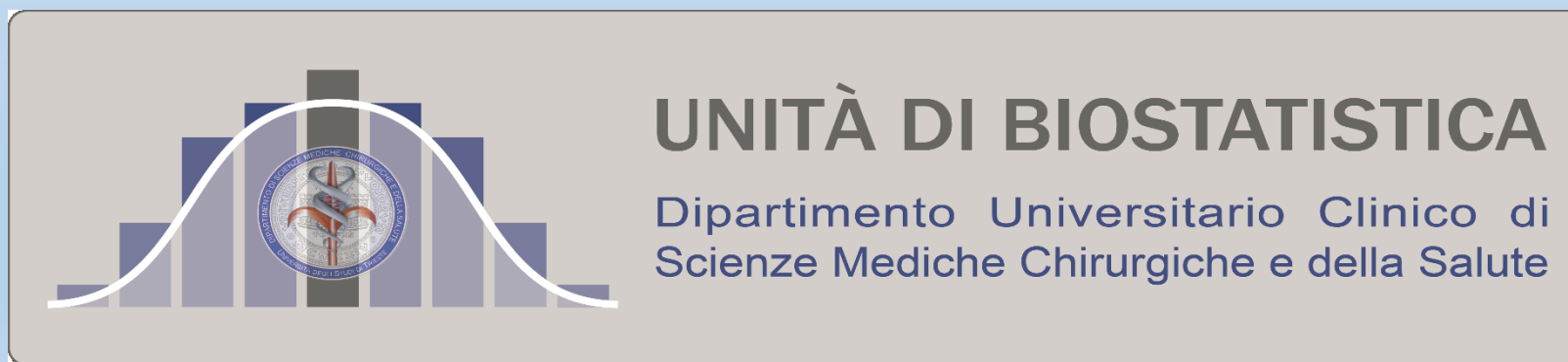


CDL in SCIENZE RIABILITATIVE DELLE PROFESSIONI SANITARIE

STATISTICA PER LA RICERCA

gbarbati@units.it

A.A. 2025-26



Il materiale di questa lezione è stato gentilmente fornito dal Prof. Roberto D'Amico, Università degli Studi di Modena e Reggio Emilia

La ricerca clinica e l'incertezza...

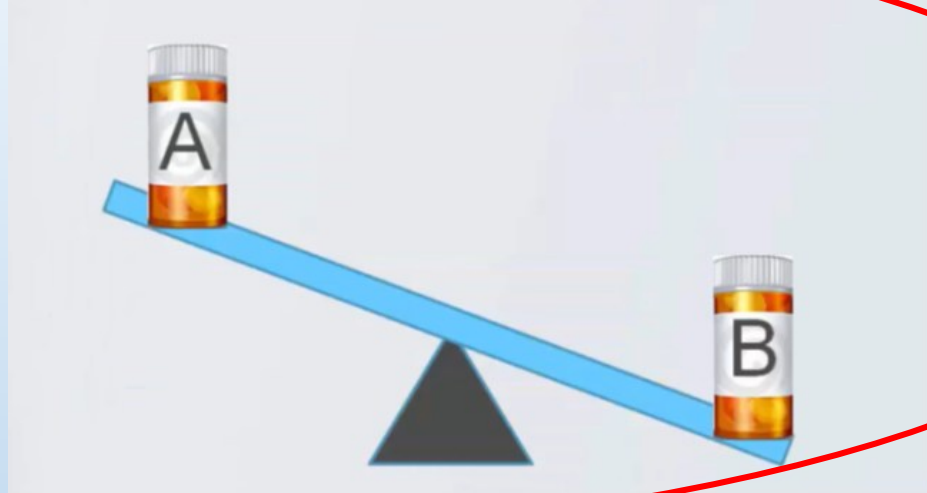
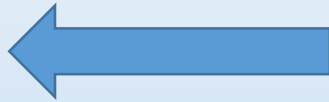
“In medicina l'incertezza è fisiologica, un certo grado di incertezza è fisiologico, però dentro questa incertezza ce n'è certamente una quota che potrebbe essere evitata, che potrebbe essere ridotta. In fin dei conti, il ruolo del ricercatore è proprio questo: è quello di ridurre quell'incertezza che è riducibile”

Alessandro Liberati, 2012

Revisioni Sistematiche & Metanalisi

I quesiti clinici più frequenti:

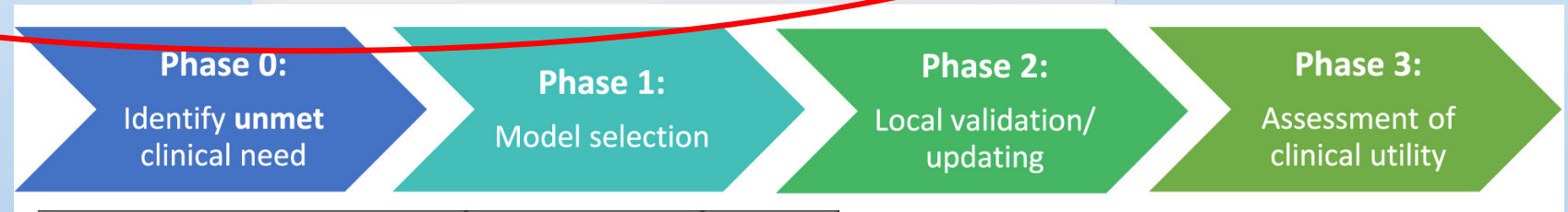
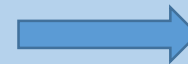
- Efficacia



- Prognosi



- Diagnosi



		True result (Did he/she have the outcome?)		Total
		Yes	No	
Predicted result (Is his/her predicted probability/score larger than the cut-off value?)	Yes	a	b	a+b
	No	c	d	c+d
Total number of people		a+c	b+d	a+b+c+d

Il quesito clinico di **efficacia** :

Nelle persone affette da una determinata patologia o condizione, quale intervento (tra A e B) è quello più efficace e sicuro?

Nelle **persone ricoverate in terapia intensiva** il rischio di morte è molto elevato.
Una delle cause di **mortalità** è lo sviluppo di **infezioni nosocomiali**.

Somministrare la **profilassi antibiotica** rispetto a non somministrarla potrebbe:

- a) Ridurre il rischio di infezioni nosocomiali e quindi il rischio di morte (**efficacia**)
- b) Aumentare le resistenze antibiotiche (**sicurezza**)

Disegno di studio più appropriato: RCT (Studio Clinico Randomizzato)

L'identificazione del quesito clinico

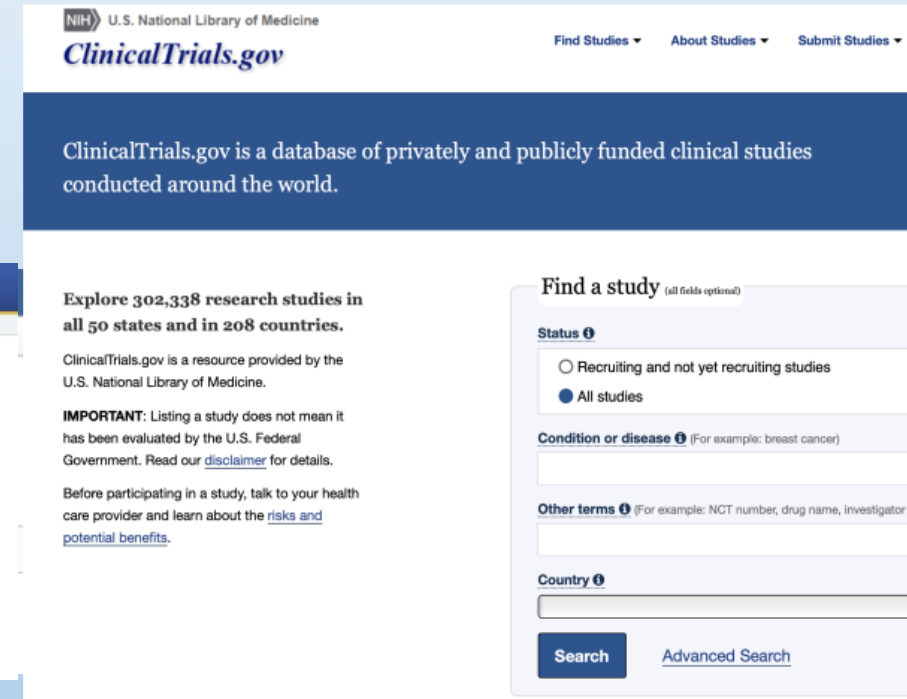
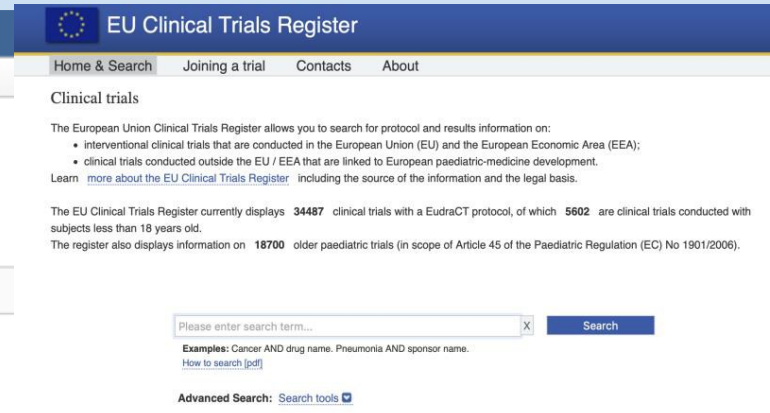
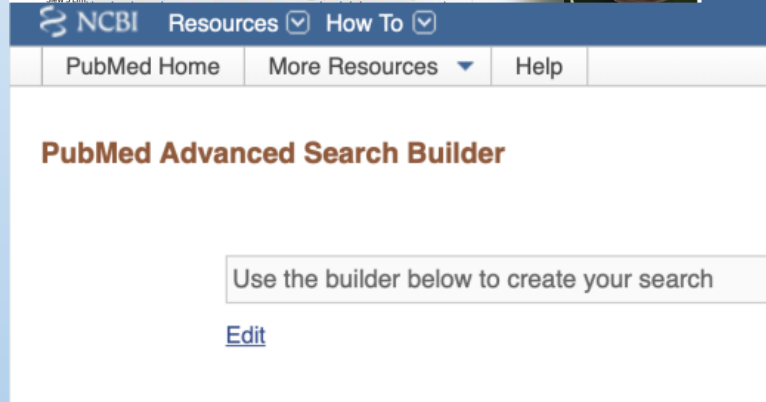
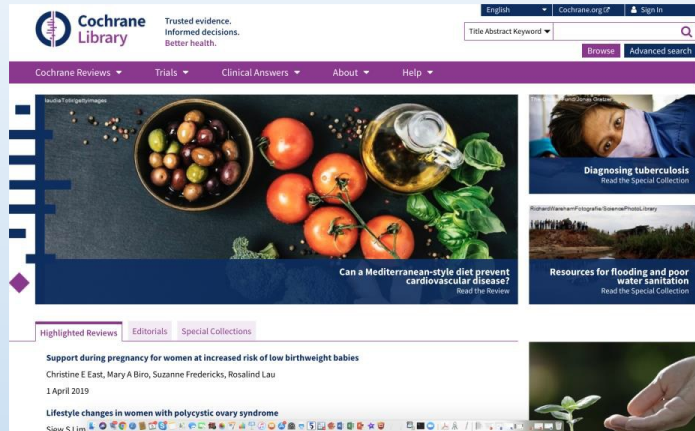
- Incertezza : risposte assenti/insoddisfacenti
- Rilevanza : paziente/medico/decisore sanitario

PICO (s)

- Qual è la **P**opolazione che si vuole studiare?
- Qual è l'**I**ntervento che si vuole valutare?
- Qual è il **C**omparatore?
- Qual è l'**O**utcome (esito clinico) che si vuole misurare?
- Qual è il disegno di **S**tudio che si vuole includere?

Cosa è stato **pubblicato** sull'argomento ?

Ogni anno vengono pubblicati più di **1 milione** di papers ... Più di **30.000** tra riviste scientifiche e divulgative ...



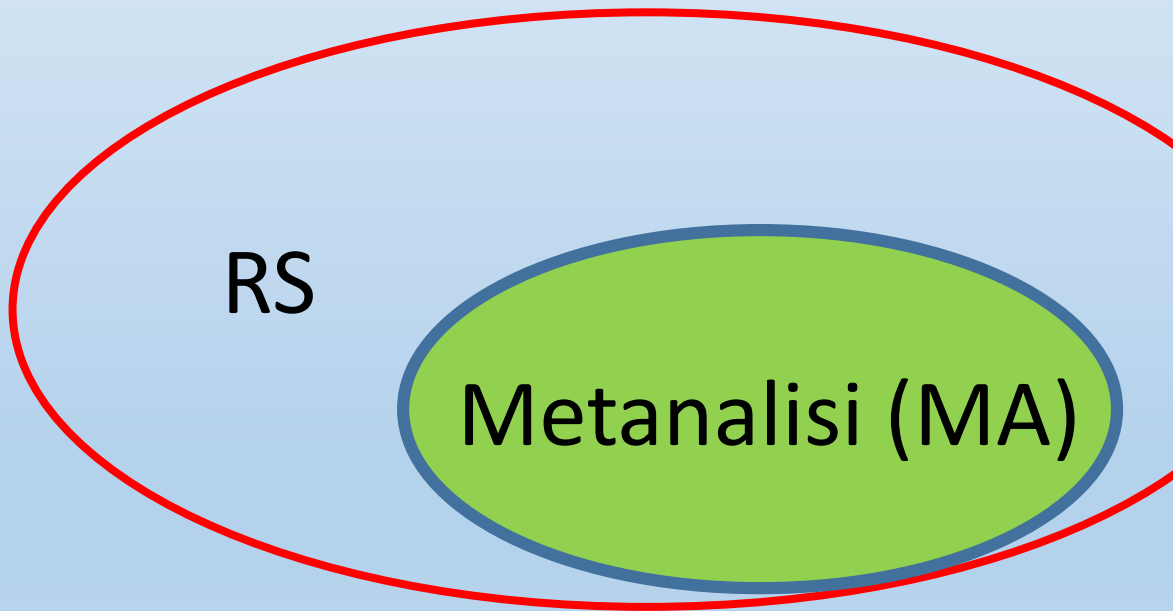
Esigenza di una metodologia per la raccolta, valutazione e **sintesi** delle migliori evidenze scientifiche

- Review *tradizionali* vs **Revisioni sistematiche**
- Oggi si usano le **review sistematiche** perché sono **riproducibili** (ricerche **automatizzate** con dei filtri da database)

Le revisioni sistematiche (RS)

- Individuare nella letteratura scientifica gli studi che hanno affrontato il quesito
- Valutarne la qualità metodologica
- Sintetizzarne i risultati (metanalisi)

Va scritto un **protocollo** della revisione sistematica !!



The screenshot shows the PROSPERO website homepage. At the top, the NIHR logo (National Institute for Health and Care Research) is on the left, and the PROSPERO logo (International prospective register of systematic reviews) is on the right. Below the logos is a navigation bar with links for Home, FAQs, and About PROSPERO. On the right side of the navigation bar are links for Search, Login, and Join us. The main content area features a large banner image with a blue background and white dots, and a text box that says 'PROSPERO home page'. Below the banner is a section titled 'What is PROSPERO?' which describes the registry's purpose and current status. To the right of this section is a dark blue button labeled 'Registering a review' and a text box that says 'Registering a review is quick and easy and free of charge. You must be logged in to register a review.' At the bottom right is a button labeled 'Register your review now'.

NIHR | National Institute for Health and Care Research

PROSPERO
International prospective register of systematic reviews

Home FAQs About PROSPERO ▾ Search Login Join us

PROSPERO home page

What is PROSPERO?

PROSPERO is an international systematic review registry that aims to promote transparency and open science, reduce reporting bias and help prevent unintended duplication and research waste. The PROSPERO database currently includes records of over **393000** prospectively registered systematic reviews with health related outcomes, providing easy access to key information about planned, in-progress and completed reviews.

PROSPERO is produced by the Centre for Reviews and Dissemination (CRD) and funded by the National Institute for Health and Care Research (NIHR)

Registering a review

Registering a review is quick and easy and free of charge. You must be logged in to register a review.

Register your review now

Lo **strumento** che ci guida nella selezione (automatizzata) dei papers :

PICO (s)

- Qual è la **P**opolazione che si vuole studiare?
- Qual è l'**I**ntervento che si vuole valutare?
- Qual è il **C**omparatore?
- Qual è l'**O**utcome (esito clinico) che si vuole misurare?
- Qual è il disegno di **S**tudio che si vuole includere?

Esempio

La *decontaminazione selettiva* del tratto digerente (SDD) riduce il rischio di sviluppare *infezioni nosocomiali* nei *pazienti in terapia intensiva*?

P= Pazienti in terapia intensiva

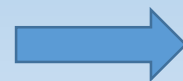
I = Profilassi antibiotica (SDD)

C = Standard Care

O = Infezioni nosocomiali (mortalità)

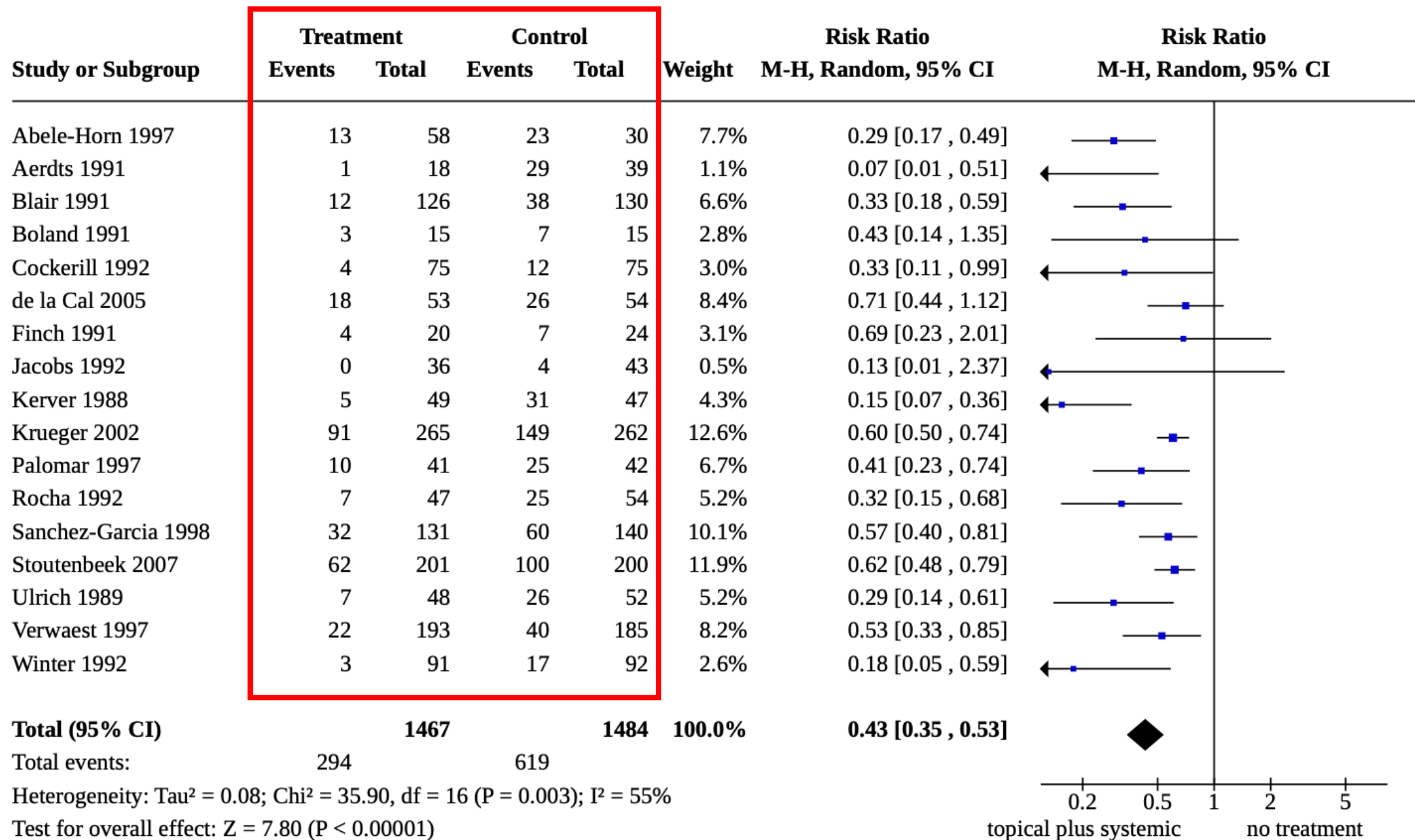
(**S**= RCT)

Processo di revisione sistematica



17 studi

Le informazioni essenziali



La scelta della **misura di associazione** per la MA

Per esiti/outcome binari:

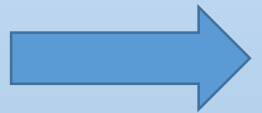
- Il rischio relativo (RR)✓
- L'Odds Ratio (OR) ✓
- *La differenza tra Rischi (DR) ...meglio di no**

* ...Se facciamo una MA i risultati che **combiniamo** sono in generale le **misure relative** più che **assolute**; questo perché l'incidenza «basale» degli eventi di interesse può essere diversa tra le varie popolazioni studiate...

	Trattamento	Controllo	Totale
Evento	a	b	$a+b$
Non evento	c	d	$c+d$
Totale	$a+c$ N° trattati	$b+d$ N° controlli	N

Per ogni studio (outcome binario) si stima/estrae:

- Il rischio relativo (RR) [oppure l'odds ratio (OR)]
- Il logaritmo naturale (ln) del RR [o OR]
- La varianza campionaria di (ln) RR [o OR]



La distribuzione campionaria sulla scala del logaritmo è approssimabile da una gaussiana

$$\log(\widetilde{RR}) \sim \text{Normal}(\log(RR), \text{var}(\log(RR))) \quad \log(\widetilde{OR}) \sim \text{Normal}(\log(OR), \text{var}(\log(OR)))$$

	Trattamento	Controllo	Totale
Evento	a	b	a+b
Non evento	c	d	c+d
Totale	$N_T = a+c$	$N_C = b+d$	N

$$RR = a/N_T : b/N_C$$

$\ln RR$

$$\text{var}(\ln RR) = \frac{1}{a} + \frac{1}{b} - \frac{1}{N_T} - \frac{1}{N_C}$$

	SDD	Controllo	Totale
Infezione si	3	17	20
Infezione no	88	75	163
Totale	91	92	183

Interpretazione: il rischio passa da 0.19 a 0.03 che equivale a dire che c'è stata una riduzione relativa di 82% (1-RR)

Il rischio tra i trattati

$$R_T = 3/91 = 0.033 \text{ (3.3\%)}$$

Il rischio tra i controlli

$$R_C = 17/92 = 0.185 \text{ (18.5\%)}$$

$$RR = R_T / R_C = 0.18$$

$$\ln RR = -3.41$$

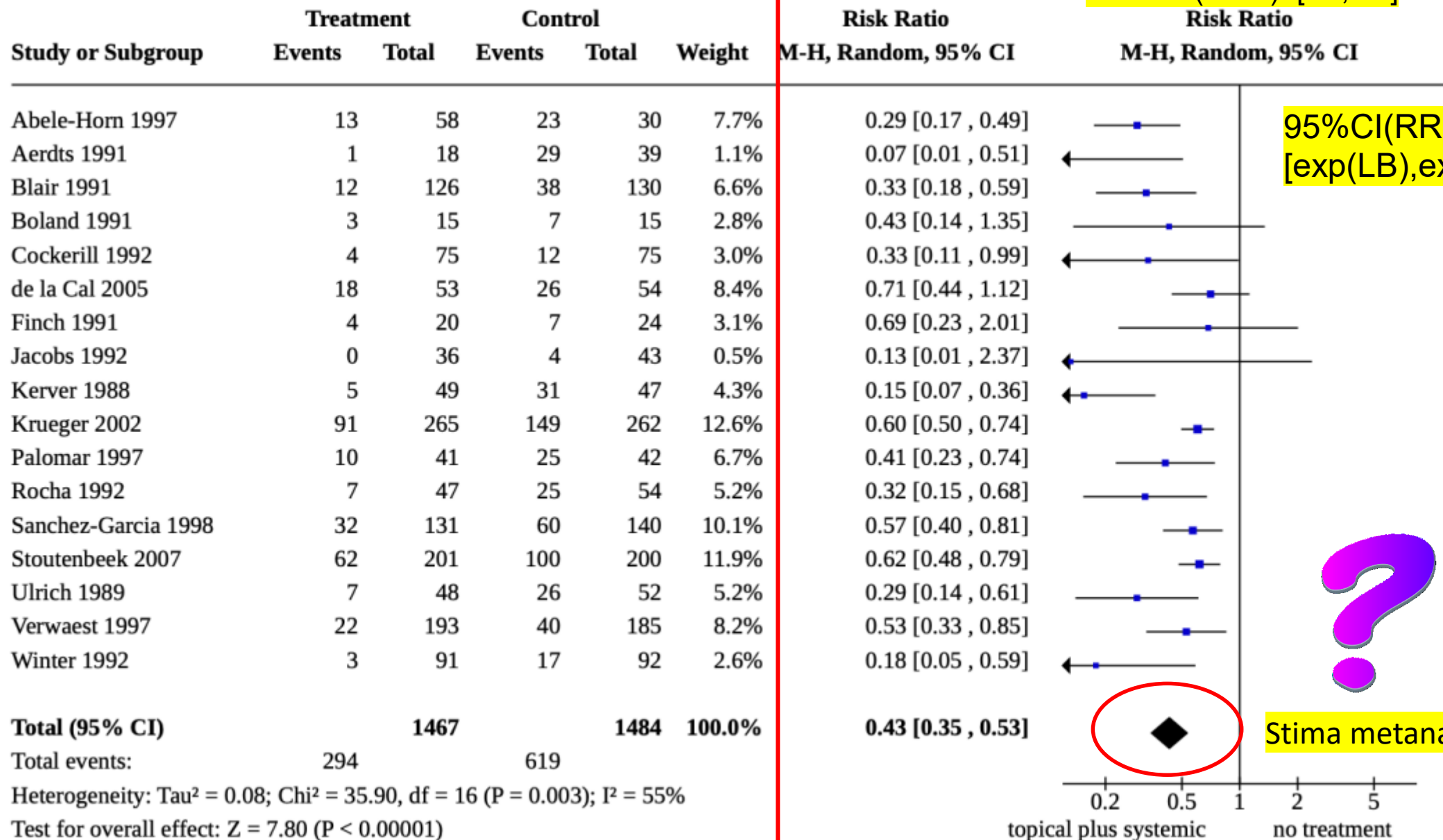
	Profilassi antibiotica		Controllo	
	N° pts con infezioni	N° pazienti	N° pts con infezioni	N° pazienti
1. Abele	...	...	...	...
16. Winter	3	91	17	92

Stima del RR e del relativo 95%CI

$$95\%CI(\ln RR) = \ln RR \pm 1.96 * SE(\ln RR) = [LB, UB]$$

$$95\%CI(RR) = [\exp(LB), \exp(UB)]$$

Forest plot



Stima metanalitica

Come meta-analizzo ??

Studio	Eventi/Trattamento	Eventi/Controllo	RR
A	40/100	50/200	1.6
B	15/200	5/100	1.5
Tot?	55/300	55/300	1

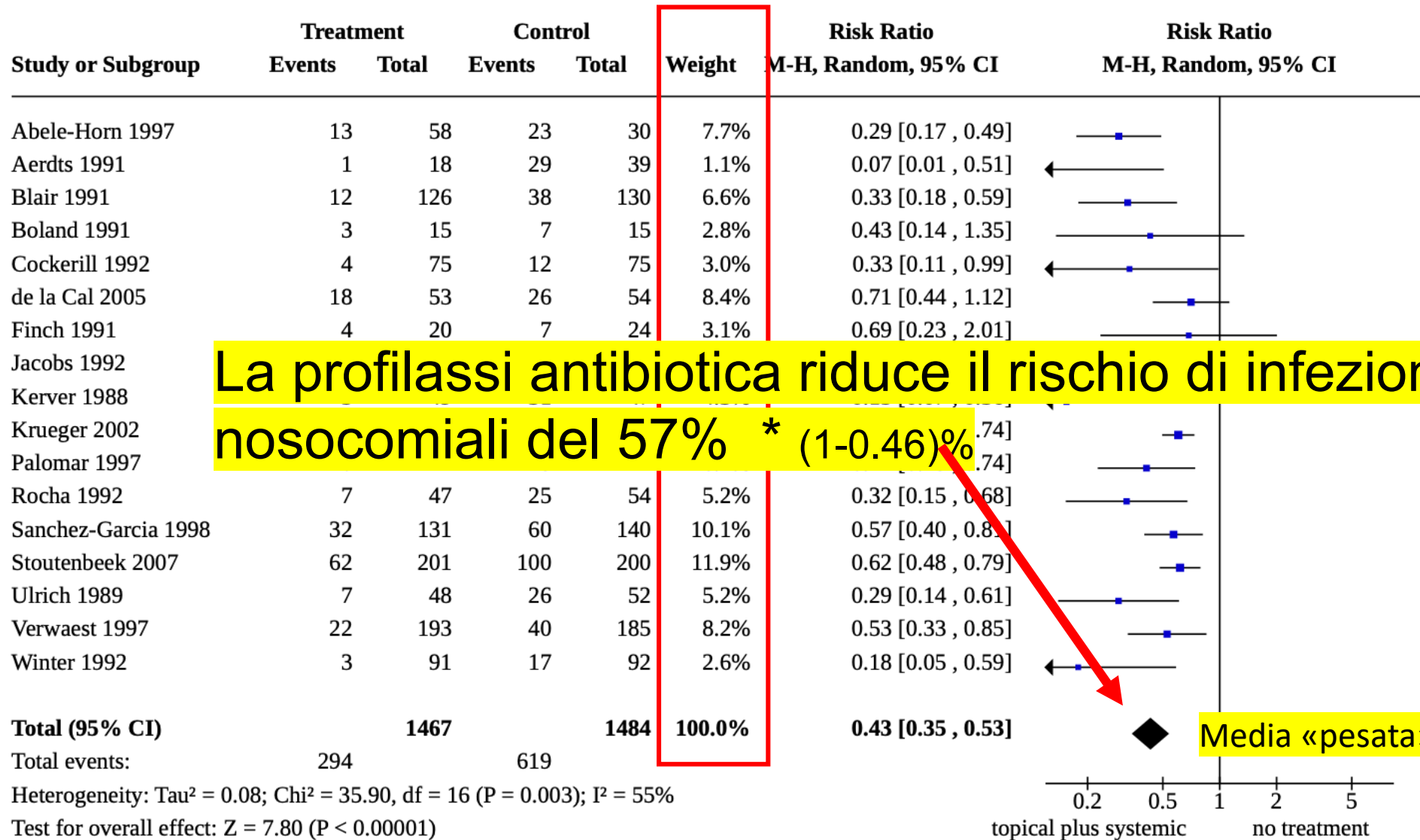


I metodi per combinare le stime:

- Il metodo dell' **inverso della varianza** (più comune)
- *Metodo di Peto (Yusuf-Peto) (solo per odds ratio)*
Sviluppato per combinare RCT con effetti piccoli. Esteso all'analisi dei dati di sopravvivenza.
- *Il metodo di Mantel-Haenszel (fixed-effects)*
Originariamente sviluppato per l'analisi stratificata.

La logica della combinazione dei risultati

Analysis 1.2. Comparison 1: Topical plus systemic prophylaxis versus placebo or no treatment, Outcome 2: Respiratory tract infections



La Media Pesata*:

$$Media\ pesata = \frac{\sum_{i=1}^n v_i p_i}{\sum_{i=1}^n p_i}$$

v è il valore (stima dello studio)

p è il peso

- Somma (Stima(studio i-esimo)*peso) / Somma pesi
- In caso di OR e RR si considera il loro logaritmo naturale.
- Alla fine della procedura si calcola l'esponenziale della media pesata e degli estremi del 95%CI

*c'è anche una misura di variabilità intorno alla media pesata rappresentata graficamente dall'ampiezza del simbolo

La Media Pesata

Il peso associato ad ogni studio è funzione:

- Della **dimensione** del campione e della **frequenza** dell' evento (in caso di esito binario)
- [Nel caso si usasse il *modello ad effetti random*, anche della *variabilità* dell'effetto del trattamento]

La dimensione del campione e frequenza dell'evento vengono considerati nel calcolo dell' **inverso della varianza** della stima.

$$p_i = \frac{1}{var(\ln(RR)_i)}$$

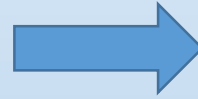
Requisiti

- Almeno due studi che riportano la stima dell'effetto del trattamento
- *Omogeneità* tra studi in termini di PICO
- Stessa misura dell'esito tra studi
- Disponibilità dei dati

Eterogeneità

Differenze **cliniche** tra studi

Formulazione, interventi, dose, durata della terapia



Eterogeneità statistica

Eccesso di variabilità tra i risultati degli studi
non compatibile con la variazione casuale

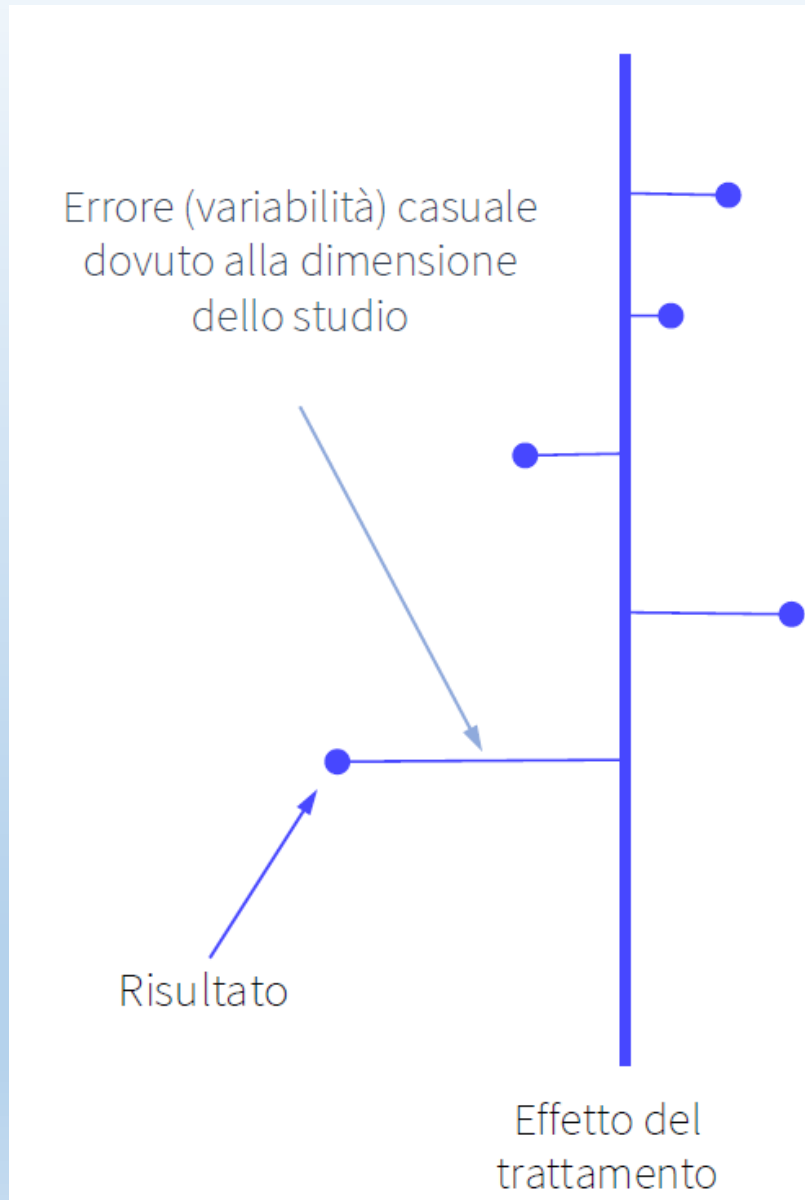
Differenze **metodologiche** tra studi

Disegno di studio

Cecità nell'assegnazione dei pazienti al trattamento

Qualità dello studio

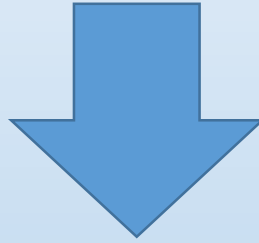
Omogeneità statistica



- Tutti gli studi valutano *lo stesso effetto* del trattamento
- i.e. se tutti gli studi fossero *di grandi dimensioni* si otterrebbero gli stessi risultati !

Eterogeneità statistica

Eccessiva *variazione* dei risultati degli studi



Variabilità dell'effetto del trattamento
nei vari studi *superiore* a quella
aspettata per il solo effetto del caso...

Identificare l'eterogeneità graficamente

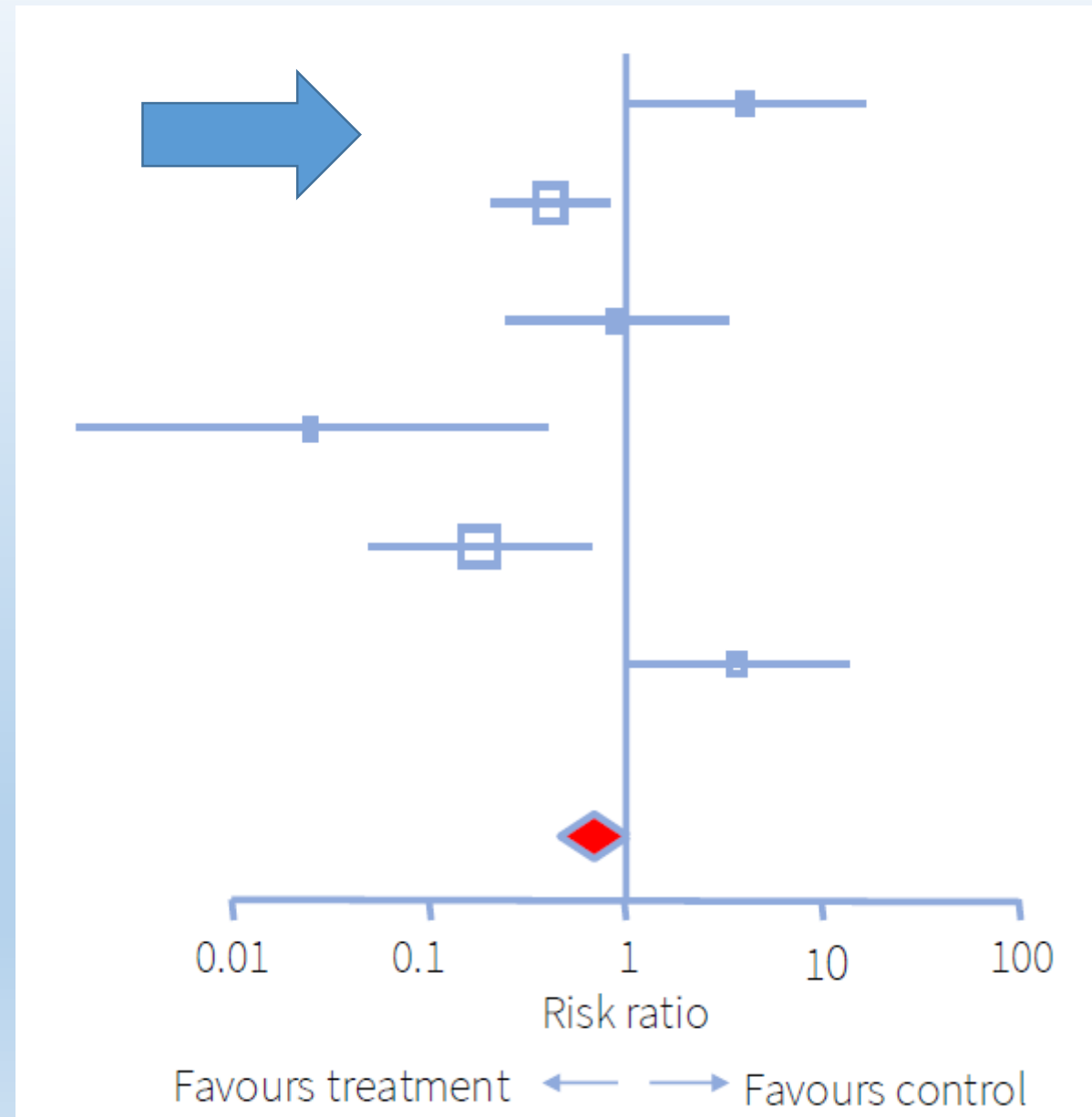
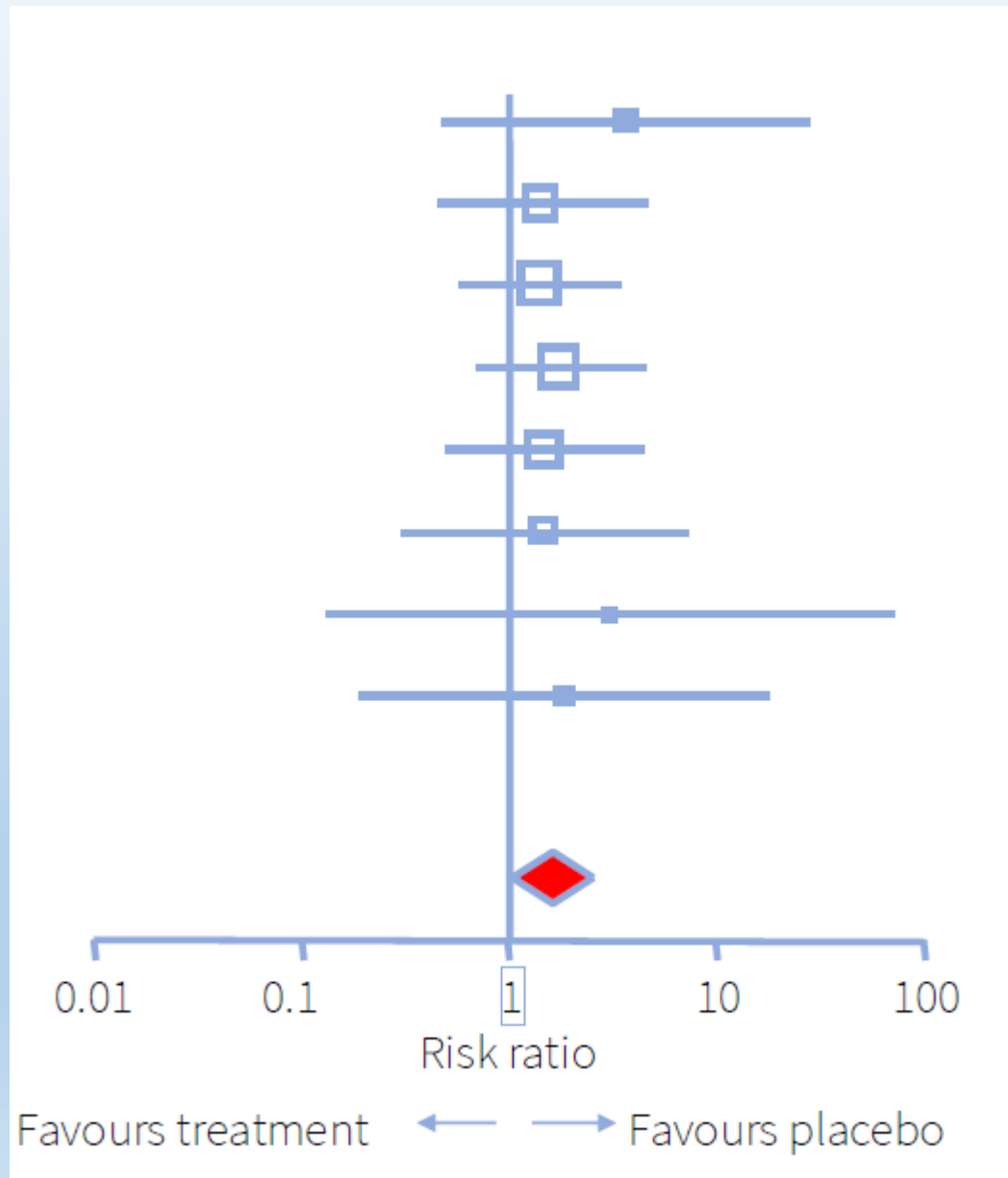
Se gli studi sono *simili* ci si aspetta che gli intervalli di confidenza si sovrappongano ...

L'eterogeneità statistica si manifesta nella *manca* di sovrapposizione degli intervalli di confidenza

(Oppure in presenza di risultati anomali)

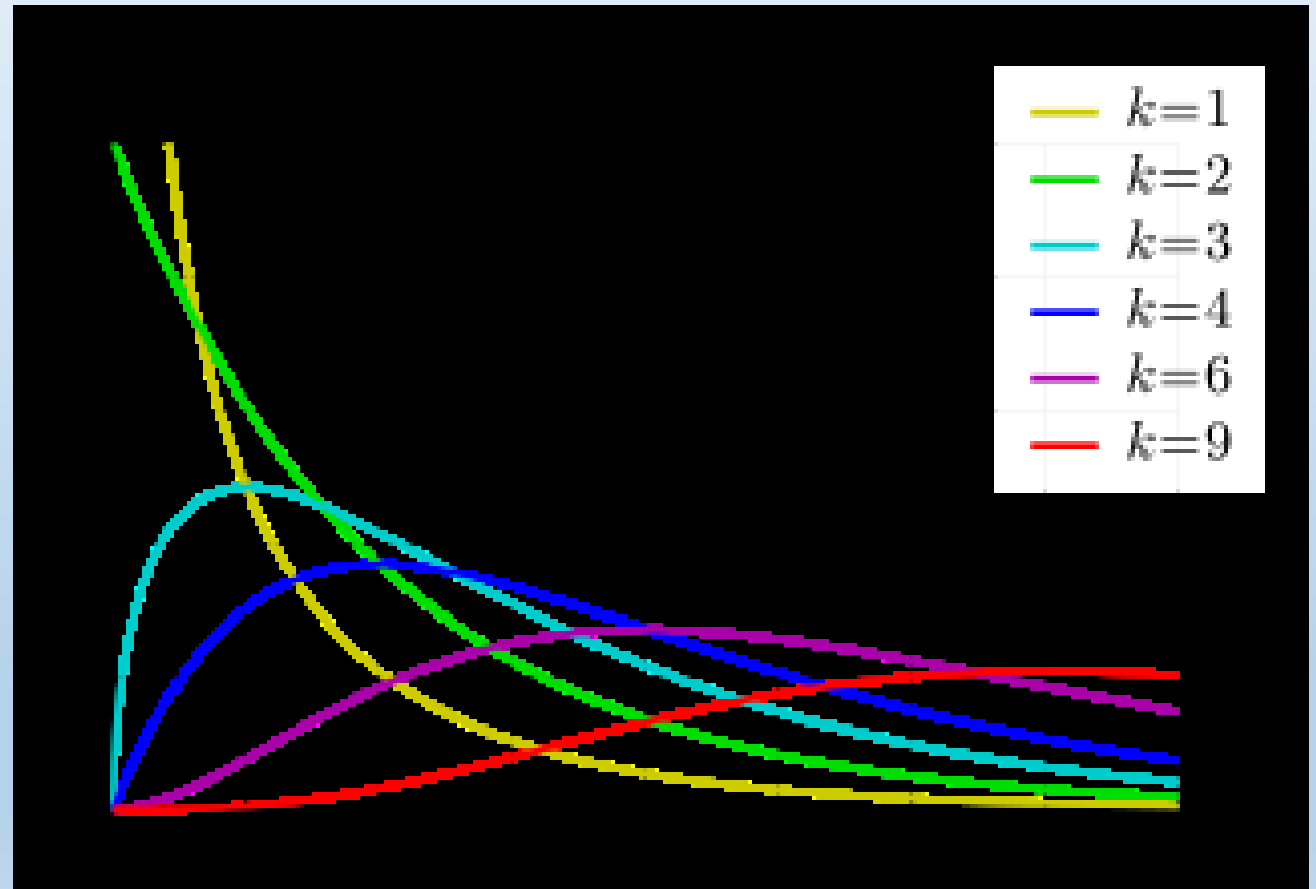


Identificare l'eterogeneità graficamente

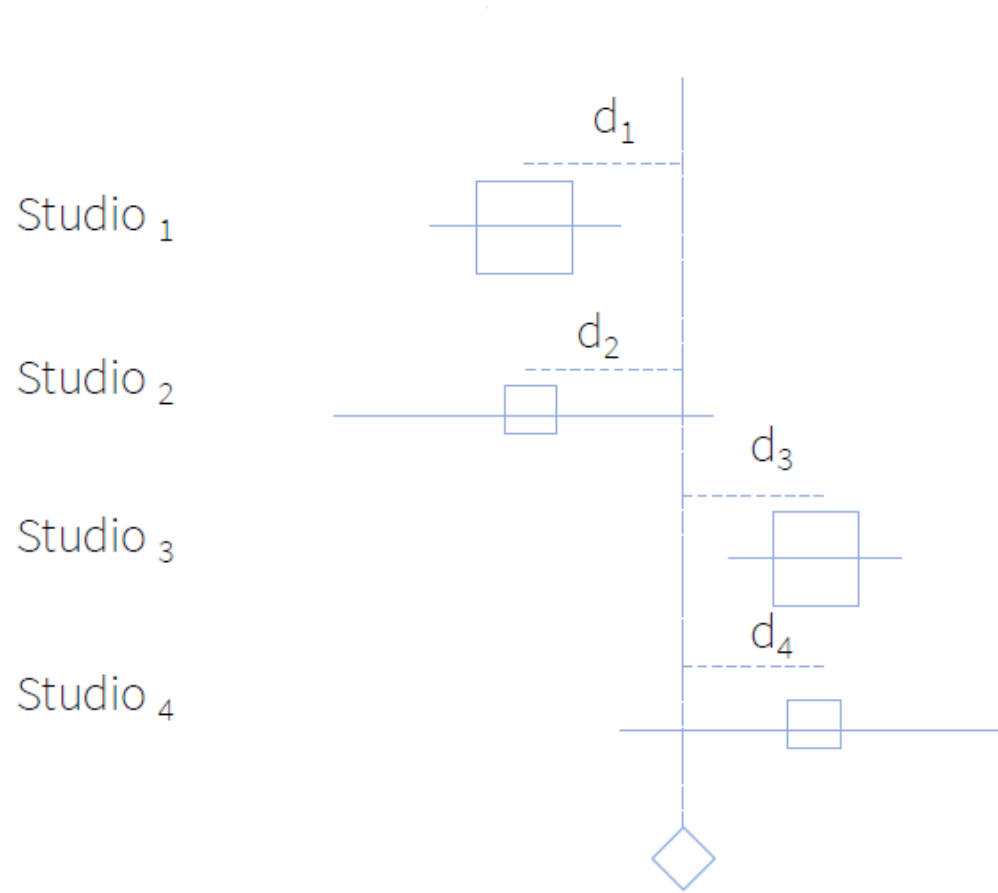


Identificare la presenza di eterogeneità attraverso l'uso di un test statistico

- Esiste un test statistico per valutare se i risultati degli studi considerati sono compatibili con l'**ipotesi nulla** di omogeneità
- Si calcola una misura di variabilità (Q). Questa quantità esprime la **distanza tra i risultati degli studi e la loro media**
- **Valori grandi** di Q suggeriscono eterogeneità (ovvero una variabilità più grande di quella aspettata solo per effetto del caso)



Il test che valuta l'ipotesi di omogeneità tra i risultati degli studi



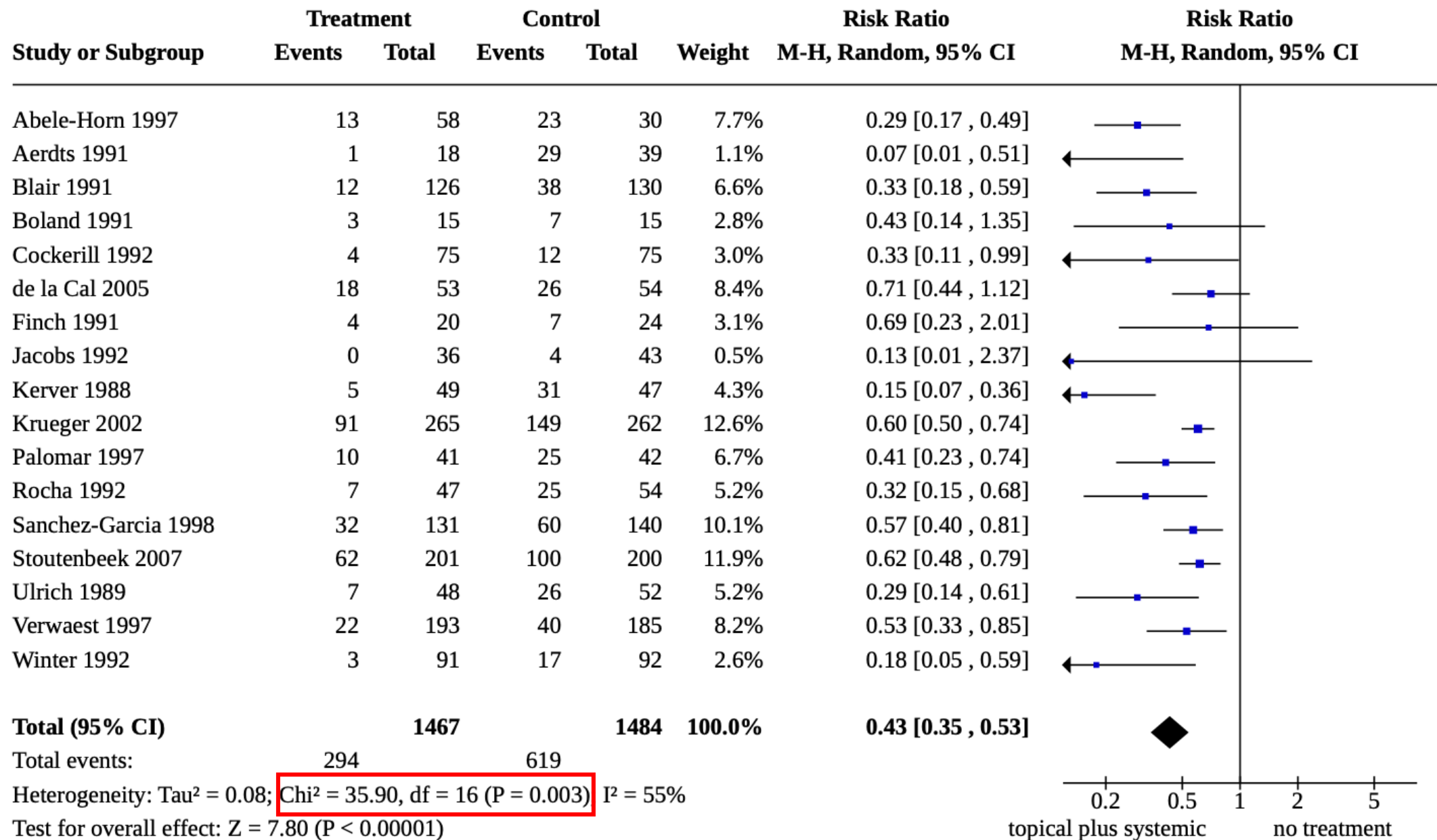
di rappresenta la **distanza** tra la stima dell'effetto del trattamento ottenuta nell' i -esimo studio e la **media pesata**

La somma **pesata** delle distanze:

$$Q = d_1p_1 + d_2p_2 + d_3p_3 + d_4p_4$$

Questa statistica di test si distribuisce secondo una distribuzione nota sotto l'ipotesi nulla (Chi-quadro con $k-1$ gdl) e quindi posso valutarne il p-value (dove k =numero degli studi)

Analysis 1.2. Comparison 1: Topical plus systemic prophylaxis versus placebo or no treatment, Outcome 2: Respiratory tract infections



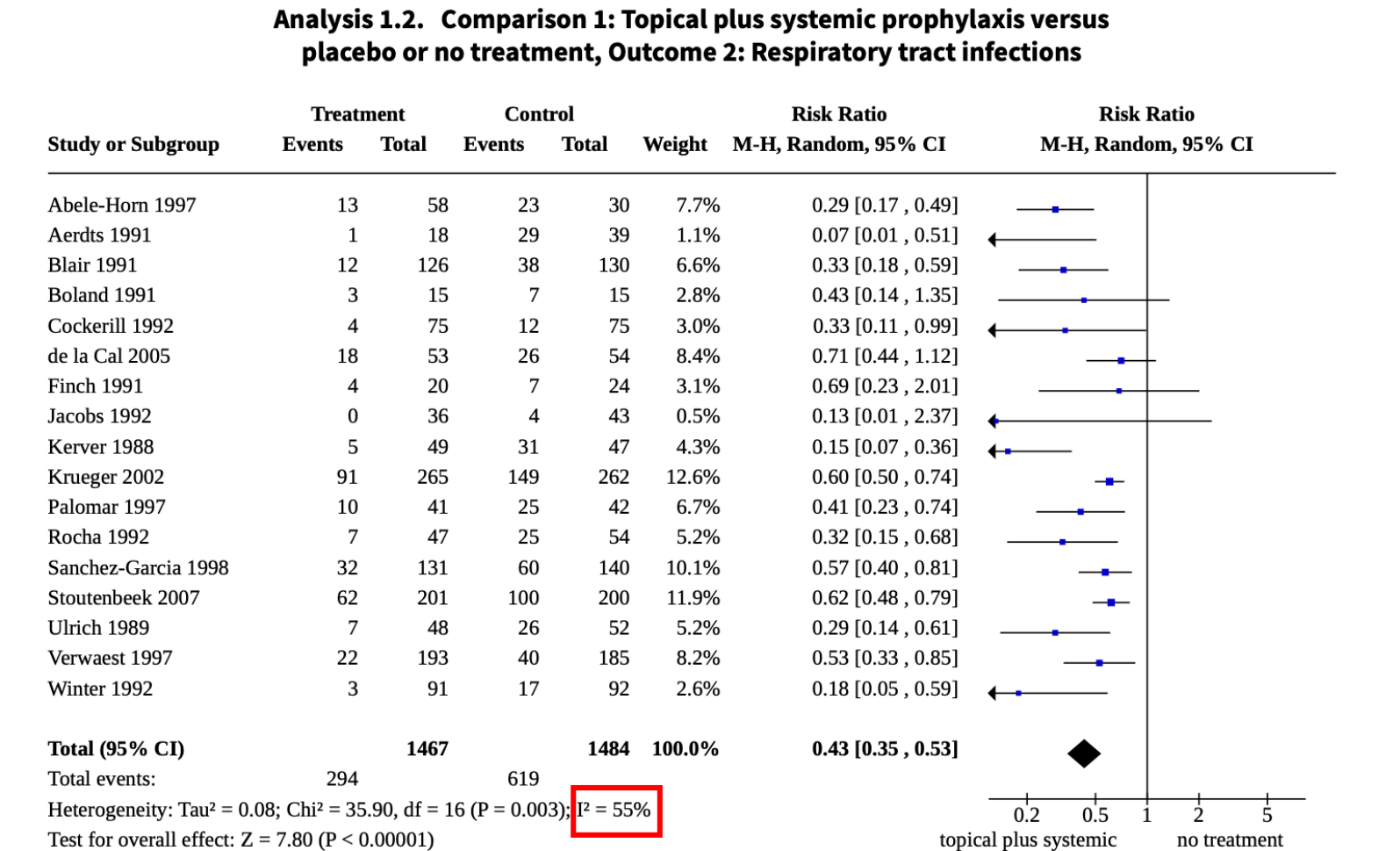
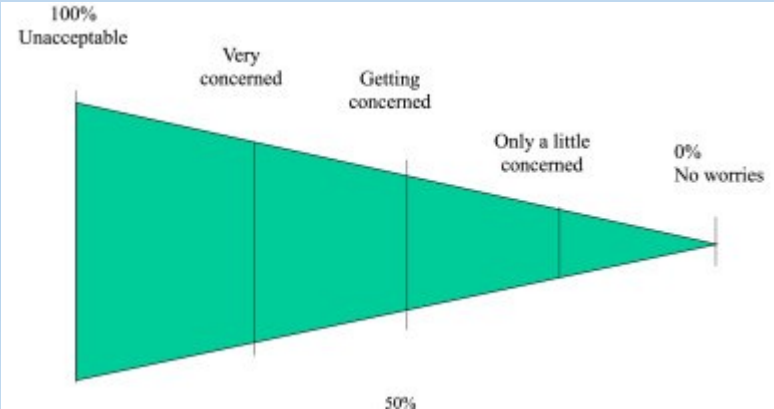
Misurare l'eterogeneità statistica: l'indice I^2

Posta pari a 100 la variabilità osservata tra i risultati, I^2 esprime quanta parte di tale variabilità *non è attribuibile* al caso.

Se c'è eterogeneità il più delle volte è attribuibile alla differenza clinica/metodologica ma il viceversa non è detto ...

$$I^2 = \frac{Q - (k - 1)}{Q}$$

$$I^2 = \frac{Q - (k - 1)}{Q} = \frac{35.90 - 16}{35.90} = 0.55$$

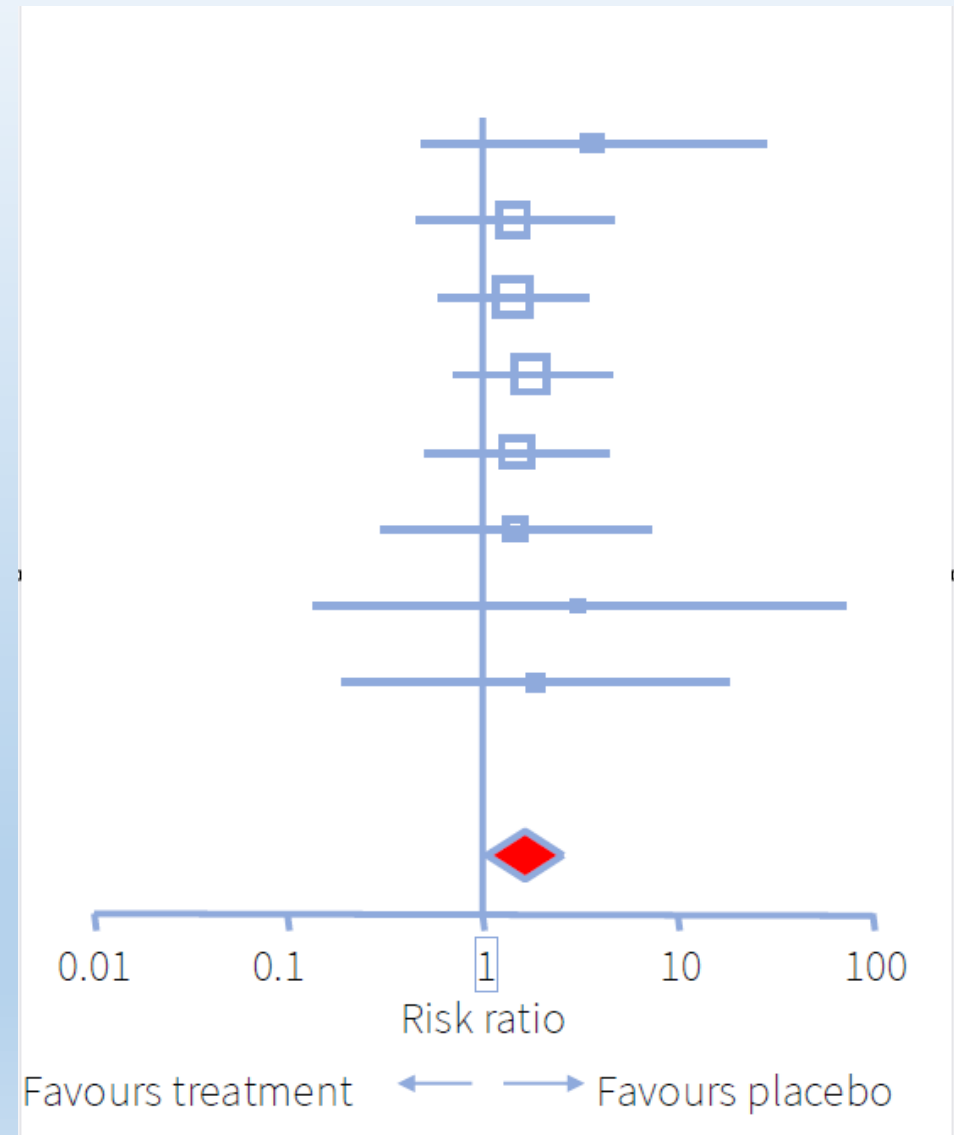


Esempio di eterogeneità clinica ma non statistica

Revisione Cochrane sulla albumina umana:
hypoalbuminaemia, morte

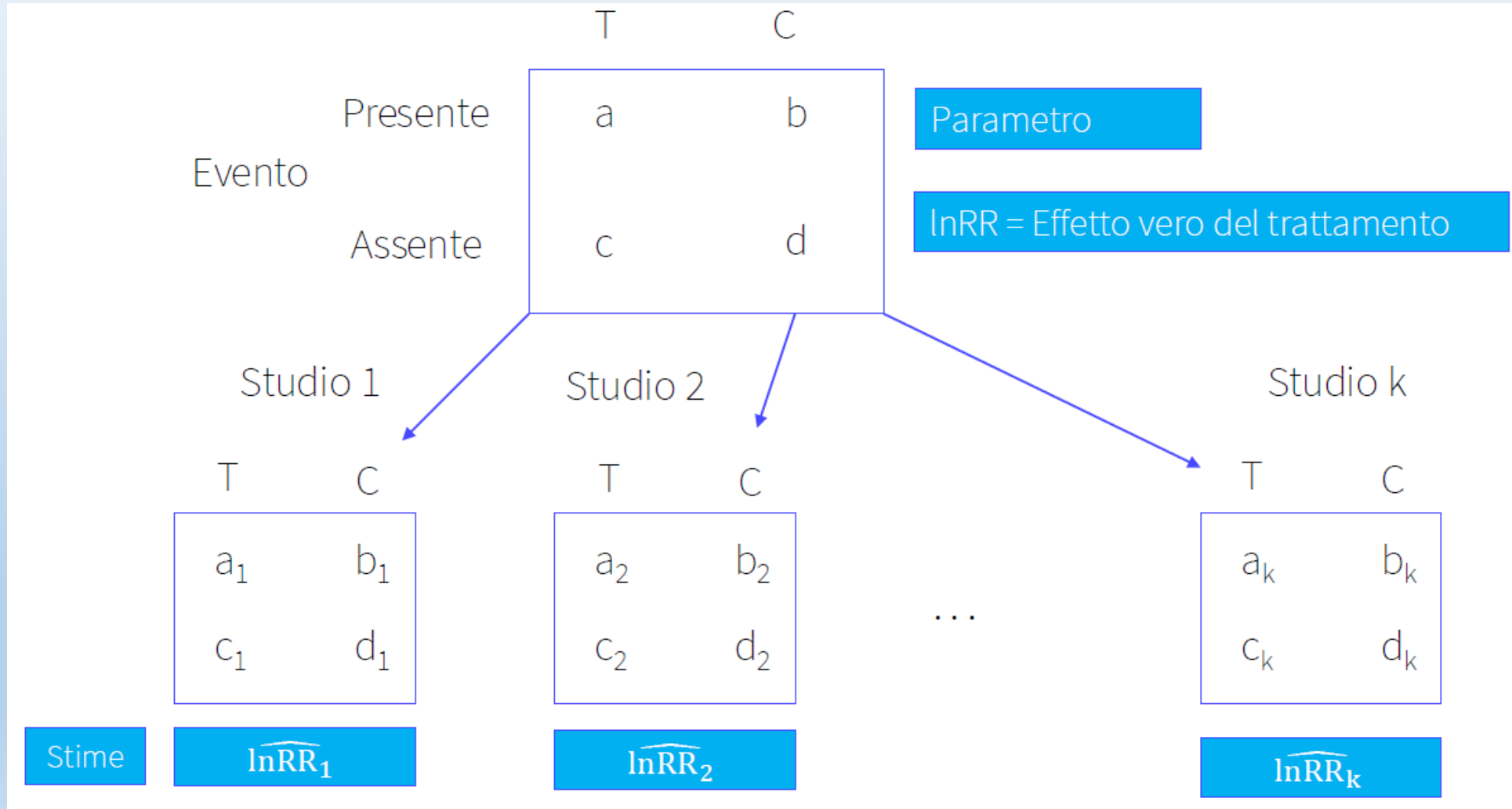
<i>Study</i>	<i>Participants</i>	<i>Study</i>	<i>Participants</i>
Bland 1976	Premature babies	Kanarek 1992	Premature babies
Brown 1988	Patients requiring TPN	Nilsson 1980	Colorectal cancer patients
Foley 1990	Critically ill patients	Rubin 1997	Patients requiring TPN
Golub 1994	Patients in surgical intensive care	Wojtysiak 1992	Adults requiring TPN

Different populations but statistical homogeneity

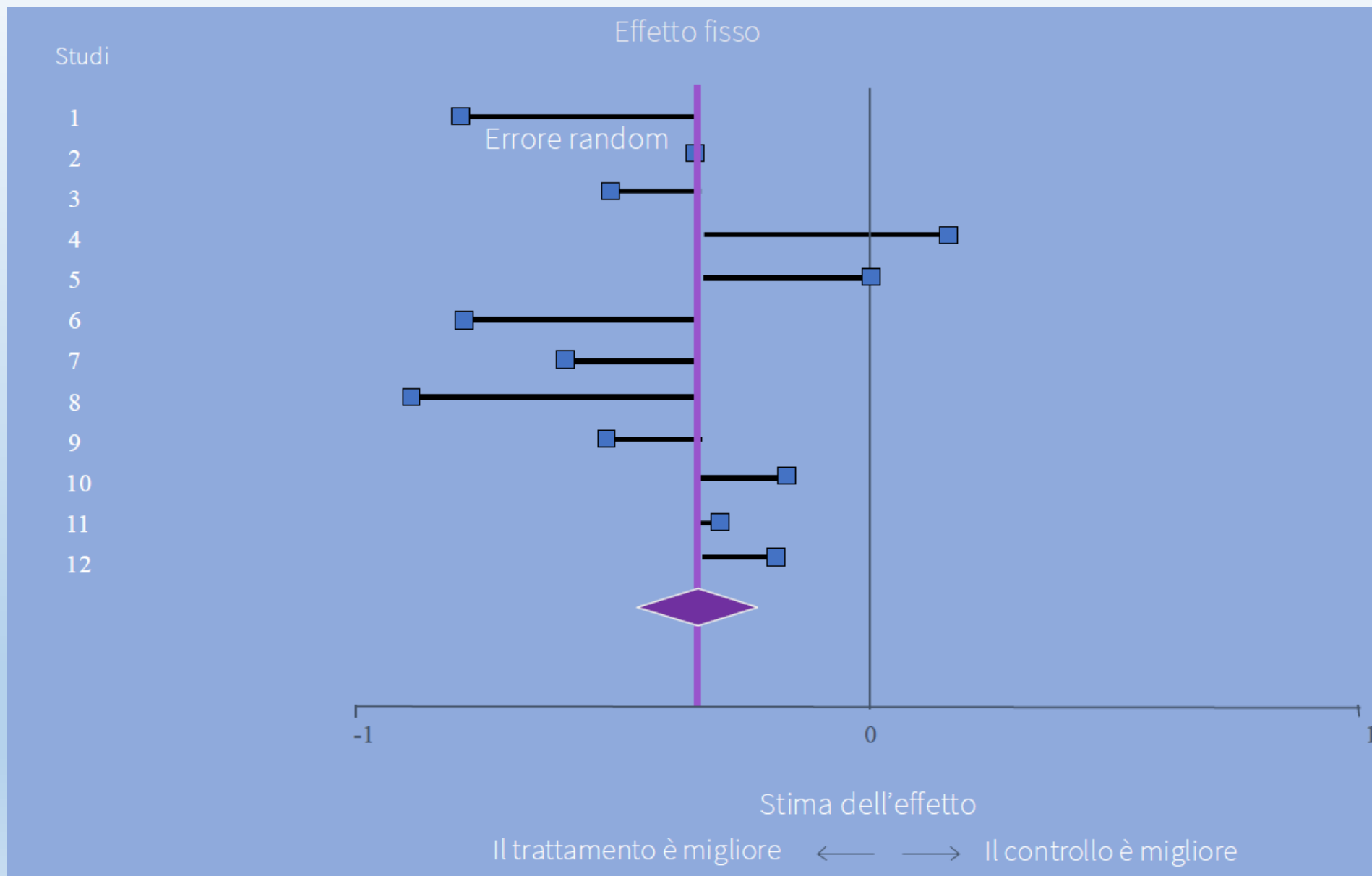


La Meta-analisi col Modello ad Effetti Fissi (Fixed/Common effects)

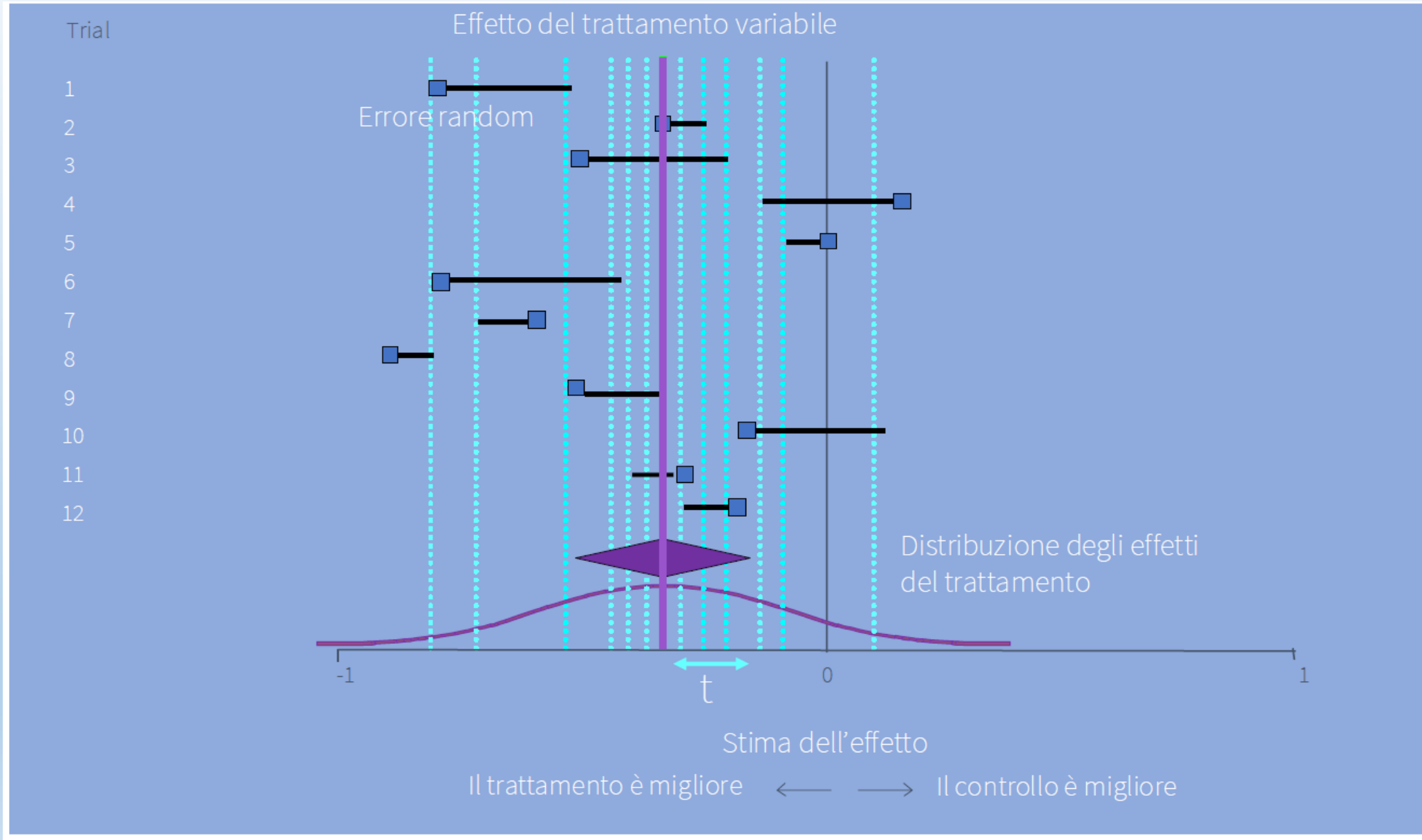
Il modello ad effetti fissi ipotizza che tutti gli studi forniscano una stima dello stesso parametro



La Meta-analisi col Modello ad Effetti Fissi



La Meta-analisi col Modello ad Effetti Casuali (random)



Differenze nel calcolo della stima tra i due modelli

Nel modello ad effetti casuali il peso viene influenzato anche dalla stima del **parametro di eterogeneità** (τ^2)

Tre steps:

- Calcolo del τ^2
- Ricalcolo dei pesi
- Calcolo la stima combinata (e la sua varianza) utilizzando i nuovi pesi

Total (95% CI)

1467

1484 100.0%

0.43 [0.35, 0.53]

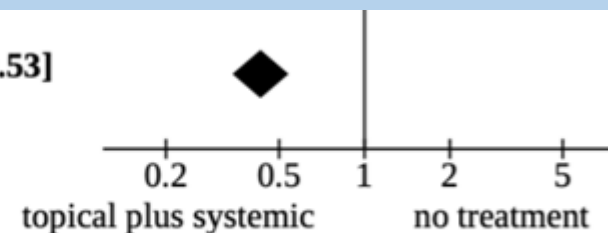
Total events:

294

619

Heterogeneity: $\tau^2 = 0.08$; $\chi^2 = 35.90$, $df = 16$ ($P = 0.003$); $I^2 = 55\%$

Test for overall effect: $Z = 7.80$ ($P < 0.00001$)



Il calcolo del peso nel modello ad effetti casuali

$$\tau^2 = \frac{Q - (k - 1)}{\sum_{i=1}^k p_i - \left(\frac{\sum_{i=1}^k p_i^2}{\sum_{i=1}^k p_i} \right)}$$

Se $Q < (k - 1) \Rightarrow \tau^2 = 0$

Nuova formula* del peso:

$$p_i' = \frac{1}{\text{var}(\ln(RR)_i) + \tau^2}$$

In assenza di eterogeneità statistica il peso del modello ad effetti casuali coincide con il peso del modello ad effetti fissi

* Esistono varie formule questa è quella più semplice

La «*Summary of Findings*»

E' una tabella che riporta per ogni confronto ed esito la sintesi dei risultati e le informazioni sulla *qualità* dell'evidenza

Per ogni outcome:

- **Grandezza** dell'effetto dell'intervento stimato (RR, OR...) e 95%CI
- **Numero** di pazienti (e di studi) inclusi per quell'esito
- Informazioni sulla *qualità* dell'evidenza (**GRADE**)

<https://www.gradeworkinggroup.org/>

La comunicazione
dei risultati a chi
deve prendere
decisioni cliniche :
*La summary of
findings*

Topical plus systemic prophylaxis versus placebo or no treatment in adults receiving mechanical ventilation for at least 48 hours

Patient or population: adults receiving mechanical ventilation for at least 48 hours

Setting: ICU in The Netherlands, France, Spain, Germany, USA, UK, Egypt, Ireland, Tunisia, South Africa, Austria, Greece, Switzerland, Belgium, Australia and New Zealand

Intervention: topical and systemic antibiotic prophylaxis

Comparison: no prophylaxis

Outcomes	Anticipated absolute effects ^a (95% CI)		Relative effect (95% CI)	No of participants (studies)	Certainty of the evidence (GRADE)
	Risk with no prophylaxis	Risk with topical plus systemic			
Overall mortality	Study population		RR 0.84 (0.73 to 0.96)	5290 (18 RCTs)	⊕⊕⊕⊕ HIGH
	303 per 1000	255 per 1000 (224 to 288)			
Respiratory tract infections	Study population		RR 0.43 (0.35 to 0.53)	2951 (17 RCTs)	⊕⊕⊕⊖ MODERATE ^a
	417 per 1000	179 per 1000 (146 to 221)			
Dropouts due to adverse events	Study population		RR 1.06 (0.30 to 3.76)	1287 (4 RCTs)	⊕⊕⊖⊖ LOW ^b
	6 per 1000	7 per 1000 (2 to 24)			
Gastrointestinal adverse events	Study population		RR 1.08 (0.57 to 2.04)	2637 (6 RCTs)	⊕⊕⊖⊖ LOW ^b
	44 per 1000	48 per 1000 (25 to 90)			
Allergic adverse events	Study population		RR 1.49 (0.09 to 25.33)	2981 (6 RCTs)	⊕⊕⊖⊖ LOW ^b
	0 per 1000	1 per 1000 (0 to 9)			

^aThe risk in the intervention group (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: confidence interval; ICU: intensive care unit; RCT: randomised controlled trial; RR: risk ratio

GRADE Working Group grades of evidence

High certainty: We are very confident that the true effect lies close to that of the estimate of the effect.

Moderate certainty: We are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.

Low certainty: Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.

Very low certainty: We have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

^aDowngraded 1 level for suspected publication bias, as indicated by asymmetry in the funnel plot.

^bDowngraded 2 levels due to very serious imprecision: sparse data.

