidenza senza perdere di vista gli "alberi" dei singoli studi.

Il grafico facilita infatti la comprensione intuitiva dei risultati e fornisce, nel contempo, a chi lo legga con attenzione, tutte le informazioni principali su obiettivi, modalità e risultati della metaanalisi.



Interpretation of sample meta-analysis graph results

- Point estimates lie on the left of the vertical line, indicating a risk reduction from treatment.
- Confidence intervals that cross the vertical line include the possibility of no treatment effect.
- The diamond, the overall summary effect derived from the meta-analysis, indicates risk reduction. Its confidence intervals do not cross the vertical line.

Se CI attraversa la linea verticale, i risultati dello studio, convenzionalmente, vanno considerati come statisticamente non significativi ($\alpha=0.05$). Questo non significa che non siano "veri" o "validi" ma piuttosto che il numero di pazienti analizzati nello studio non è sufficiente a garantire che i risultati siano indipendenti dal "caso".

La qualità di una revisione sistematica

- ✓ La validità di una RS e la rilevanza dei suoi risultati sono dipendenti da quanto i singoli studi sono stati ben disegnati e condotti da un punto di vista metodologico e da quanto i loro risultati riguardano pazienti rappresentativi della realtà nella quale i risultati devono essere applicati.
- ✓ Combinare sia in modo qualitativo (RS) o anche quantitativo (MA) un insieme di studi gravati da bias importanti conduce a conclusioni gravate dagli stessi bias
- ✓ Il rischio più grande è che queste conclusioni vengano accettate e a queste venga dato ancora più peso di quanto si darebbe a un singolo studio.
- ✓ Una volta che gli studi primari hanno superato l'esame di inclusione / esclusione è quindi sempre importante valutarne la qualità metodologica.

OPEN ACCESS Freely available online

PLOS MEDICINE

Guidelines and Guidance

July 2009 | Volume 6 | Issue 7 | e1000100

The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration

Alessandro Liberati^{1,2*}, Douglas G. Altman³, Jennifer Tetzlaff⁴, Cynthia Mulrow⁵, Peter C. Gøtzsche⁶, John P. A. Ioannidis⁷, Mike Clarke^{8,9}, P. J. Devereaux¹⁰, Jos Kleijnen^{11,12}, David Moher^{4,13}

Preferred Reporting Items for Systematic reviews and Meta-Analyses:
• 27-item checklist
• four-phase flow diagram.

Bibliografia

- ✓ Stuart J Pocock. Clinical Trials- a practical approach. 1996, Centro Scientifico Torinese
- ✓ Bobbio M., Trial clinici. Come interpretare e applicare i risultati di una ricerca scientifica 1996, Centro Scientifico Editore
- ✓M. Belfiglio "Gli elementi essenziali di uno studio clinico" - Corso di aggiornamento UVEF "La lettura e l'interpretazione degli studi clinici: approfondimenti teorici ed applicazioni pratiche". Verona 18-19 giugno 2007
- ✓ Smith F.G. Clinical Research 2003, BIOS Scientific Publishers Limited
- ✓ Faggiano F. et al, Manuale di epidemiologia per la sanità pubblica 2005, Centro Scientifico Editore
- ✓<http://www.cochrane.org/>
- ✓Liberati A. et al, PloS Med 6 (7): e1000100.doi:10.1371/journal.pmed.1000100