

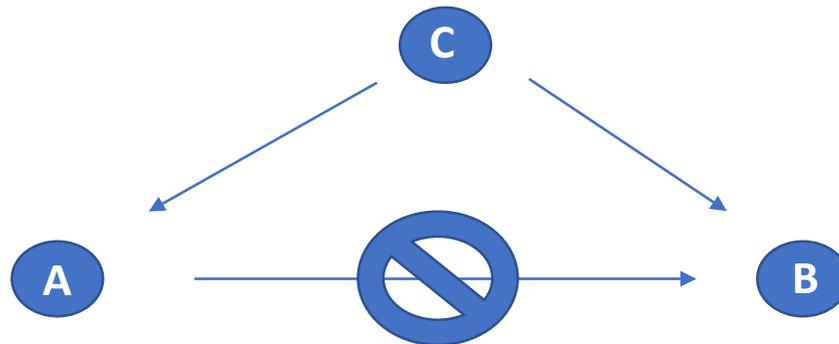
# Regressione logistica

Il termine **associazione** è largamente usato nella letteratura scientifica ed esprime la relazione che esiste tra due variabili

Per studiare l'associazione tra due variabili bisogna pensare almeno a due livelli di analisi:

1. La relazione tra le due variabili in studio può essere spiegata da una terza variabile  
(la terza variabile deve essere inclusa nello studio!!!)
2. Utilizzare i metodi statistici appropriati per studiare la relazione

1. La relazione tra le due variabili in studio può essere spiegata da una terza variabile  
(la terza variabile deve essere inclusa nello studio!!!)



630

THE NEW ENGLAND JOURNAL OF MEDICINE

March 12, 1981

### COFFEE AND CANCER OF THE PANCREAS

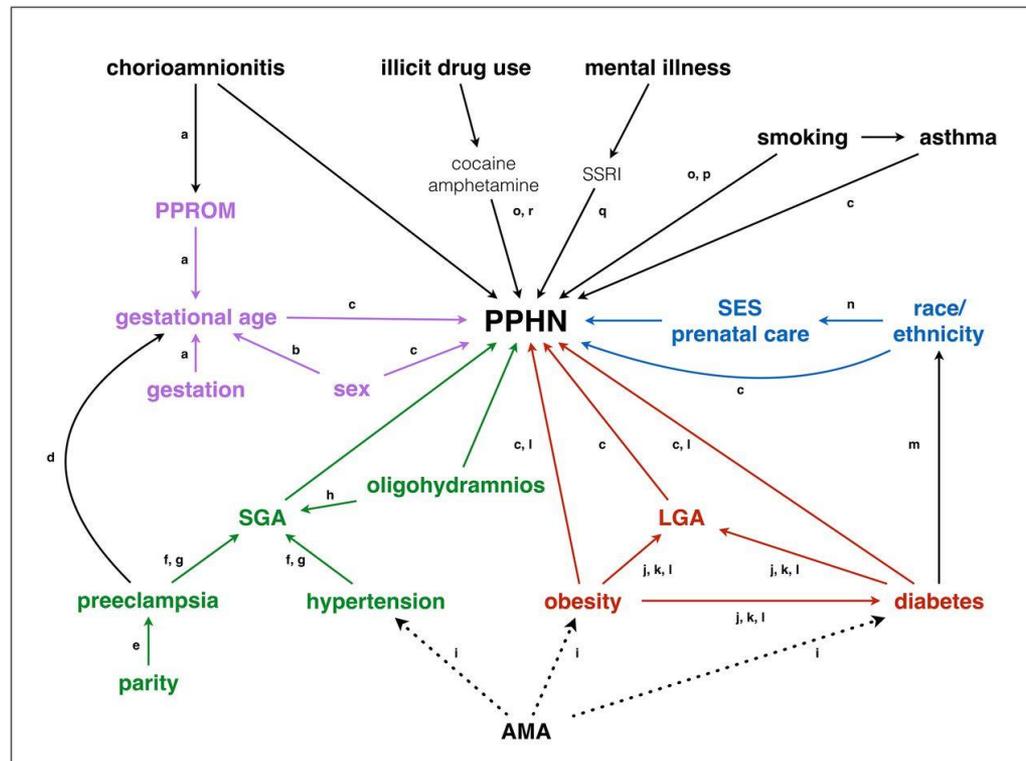
BRIAN MACMAHON, M.D., STELLA YEN, M.D., DIMITRIOS TRICHOPOULOS, M.D., KENNETH WARREN, M.D.,  
AND GEORGE NARDI, M.D.

# 1. La relazione tra le due variabili in studio può essere spiegata da una terza variabile (la terza variabile deve essere inclusa nello studio!!!)

Pediatrics  
January 2017, VOLUME 139 / ISSUE 1  
From the American Academy of Pediatrics  
Article

## Persistent Pulmonary Hypertension of the Newborn in Late Preterm and Term Infants in California

Martina A. Steurer, Laura L. Jelliffe-Pawlowski, Rebecca J. Baer, J. Colin Partridge, Elizabeth E. Rogers, Roberta L. Keller



## 2. Utilizzare i metodi statistici appropriati per studiare la relazione

### Response variable

Categorical

Quantitative

Categorical

C-C

C-Q

$Q \rightarrow C$

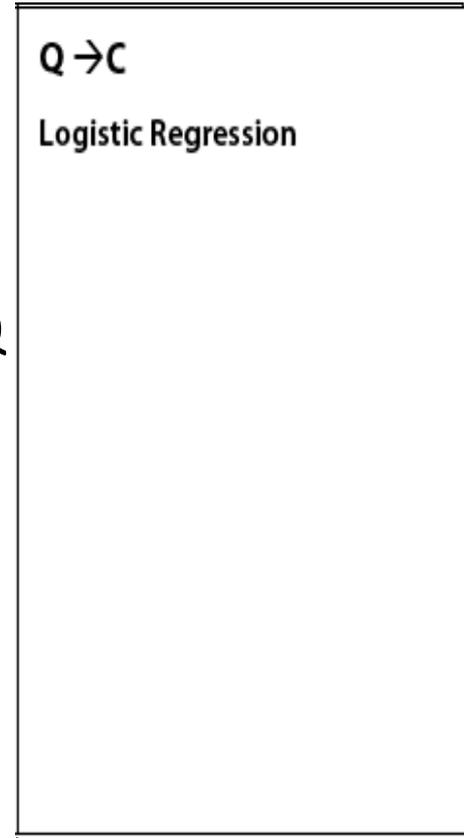
Logistic Regression

### Explanatory variable

Quantitative

Q-C

Q-Q



# Regressione logistica

Variabile di outcome dicotomica  
(es. malati / non malati, caso / controllo)

Variabili esplicative categoriche o continue

La regressione logistica è usata per determinare quali variabili esplicative e in che misura predicono l'outcome, quando l'outcome è una variabile dicotomica

## Obiettivi della regressione logistica

1. Costruire un modello di predizione, in grado di prevedere l'outcome di interesse conoscendo altre informazioni dei soggetti inclusi nel campione (variabili esplicative)
2. Valutare l'effetto di una variabile esplicativa (es. un fattore di rischio) su un outcome di interesse (dicotomico)

Modello molto versatile, utilizzato sia negli studi trasversali, caso controllo e studi sperimentali nei quali la variabile di outcome è dicotomica (caratteristica frequente in epidemiologia)

## regressione logistica

Si vuole modellare la relazione tra una variabile dipendente dicotomica e un insieme di altre variabili (variabili esplicative) che si ritiene influenzino la variabile dipendente

La variabile dipendente è dicotomica (es. presenza / assenza di malattia)

$$\text{Logit}(P) = \ln(P/1-P) = a + b_1x_1 + b_2x_2 + b_ix_i + b_nx_n$$

P= probabilità che si verifichi un evento

The spreadsheet **asthma.sav** contains data from a random cross-sectional sample of 2464 children aged 8 to 10 years in which the exposure of allergy to housedust mites (HDM), the exposure to respiratory infection in early life, the characteristic gender and the presence of the disease (asthma) were measured in all children.

# Tabella di contingenza: asma e genere

**Tavola di contingenza Gender \* Diagnosed asthma**

			Diagnosed asthma		Totale
			No	Yes	
Gender	Female	Conteggio	965	223	1188
		Conteggio atteso	936,8	251,2	1188,0
		% entro Gender	81,2%	18,8%	100,0%
	Male	Conteggio	978	298	1276
		Conteggio atteso	1006,2	269,8	1276,0
		% entro Gender	76,6%	23,4%	100,0%
Totale	Conteggio	1943	521	2464	
	Conteggio atteso	1943,0	521,0	2464,0	
	% entro Gender	78,9%	21,1%	100,0%	

## Chi-quadrato

	Valore	df	Sig. asint. (2 vie)	Sig. esatta (2 vie)	Sig. esatta (1 via)
Chi-quadrato di Pearson	7,751 <sup>a</sup>	1	,005		

## Stima di rischio

	Valore	Intervallo di confidenza 95%	
		Inferiore	Superiore
Rapporto odd per Gender (Female / Male)	1,319	1,085	1,602

# Tabella di contingenza: asma e allergia agli acari

Tavola di contingenza Allergy to HDM ^ Diagnosed asthma

			Diagnosed asthma		Totale
			No	Yes	
Allergy to HDM	No	Conteggio	1414	125	1539
		Conteggio atteso	1213,6	325,4	1539,0
		% entro Allergy to HDM	91,9%	8,1%	100,0%
	Yes	Conteggio	529	396	925
		Conteggio atteso	729,4	195,6	925,0
		% entro Allergy to HDM	57,2%	42,8%	100,0%
Totale	Conteggio	1943	521	2464	
	Conteggio atteso	1943,0	521,0	2464,0	
	% entro Allergy to HDM	78,9%	21,1%	100,0%	

## Chi-quadrato

	Valore	df	Sig. asint. (2 vie)	Sig. esatta (2 vie)	Sig. esatta (1 via)
Chi-quadrato di Pearson	416,951 <sup>a</sup>	1	,000		

## Stima di rischio

	Valore	Intervallo di confidenza 95%	
		Inferiore	Superiore
Rapporto odd per Allergy to HDM (No / Yes)	8,468	6,765	10,600

# Tabella di contingenza: asma e infezioni respiratorie

**Tavola di contingenza Early infection \* Diagnosed asthma**

			Diagnosed asthma		Totale
			No	Yes	
Early infection	No	Conteggio	1622	399	2021
		Conteggio atteso	1593,7	427,3	2021,0
		% entro Early infection	80,3%	19,7%	100,0%
	Yes	Conteggio	321	122	443
		Conteggio atteso	349,3	93,7	443,0
		% entro Early infection	72,5%	27,5%	100,0%
Totale	Conteggio	1943	521	2464	
	Conteggio atteso	1943,0	521,0	2464,0	
	% entro Early infection	78,9%	21,1%	100,0%	

## Chi-quadrato

	Valore	df	Sig. asint. (2 vie)	Sig. esatta (2 vie)	Sig. esatta (1 via)
Chi-quadrato di Pearson	13,247 <sup>a</sup>	1	,000		

## Stima di rischio

	Valore	Intervallo di confidenza 95%	
		Inferiore	Superiore
Rapporto odd per Early infection (No / Yes)	1,545	1,221	1,955

# Modello logistico: asma variabile dipendente + covariate (allergia, infezioni e genere)

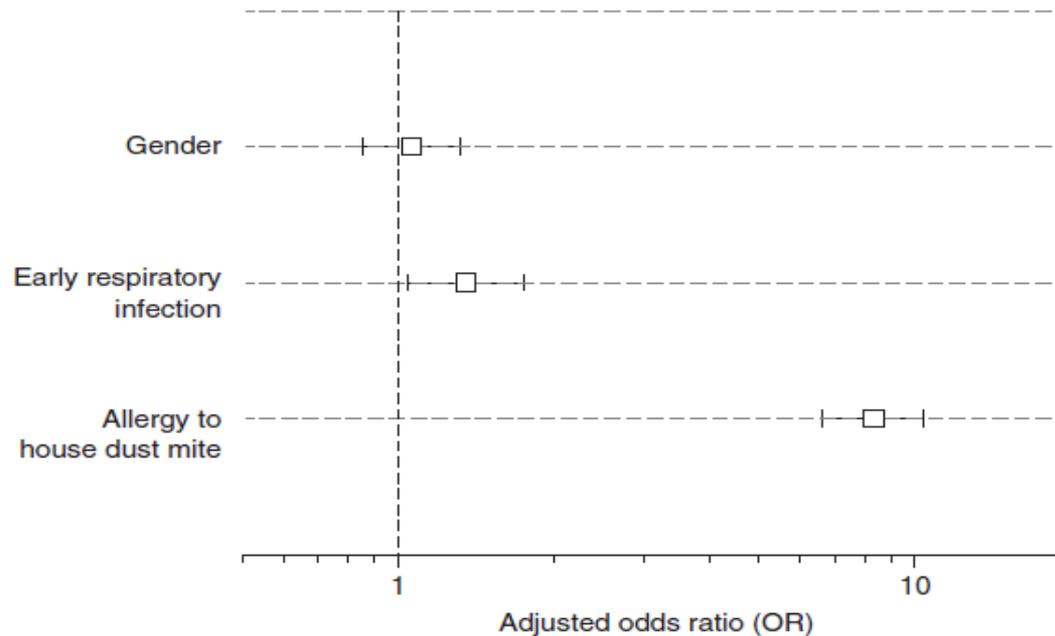
Variabili nell'equazione

	B	E.S.	Wald	df	Sig.	Exp(B)	95% CI per EXP(B)	
							Inferiore	Superiore
Passo 1 <sup>a</sup>								
hdm	2,118	,115	338,103	1	,000	8,313	6,633	10,419
infect	,302	,133	5,155	1	,023	1,353	1,042	1,756
gender	,058	,110	,274	1	,600	1,059	,854	1,314
Costante	-4,985	,289	297,409	1	,000	,007		

Or  
non aggiustati

8,47  
1,55  
1,32

a. Variabili immesse al passo 1: hdm, infect, gender.



Independent risk factors for diagnosed asthma in children.