

# Pianificazione di uno studio caso-controllo multicentrico: un esempio

Chiara Peila<sup>1\*</sup>, Elena Spada<sup>1,2\*</sup>,  
Lorenzo Riboldi<sup>1</sup>, Roberto Bellù<sup>2,3</sup>,  
Giancarlo Gargano<sup>2,4</sup>, Alessandra Coscia<sup>1,2</sup>

\* Hanno contribuito in egual misura alla stesura del lavoro  
<sup>1</sup>AOU Città della Salute e della Scienza di Torino, Neonatologia Universitaria, Dipartimento di Scienze della Sanità Pubblica e Pediatriche, Torino

<sup>2</sup>Laboratorio della Conoscenza Carlo Corchia, Firenze

<sup>3</sup>Dipartimento area della donna e materno Infantile ASST, Lecco

<sup>4</sup>Dipartimento Materno Infantile Direttore Struttura Complessa Neonatologia AUSL-IRCCS, Reggio Emilia

Vogliamo pianificare uno studio osservazionale per rispondere al quesito clinico: la gemellarità è un fattore di rischio per patologie acute intestinali chirurgiche (PAIC) neonatali? Tra i diversi studi osservazionali, il caso-controllo multicentrico risulta il più adeguato. La popolazione target è costituita dai neonati pretermine (età gestazionale <37 settimane). L'endpoint primario è l'insorgenza di PAIC tra la nascita e le dimissioni: i casi sono i neonati con PAIC, i controlli i neonati senza PAIC. Raccoglieremo i dati dai registri ospedalieri, da cui saranno estratti tutti i casi e, in modo casuale, 4 controlli dello stesso centro per ogni caso. Useremo la regressione logistica per stimare l'odds ratio di avere PAIC nei gemelli rispetto ai singoli, dopo aver scelto il modello con un'attenta valutazione delle relazioni causali tra le variabili. Il limite principale di questo studio è la bassa potenza statistica, difficile da superare anche pianificando uno studio multicentrico.

*We planned an observational study to address the clinical question: is twinning a risk factor for acute neonatal surgical intestinal disorders (PAIC)? Multicentric case-control studies appear the most appropriate study design among observational ones. The target population consists of preterm infants (gestational age <37 weeks). The primary endpoint is PAIC occurrence from birth to hospital discharge: cases are neonates with PAIC, and controls are neonates without PAIC. Data will be collected from hospital records, extracting all cases and randomly selecting 4 controls from the same center for each case. We will use logistic regression to estimate the odds ratio of having PAIC in twins compared to singletons, after carefully evaluating the causal relationships among the involved variables. The main limitation of this study is the low statistical power, despite the multicentric planning of the study.*

## Introduzione

Le patologie acute intestinali chirurgiche (PAIC) a insorgenza neonatale sono fra le più gravi complicanze nei nati pretermine, e hanno una elevata mortalità e morbilità a lungo termine [1]. Fra queste, l'enterocolite necrotizzante (Necrotising Enterocolitis, NEC) e la perforazione intestinale spontanea (Spontaneous Intestinal Perforation, SIP) sono le più frequenti [2,3].

Uno studio [4] suggerisce una possibile associazione tra gemellarità e insorgenza di PAIC, ma è uno studio monocentrico basato su un campione di piccole dimensioni. Per questo motivo, abbiamo pianificato uno studio multicentrico con l'obiettivo di indagare l'associazione tra gemellarità (fattore di rischio) e PAIC (outcome) nei neonati pretermine. Scopo di questo articolo è descrivere la pianificazione di tale studio.

## Quesito clinico

La gemellarità è un fattore di rischio per PAIC?

## Disegno dello studio

Il quesito clinico richiede l'implementazione di uno studio osservazionale (non è possibile assegnare in via sperimentale la gemellarità).

Lo studio di coorte con raccolta dati prospettiva ha il vantaggio di consentire la scelta ad hoc delle informazioni da raccogliere e il loro formato (quali unità di misura, granularità/dettaglio delle informazioni, per esempio età in anni o per gruppi), migliorando l'uniformità della codifica delle variabili e riducendo il rischio di dati mancanti. Inoltre, il reclutamento prospettivo mitiga il rischio di bias di selezione (quando condizioni che hanno un effetto sull'outcome o sui fattori di rischio influenzano la probabilità di inclusione nello studio). La rarità delle PAIC (SIP: <2% nei pretermine e NEC: 7% a livello mondiale nei neonati ricoverati in TIN) [2,3] comporta un tempo di raccolta dati eccessivo per osservare un sufficiente numero di eventi ai fini dell'analisi statistica.

La raccolta dati retrospettiva permetterebbe di raccogliere le informazioni su un numero adeguato di eventi allargando la finestra temporale; tuttavia l'informatizzazione delle cartelle cliniche è recente e parziale, e le risorse necessarie per ottenere le informazioni dalle cartelle cliniche cartacee sono insufficienti.

Lo studio caso-controllo è uno studio osservazionale alternativo, in cui i neonati con PAIC (casi) vengono comparati con un gruppo di neonati senza PAIC che provengono dalla stessa popolazione (controlli). Le informazioni sono raccolte in maniera retrospettiva e si valuta se la distribuzione della gemellarità (fattore di rischio) è diversa nei due gruppi (casi vs controlli). Il vantaggio principale di questo tipo di studio è la possibilità di includere il maggior numero di casi possibile con le risorse disponibili; lo svantaggio è il rischio di bias di selezione e di informazione (quando condizioni che hanno un effetto sull'outcome o sui fattori di rischio influenzano la registrazione/raccolta del dato).

Per il nostro quesito clinico, lo studio caso-controllo multicentrico sembra il più appropriato ed efficiente, perché la partecipazione di più centri aumenta la possibilità di reclutare un numero di casi adeguato.

## Definizione dell'endpoint primario, di casi e controlli

L'endpoint primario è l'insorgenza di PAIC nell'intervallo di tempo tra la data di nascita e la data di dimissione. È definito caso il neonato con insorgenza di PAIC e controllo il neonato senza insorgenza di PAIC nell'intervallo definito.

## Definizione della popolazione target, criteri di inclusione e fonte dei dati

La popolazione di riferimento sono i nati pretermine (età gestazionale (EG) <37 settimane) dal 2010 negli ospedali italiani [Appendice]. Saranno pertanto eleggibili per lo studio tutti i nati pretermine da gravidanze singole o bigemine afferenti ai centri partecipanti dal 2010.

Le informazioni saranno estratte dai registri ospedalieri e saranno integrate attraverso le cartelle cliniche (elettroniche o cartacee).

### Reclutamento centri

In una prima fase il centro referente presenterà il protocollo dello studio ai centri interessati alla partecipazione (su base volontaria), i quali, a loro volta, potranno dividerlo con altri centri.

### Definizione delle variabili

Oltre a gemellarità e tipo di gemellarità (bicoriale, monocoriale), saranno raccolte le variabili associate con il fattore di rischio o con l'outcome e le variabili utili per la descrizione del campione. Raccoglieremo informazioni sulle caratteristiche della madre (es. età, peso/statura), gravidanza (es. EG, complicanze gestazionali, tipo di parto), complicanze e terapie perinatali (es. uso di surfattante, sofferenza perinatale), presenza di comorbidità, utilizzo di terapie postnatali e stato vitale del neonato alla fine del periodo di osservazione con eventuale causa di morte. Inoltre, data la natura multicentrica dello studio, per ogni centro raccoglieremo informazioni sulla presenza di protocolli nutrizionali, di gestione delle intolleranze nutrizionali, sull'uso di probiotici e sulla presenza della banca del latte. In caso di gemellarità, raccoglieremo le informazioni per entrambi i gemelli anche se solo uno appartiene al campione. Questa scelta permetterà di chiarire il ruolo di fattori di rischio condivisi (es. fattori genetici, ambiente uterino) sull'insorgenza di PAIC, in particolare nel caso in cui uno solo dei due gemelli abbia PAIC.

### Campionamento dei casi e controllo e potenza stimata

Raccoglieremo tutti i casi che soddisfano i criteri di inclusione e afferenti ai centri partecipanti. L'estrazione dei controlli, stratificata per centro, avverrà in maniera casuale con un rapporto di 4 controlli per ciascun caso.

La **Tabella 1** riporta la potenza dello studio, al variare della prevalenza di gemellarità tra il 5% e il 20% nei controlli, ipotizzando un OR di 1,5. L'ipotesi nulla del test statistico è OR=1, l'ipotesi alternativa è OR≠1; si assume una probabilità di errore di primo tipo α=0,05.

**Tabella 1. Potenza dello studio al variare della prevalenza di gemellarità nei controlli e al variare della numerosità di casi**

Prevalenza di gemelli nei controlli	N. casi	N. controlli	Potenza
0,05	200	800	0,187
0,05	250	1000	0,233
0,05	300	1200	0,278
0,05	350	1400	0,323
0,05	400	1600	0,367
0,10	200	800	0,341
0,10	250	1000	0,420
0,10	300	1200	0,494
0,10	350	1400	0,562
0,10	400	1600	0,623
0,15	200	800	0,460
0,15	250	1000	0,556
0,15	300	1200	0,640
0,15	350	1400	0,711
0,15	400	1600	0,771
0,20	200	800	0,548
0,20	250	1000	0,650
0,20	300	1200	0,734
0,20	350	1400	0,801
0,20	400	1600	0,852

Rapporto casi:controlli=1:4; OR ipotizzato=1,5; ipotesi nulla del test statistico: OR=1; ipotesi alternativa: OR≠1; probabilità di errore di primo tipo α=0,05.

### Gestione dei dati

I dati, pseudo-anonimizzati dal centro partecipante tramite un codice identificativo, saranno inseriti tramite piattaforma RedCap, che permette la realizzazione di form personalizzati per la raccolta dati e garantisce omogeneità nel formato delle variabili.

### Analisi statistica

Per la costruzione del modello statistico utilizzeremo un approccio di inferenza causale tramite l'impiego di Directed Acyclic Graphs (DAGs). Per l'implementazione, utilizzeremo il pacchetto R "dagitty" (<https://www.dagitty.net>) [5].

Negli studi caso-controllo la proporzione di casi e controlli inclusi nello studio è fissata dallo sperimentatore, che definisce l'intervallo temporale di osservazione e il numero di casi/controlli da includere. Per questo motivo, il rischio relativo

$$RR = \frac{p(\text{PAIC/gemello})}{p(\text{PAIC/singolo})}$$

ossia il rapporto tra la probabilità di avere PAIC nei gemelli e la probabilità di avere PAIC nei singoli) è una misura distorta dell'associazione fra outcome e fattore di rischio.

L'odds ratio (OR)

$$OR = \frac{n(\text{PAIC gemelli})/n(\text{non-PAIC gemelli})}{n(\text{PAIC singoli})/n(\text{non-PAIC singoli})}$$

(ossia il rapporto tra il numero di PAIC e non-PAIC nei gemelli e nei singoli) è la statistica, non distorta, spesso utilizzata per stimare tale associazione quando l'outcome è una variabile dicotomica (nel nostro caso: presenza/assenza di PAIC) (nota: si può ribaltare il ragionamento, vedi **Appendice**).

Stimeremo l'OR attraverso una regressione logistica in cui l'outcome è l'evento PAIC e le variabili esplicative sono la gemellarità (fattore di rischio) e variabili di aggiustamento. Per tenere conto della non-indipendenza tra gemelli appartenenti alla stessa coppia (es. stesso ambiente uterino, stesso DNA per gemelli monocoriali), eseguiremo un'analisi per misure ripetute.

Infine, se la numerosità sarà sufficiente il fattore di rischio verrà definito in tre classi: (a) singolo, (b) gemello dicoriale, (c) gemello monocoriale.

### Discussione

Lo studio caso-controllo ha il vantaggio di richiedere, in generale, poche risorse (anche economiche). Permette di avere informazioni sull'associazione tra due variabili in tempi inferiori rispetto a uno studio prospettivo, ed è utile a scopo esplorativo soprattutto nei casi in cui l'outcome di interesse è raro. Uno degli svantaggi è il rischio di ottenere dei risultati affetti da bias di selezione e di informazione. Nella nostra proposta di studio, abbiamo mitigato il bias di selezione tramite l'inclusione di tutti i casi referenti ai centri partecipante e il campionamento dei controlli dall'intera popolazione ed entrocentro; riteniamo il bias di informazione trascurabile perché le variabili di interesse sono raccolte al momento dell'evento nel registro ospedaliero.

Il limite difficilmente superabile di questo studio è che la potenza raggiunge la usuale soglia dell'80% solo in condizioni molto particolari [**Tabella 1**], con numerosità campionarie difficili da ottenere anche tramite uno studio multicentrico. Le conseguenze di una numerosità inferiore dovranno essere discusse in fase di reclutamento dei centri, per valutare se la numerosità raggiungibile corrisponda ad una potenza dello studio accettabile.

### Ringraziamenti

Ringraziamo Laura Viviani per l'attenta revisione del testo e dei suoi contenuti.

## Bibliografia

1. Horbar JD, Carpenter JH, Badger GJ, et al. Mortality and neonatal morbidity among infants 501 to 1500 grams from 2000 to 2009. *Pediatrics*. 2012 Jun;129(6):1019-26.
2. Alsaied A, Islam N, Thalib L. Global incidence of Necrotizing Enterocolitis: a systematic review and Meta-analysis. *BMC Pediatr*. 2020 Jul 13;20(1):344.
3. Vongbhavit K, Underwood MA. Intestinal perforation in the premature infant. *J Neonatal Perinatal Med*. 2017;10(3):281-89.
4. Peila C, Spada E, Riboldi L, et al. Twinning as a risk factor for neonatal acute intestinal diseases: a case-control study. *Front Pediatr*. 2023 Dec 14;11:1308538.
5. Textor J, van der Zander B, Gilthorpe MS, et al. Robust causal inference using directed acyclic graphs: the R package 'dagitty'. *Int J Epidemiol*. 2016 Dec 1;45(6):1887-94.

## APPENDICE

### Definizione della popolazione target

Uno dei punti critici della pianificazione di uno studio caso-controllo è la scelta dei controlli. La loro selezione dovrebbe essere fatta seguendo degli assunti schematizzabili in tre principi che elenchiamo di seguito.

Un controllo:

- deve essere estratto dalla stessa popolazione da cui derivano i casi (popolazione target);
- deve essere selezionato indipendentemente dall'esposizione al fattore di rischio (nel nostro studio, indipendentemente dall'essere o meno un gemello);
- ha la stessa probabilità di essere estratto di qualunque neonato – senza outcome – appartenente alla popolazione target, e tale probabilità deve essere indipendente dall'eventuale selezione di altri neonati (la selezione di ciascun neonato non altera la probabilità di estrarre nessuno degli altri neonati).

Per ottenere un campione in cui casi e controlli hanno distribuzione simile (sono bilanciati) per una definita variabile confondente, è possibile estrarre i controlli "appaiano" con i casi per quella variabile, cioè scegliendoli in base al valore della variabile di appaiamento che li caratterizza. Questa tecnica comporta – spesso, ma non sempre – un aumento dell'efficienza statistica, in cambio, però, della rinuncia della condizione [3]: di fatto si inserisce un bias di selezione perché la probabilità di un controllo di essere estratto dipende dal valore della sua variabile di appaiamento. In fase di analisi occorre tenere conto di tale bias, inserendo nel modello la variabile di appaiamento come confondente o facendo un'analisi stratificata.

Durante la pianificazione del nostro studio, avevamo inizialmente pensato di appaiare i controlli con i casi per EG (principale fattore di rischio di PAIC). Tuttavia, la valutazione delle relazioni causali tra le variabili ha evidenziato che l'EG è un mediatore tra la gemellarità e il rischio di PAIC [Figura 1]. In quanto tale, non è possibile appaiare per EG: l'appaiamento comporta l'inserimento di EG come covariata nel modello di analisi per correggere il bias di selezione (la linea tratteggiata).

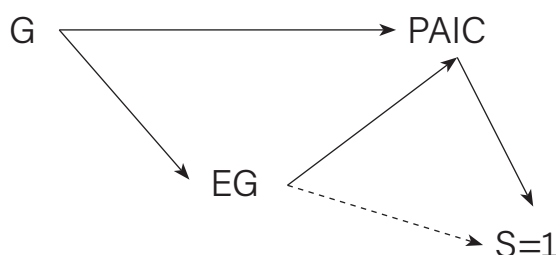


Figura 1. Relazioni casuali tra gemellarità (G), PAIC, EG, e appartenenza al campione (S=1). La linea tratteggiata mostra la relazione generata dall'appaiamento per EG.

ta della Figura 1); la correzione interrompe il percorso G, EG, PAIC, che stima l'effetto della gemellarità mediato dall'EG; il risultato è un modello sovracorretto [1A, 2A].

Sempre durante la pianificazione è emersa un'altra riflessione: i neonati a termine (EG ≥ 37 settimane) hanno sia probabilità di esposizione al fattore di rischio (gemellarità), sia probabilità di PAIC quasi nulle, e rappresentano quasi il 94% di tutti i neonati [3A]. Ne consegue che i casi provengono da una popolazione prevalentemente pretermine a cui appartengono la maggior parte dei gemelli. Per non violare la condizione [1] per la scelta dei controlli, abbiamo quindi deciso di considerare come popolazione target i soli nati pretermine.

### Stima dell'odds ratio (OR) e sua reversibilità

Gli studi caso-controllo sono un tipo particolare di studi osservazionali, in cui la proporzione tra casi (evento=sì) e controlli (evento=no) è definita dallo sperimentatore; per cui la variabile casuale è la presenza/assenza del fattore di rischio nei casi rispetto ai controlli. Al contrario, negli altri studi la variabile casuale è evento sì/evento no negli esposti al fattore di rischio rispetto ai non esposti. Seguendo questa logica, mentre in generale si stima la proporzione di eventi in funzione del fattore di rischio, negli studi caso-controllo l'analisi dovrebbe essere eseguita stimando la proporzione del fattore di rischio in funzione dell'evento.

Di seguito mostreremo che i due procedimenti portano allo stesso risultato, differendo solo per l'approccio filosofico e per l'interpretazione.

Per semplicità raggruppiamo le osservazioni in una tabella 2x2 [Tabella 2].

Tabella 2. Schema delle osservazioni in un campione di uno studio caso-controllo. Dimensione campionaria dei casi=a+c; dimensione campionaria dei controlli=b+d

Esposto al fattore di rischio	Outcome	
	Caso Sì (PAIC)	Controllo No (non-PAIC)
Sì (gemello)	a	b
No (singolo)	c	d

Primo scenario: valutazione dell'associazione tra evento e fattore di rischio come odds di essere gemello nei casi vs odds di essere gemello nei controlli.

$$\begin{aligned} \text{Ipotesi nulla } H_0: \text{OR}=1 \\ \text{Ipotesi alternativa } H_1: \text{OR} \neq 1 \end{aligned}$$

dove:

$$\text{OR} = \frac{\text{Odds}(\text{gemellarità} | \text{caso})}{\text{Odds}(\text{gemellarità} | \text{controllo})} = \frac{\frac{a}{c}}{\frac{b}{d}} = \frac{a}{c} \times \frac{d}{b} = \frac{a \times d}{b \times c}$$

Secondo scenario: valutazione dell'associazione tra fattore di rischio ed evento come odds di essere caso nei gemelli vs odds di essere caso nei nati singoli.

$$\begin{aligned} \text{Ipotesi nulla } H_0: \text{OR}=1 \\ \text{Ipotesi alternativa } H_1: \text{OR} \neq 1 \end{aligned}$$

dove:

$$\text{OR} = \frac{\text{Odds}(\text{caso} | \text{gemello})}{\text{Odds}(\text{caso} | \text{singolo})} = \frac{\frac{a}{b}}{\frac{c}{d}} = \frac{a}{b} \times \frac{d}{c} = \frac{a \times d}{b \times c}$$

Numericamente i due OR coincidono. Cambia l'interpretazione: OR > 1 nel primo scenario indica che l'odds di essere gemello (avere il fattore di rischio) nei neonati con PAIC (nella popo-

lazione dei casi) è maggiore dell'odds di essere gemello (avere il fattore di rischio) nei neonati senza PAIC (nella popolazione dei controlli); mentre  $OR > 1$  nel secondo scenario indica che l'odds di avere PAIC (essere caso) nei nati gemelli (nella popolazione con il fattore di rischio) è maggiore dell'odds di avere PAIC (essere caso) nei nati singoli (nella popolazione senza fattore di rischio).

Per semplicità di interpretazione preferiamo utilizzare l'approccio mostrato nel secondo scenario. ■

### Bibliografia

1. Mansournia MA, Hernán MA, Greenland S. Matched designs and causal diagrams. *Int J Epidemiol.* 2013 Jun;42(3):860-9.
2. Buzzetti R, Gagliardi L. Il sesso femminile espone il neonato con fibrosi cistica a un esito peggiore? Un processo in tre atti. *Quaderni acp* 2024;3:105-7.
3. Rapporto sull'evento nascita in Italia (CeDAP) - anno 2022. <https://www.epicentro.iss.it/materno/cedap-2022>.

*elenaspada.bios@gmail.com*

## blister Call for abstracts - Giornate di Epidemiologia 2025

Le *Giornate di Epidemiologia "Carlo Corchia" 2025*, organizzate dal Laboratorio della Conoscenza, rappresentano un'occasione unica di incontro e confronto per i professionisti impegnati nel campo della salute perinatale e pediatrica. Quest'anno, l'evento residenziale si terrà dal **12 al 14 giugno 2025** nella suggestiva cornice di **Fiesole (FI)**, presso il Centro Studi Ricerca e Formazione CISL.

Seguendo la tradizione di Carlo Corchia, promotore di un approccio innovativo all'epidemiologia pediatrica, queste giornate sono progettate per favorire uno scambio orizzontale di idee e metodologie. L'obiettivo non è solo trasmettere conoscenze, ma costruire cultura attraverso un dialogo attivo e paritario tra i partecipanti.

### Un invito a partecipare attivamente

La call for abstracts è ora aperta! Professionisti e ricercatori sono invitati a inviare proposte di comunicazione entro il **20 febbraio 2025**. Gli abstract saranno pubblicati in un supplemento di *Quaderni acp*, dando visibilità alle ricerche e ai progetti presentati.

### Cosa cerchiamo?

Accogliamo con entusiasmo contributi su:

- Studi epidemiologici conclusi o in corso;
- Progetti di ricerca innovativi;
- Riflessioni metodologiche su temi specifici.

Gli interventi avranno uno spazio dedicato di **10 minuti di presentazione**, seguiti da **15 minuti di discussione**, per promuovere un dialogo aperto e costruttivo.

### Come inviare un abstract?

Gli abstract devono essere strutturati e includere sezioni come *background*, *obiettivo*, *metodi*, *risultati* e *conclusioni*, con un massimo di **300 parole**. Eventuali tabelle o figure possono essere allegate, ma non saranno pubblicate. Maggiori dettagli sono disponibili sul sito ufficiale dell'evento:

[www.labcorchia.it](http://www.labcorchia.it)

### Un contesto unico per un dialogo interdisciplinare

Le *Giornate di Epidemiologia 2025* riuniranno professionisti da diverse discipline (neonatologi, pediatri, ostetrici, neuropsichiatri, genetisti, epidemiologi e infermieri) per affrontare insieme le sfide della salute perinatale. La formula residenziale dell'evento, limitata a soli **60 partecipanti**, favorisce un'interazione autentica e approfondita, anche al di fuori delle sessioni seminariali.

### Unisciti a noi a Fiesole!

Invitiamo tutti i lettori dei *Quaderni acp* a contribuire a rendere queste giornate un momento di scambio e crescita per l'intera comunità scientifica. Che tu sia un professionista esperto o un giovane ricercatore in formazione, questa è un'opportunità unica per presentare il tuo lavoro e costruire collaborazioni preziose.

### Scadenze importanti

- Invio degli abstract: **20 febbraio 2025**
- Comunicazione di accettazione: **15 marzo 2025**

Ci auguriamo di vedervi numerosi a Fiesole, per continuare a costruire insieme il futuro della salute perinatale e pediatrica.

*Il Direttivo del Laboratorio della Conoscenza "Carlo Corchia"*