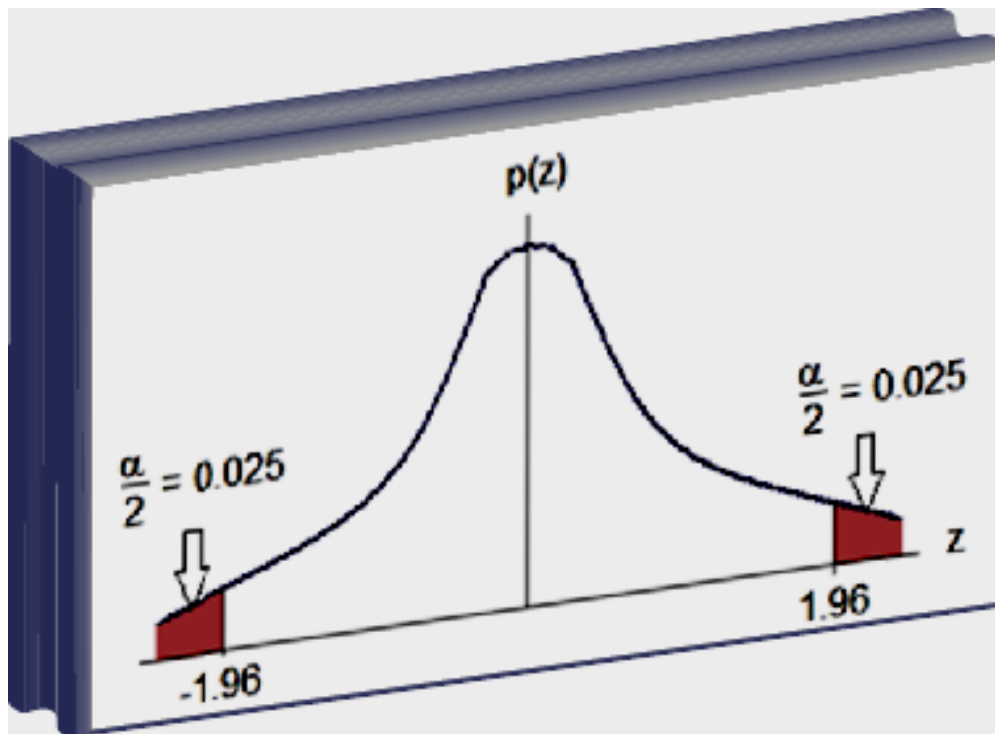


## METODI STATISTICI PER IL CONTROLLO DEI FALSI POSITIVI IN EPIDEMIOLOGIA AMBIENTALE

K. DI BIAGIO, M. BALDINI, S. BARTOLACCI, H'J "GIMEONI, M. MARIOTTINI  
AGENZIA REGIONALE PER LA PROTEZIONE AMBIENTALE DELLE MARCHE



Dicembre 2013

## SOMMARIO

---

SOMMARIO .....	2
ABSTRACT .....	3
1. INTRODUZIONE .....	4
1.1 Obiettivi del presente studio.....	5
2. MATERIALI E METODI .....	6
3. RISULTATI.....	9
4. DISCUSSIONE E CONCLUSIONI.....	12
BIBLIOGRAFIA.....	14

---

## ABSTRACT

---

### **Obiettivi**

Nel contesto dell'epidemiologia ambientale di tipo esplorativo si è interessati ad investigare più ipotesi contemporaneamente, come l'indagine dei rischi di diverse patologie sui residenti di uno o più comuni. La verifica di più ipotesi scientifiche sullo stesso campione di dati espone ad un elevato rischio di avere risultati statisticamente significativi (s.s.) solo per effetto del caso (falsi positivi), compromettendo la misurazione dell'evidenza con la conseguente produzione di risultati inconsistenti.

L'obiettivo di questo studio è proseguire la sorveglianza dello stato di salute dei residenti del comune di Potenza Picena, e in generale della popolazione marchigiana, attraverso il monitoraggio di una delle patologie con maggior tasso di mortalità nella regione Marche, utilizzando e comparando le principali metodologie statistiche per il controllo del rischio di avere dei risultati falsi positivi.

### **Dati e metodi**

Sono stati analizzati i soggetti residenti nei comuni della regione Marche nel periodo 2006-2009 che presentavano nella causa iniziale di decesso la malattia ischemica del cuore (ICD X I20-I25); per ogni comune e genere sono stati calcolati i rapporti di mortalità standardizzati per classi di età quinquennali e indice di deprivazione e i valori p aggiustati per la molteplicità dei test, secondo il criterio di Bonferroni e secondo l'approccio del tasso di false scoperte positivo (valore q).

### **Risultati**

Su un totale di 478 ipotesi valutate sul campione in studio, il 7% (n. 34) mostrava un rischio di decesso statisticamente significativo al livello di confidenza del 5%; dopo l'aggiustamento secondo il criterio di Bonferroni le ipotesi risultate s.s. (in eccesso o in difetto) sono diminuite drasticamente, passando da 34 a 2, mentre secondo l'approccio del valore q le ipotesi s.s. sono risultate 3. L'eccesso di mortalità s.s. secondo entrambi i metodi di correzione è stato rilevato nel comune di Monte Vidon Corrado con un SMR% di 488.

### **Conclusioni**

I risultati delle analisi esplorative corrette per la molteplicità dei test sono utili come elemento aggiuntivo per la formulazione di ipotesi epidemiologiche da valutare in un successivo studio analitico. La correzione per la molteplicità delle ipotesi possono, invece, essere necessarie nelle indagini epidemiologiche previste dal Decreto 24 aprile 2013 dei Ministeri della Salute e dell'Ambiente (VDS) poiché permetterebbe di concentrare l'attenzione solo sugli eventi che presentano maggiore indicazione d'impatto.

### **KEYWORDS :**

***valore-p, confronti multipli, valore-q***

## 1. INTRODUZIONE

Uno degli obiettivi dell'epidemiologia è quello di indagare la presenza di un eccesso di rischio sanitario attraverso un criterio scientifico che consiste, in prima fase, nella teorizzazione di un'ipotesi epidemiologica e, in seconda fase, nella verifica della stessa su base empirica. In accordo al principio di falsificabilità introdotto da Popper, occorre formalizzare il procedimento scientifico con la formulazione di un'ipotesi sul non verificarsi di un effetto, indicata come ipotesi nulla ( $H_0$ ), e sul suo verificarsi, indicata come ipotesi alternativa ( $H_A$ ). Sulla base dei dati empirici di un campione rappresentativo della popolazione in studio, l'obiettivo della ricerca è quello di falsificare l'ipotesi nulla relativa alla mancanza dell'effetto così da ritenere verosimile quanto formulato nell'ipotesi alternativa (presenza dell'effetto).

Disponendo di dati campionari, l'ipotesi scientifica viene quindi espressa e valutata in termini probabilistici e, proprio perché la decisione sul quesito scientifico si basa su dati che rappresentano solo una parte dell'intera popolazione, rimane la reale possibilità di giungere a conclusioni errate.

Definita l'ipotesi nulla e l'ipotesi alternativa relative ad un parametro della popolazione, si stabilisce una regola che discrimina i risultati che conducono a rifiutare o a non rifiutare l'ipotesi nulla; secondo lo schema riportato in tabella 1 due sono i potenziali errori che si possono commettere:

- rifiutare l'ipotesi nulla quando è vera (errore di prima specie indicato con il simbolo  $\alpha$ ; falso positivo);
- non rifiutare l'ipotesi nulla quando è falsa (errore di seconda specie indicato con il simbolo  $\beta$ ; falso negativo).

**Tabella 1. Errore di prima e seconda specie in un test d'ipotesi.**

Test di ipotesi	Decisione	
	<i>Non rifiuto <math>H_0</math></i>	<i>Rifiuto <math>H_0</math></i>
<i>Stato reale</i> $H_0$ è vera	Decisione corretta	Errore di prima specie $\alpha$
$H_0$ è falsa	Errore di seconda specie $\beta$	Decisione corretta

Prima di procedere con la verifica empirica dell'ipotesi si stabilisce qual è il rischio di commettere un errore falso-positivo che si è disposti ad accettare, indicato come livello di significatività ( $\alpha$ ), e che convenzionalmente viene fissato al valore di 0,05. Fissato un campione di dati di una determinata numerosità, la procedura di verifica delle ipotesi non permette di rendere minimi simultaneamente gli errori di prima e di seconda specie; pertanto, dopo aver fissato a priori la soglia di accettabilità per l'errore di prima specie, la procedura determina il punto di demarcazione dei risultati che conducono ad accettare l'ipotesi nulla minimizzando però l'errore di seconda specie. Il risultato del test statistico, proveniente dai dati raccolti, fornisce la probabilità che l'effetto rilevato dai dati (una differenza rispetto al valore del parametro previsto nell'ipotesi nulla) possa essere osservato solo per caso (valore  $p$  o  $p$ -value).

Un valore  $p$  molto piccolo indica che la differenza osservata non è imputabile all'effetto del caso ma può essere verosimilmente considerata come reale; operativamente il valore  $p$  è comparato con il livello di significatività, ad esempio dello 0,05, e quando risulta inferiore ad esso si asserisce che l'effetto mostrato dai dati è statisticamente significativo con un rischio inferiore al 5% di sostenere che l'effetto ipotizzato è presente quando in realtà non lo è (errore falso positivo).

Dato un campione di dati, in generale l'indagine è focalizzata su una singola ipotesi di ricerca; tuttavia, nel contesto dell'epidemiologia ambientale di tipo esplorativo si è interessati a investigare più ipotesi contemporaneamente, come ad esempio l'indagine dei rischi di svariate tipologie di malattie su una popolazione di un determinato comune, oppure su popolazioni diverse, come possono essere quelle di tutti i comuni di una regione. La verifica di diverse ipotesi scientifiche sullo stesso campione di dati espone ad un elevato rischio di avere risultati statisticamente significativi solo per effetto del caso (incremento della probabilità di commettere un errore falso-positivo), compromettendo la misurazione dell'evidenza sperimentale con la conseguente produzione di risultati inconsistenti.

Diverse tecniche statistiche sono state formulate al fine di correggere il valore  $p$  per la molteplicità dei confronti multipli effettuati in un'indagine, a partire dall'aggiustamento secondo il criterio di Bonferroni fino al valore  $q$  ( $q$ -value) recentemente introdotto da Storey J.D. (1).

## **1.1 OBIETTIVI DEL PRESENTE STUDIO**

Nell'ambito dell'epidemiologia il quesito scientifico e il disegno dello studio determinano la scelta dei metodi statistici adeguati a rappresentare la struttura del fenomeno osservato. La verifica di una molteplicità di ipotesi su uno stesso campione di dati, pone il problema, da un punto di vista probabilistico, di riconoscere i rischi sanitari reali da quelli che risultano tali solo per effetto del caso.

L'obiettivo di questo studio è proseguire la sorveglianza dello stato di salute dei residenti del comune di Potenza Picena e in generale della popolazione marchigiana, monitorando l'andamento di una delle patologie con maggior tasso di mortalità nella regione Marche, utilizzando e comparando le principali metodologie statistiche di correzione per la molteplicità dei test.

## 2. MATERIALI E METODI

La verifica di ipotesi può essere applicata ad una singola ipotesi scientifica su un parametro di una popolazione, oppure riguardare più ipotesi contemporaneamente; poiché ciascun test ha un determinato errore di prima specie  $\alpha$ , all'aumentare del numero di test effettuati sullo stesso campione di dati, la probabilità di giungere ad una conclusione errata aumenta rapidamente all'aumentare del numero di ipotesi verificate.

Se si considera un numero di ipotesi pari ad  $m$ , i possibili risultati dei test di ipotesi possono essere (tabella 2):

**Tabella 2. Tassi di errore di prima e seconda specie nella verifica di  $m$  ipotesi.**

	Accettazione $H_0$	Rifiuto $H_0$	Totale
Ipotesi nulle vere	U	V	$m_0$
Ipotesi alternative vere	T	S	$m_1$
Totale	$m-R$	R	$m$

La probabilità di commettere almeno un errore che genera un falso positivo nella famiglia di tutti i test è il tasso di errore globale dell'esperimento ed è indicato come Family Wise Error Rate (FWER); sulla base della formalizzazione riportata in tabella 2, il FWER è rappresentato dalla probabilità che  $V$  sia superiore o uguale a 1.

La procedura tradizionale che controlla il FWER è quella di Bonferroni che, considerate  $m$  ipotesi, consiste nel controllare la probabilità di ottenere almeno un risultato falso positivo modificando il livello di significatività che, anziché essere pari ad  $\alpha$ , viene assunto pari ad  $\alpha/m$  per ogni test; operativamente, quindi, ogni test di ipotesi che presenta un valore  $p$  inferiore ad  $\alpha/m$  viene rigettato (tabella 3, procedura a). Alternativamente, la verifica di ipotesi può essere considerata in termini di valori  $p$  aggiustati per la molteplicità delle ipotesi; piuttosto che confrontare il valore  $p$  del singolo test con il livello di significatività modificato per i confronti multipli, si aggiusta il valore  $p$  del singolo test tenendo conto del FWER e si confronta con il livello di significatività globale dell'esperimento, cioè direttamente con  $\alpha$ . Il valore- $p$  aggiustato in accordo al metodo di Bonferroni è calcolato moltiplicando il valore  $p$  della singola ipotesi nulla per il numero  $m$  di ipotesi verificate (tabella 3, procedura b).

**Tabella 3. Procedure alternative di Bonferroni per la correzione dei confronti multipli.**

	Probabilità del risultato		Livello di significatività di riferimento
a)	valore- $p$	VS	$\alpha/m$
b)	valore- $p * m$	VS	$\alpha$

Nota:  $m$  = numero ipotesi;  $\alpha$  = livello di significatività dell'esperimento; valore- $p$  = risultato del singolo test di ipotesi.

Considerando che, dato un campione di dati, non è possibile ridurre contemporaneamente l'errore di prima e seconda specie, la procedura di correzione di Bonferroni tende a ridurre in maniera eccessiva la probabilità di commettere un falso positivo (errore di prima specie) alzando eccessivamente il rischio di commettere un falso negativo (errore di seconda specie).

Sono stati proposti diversi metodi di controllo del tasso di errore globale dell'esperimento meno conservativi rispetto a quello proposto da Bonferroni, quali ad esempio quello di Šidák (2), Hommel G. (3), Hochberg (4), metodi di aggiustamento del valore p basati sulla procedura di Bootstrap (6) e con l'uso delle permutazioni (6, 7).

In alternativa a questi metodi di controllo del FWER, che conducono a livelli di significatività ancora troppo stringenti per ogni test, sono stati proposti dei criteri che bilanciano in maniera più incisiva il rischio di commettere errori di primo tipo e la potenza dei test; Benjamini e Hochberg hanno introdotto il tasso di false scoperte (False Discovery Rate, FDR) (8) definendolo come la proporzione attesa del numero di risultati falsi positivi sul totale di tutti i risultati positivi e cioè, in altri termini, del numero di ipotesi nulle erroneamente rifiutate sul totale di quelle rifiutate. Facendo riferimento alla tabella 2, il False Discovery Rate si può esprimere come una funzione fattorizzata in due termini:

$$FDR = \begin{cases} 0 & \text{se } R = 0 \text{ (nessuna ipotesi è rifiutata)} \\ E\left(\frac{V}{R}\right) & \text{se } R > 0 \text{ (almeno un'ipotesi viene rifiutata)} \end{cases}$$

Il FDR è meno restrittivo del FWER e, ai fini del controllo del rischio di falsi positivi in presenza di una molteplicità di test di ipotesi sullo stesso campione, è una procedura più potente. Quando infatti molte ipotesi alternative sono in realtà vere, il tasso di false scoperte produce più risultati statisticamente significativi rispetto al test di Bonferroni dimostrandosi quindi una procedura potenzialmente più potente. Storey ha proposto di considerare solo il primo termine della scomposizione dell'FDR definendo così il tasso di false scoperte positivo (Positive False Discovery Rate, pFDR), ritenendolo più adatto in studi dove si prevede di rilevare almeno un risultato statisticamente significativo. Il valore p dei singoli test di ipotesi possono essere aggiustati tenendo conto del FDR e del pFDR in base ad una procedura ottimizzata che si basa sulle caratteristiche della distribuzione dei valori p dei test di tutte le ipotesi considerate nell'indagine; questo comporta che l'interpretazione del numero di falsi positivi per un dato valore di soglia va quindi fatta alla luce dell'intero gruppo di valori p osservati.

Il valore p aggiustato secondo l'approccio del tasso di false scoperte positivo (pFDR) ha il vantaggio di avere un'interpretazione in termini bayesiani e rappresenta la probabilità a posteriori dell'ipotesi nulla sulla base del campione osservato, corretta per la molteplicità dei test (9,10).

In questo studio sono stati analizzati i soggetti residenti nei comuni della regione Marche nel periodo 2006-2009 con malattia ischemica del cuore (ICD X I20-I25) come causa iniziale di decesso; la causa di morte per ischemia del cuore è stata selezionata alla luce dei risultati dei rapporti ISTISAN 2008/2009 (11,12) in cui emerge che la regione Marche ha avuto un netto incremento di mortalità per questa causa di morte passando dal 2008 al 2009, in una graduatoria crescente dei tassi di mortalità, dal 6° al 17° posto nel gruppo dei maschi e dal 7° al 18° posto nella graduatoria delle femmine (13).

Tenendo conto che il numero di comuni delle Marche è di 239, e che sono stati esaminati separatamente i maschi dalle femmine, il numero totale di ipotesi analizzate sono state 478 (239 comuni, 2 generi, 1 causa di decesso).

Per ogni comune e genere sono stati calcolati i rapporti di mortalità standardizzati per classi di età quinquennali e indice di deprivazione (SMR), con i relativi intervalli di confidenza al 95% e valore p.

Al fine di correggere per la molteplicità dei test, per ogni SMR osservato sono stati calcolati i valori p aggiustati secondo il metodo di Bonferroni, secondo l'approccio FDR e pFDR utilizzando la procedura Multtest del software SAS System v. 9.3; per comparazione, il valore q è stato calcolato anche con il software R-Project utilizzando la libreria "qvalue", versione 1.0 aggiornata a gennaio 2004.

In aggiunta è stato calcolato per ogni comune anche il rapporto di mortalità bayesiano (BMR) controllato per classi di età quinquennali e indice di deprivazione, con il relativo l'intervallo di credibilità al 95%, ma per questo tipo di inferenza non è stata adottata nessuna procedura di controllo dell'errore di primo tipo.



---

### 3. RISULTATI

---

Per ciascuna delle 478 ipotesi è stato calcolato il rapporto standardizzato di mortalità ed è risultato che meno dell'uno per cento dei gruppi (Comuni di Monte Grimano Terme, Camporotondo di Fiastrone, Montedinove e Monte Rinaldo) presentava un SMR pari a zero. Fissato a priori un livello di significatività di 0,05, ci si attende un numero di falsi positivi pari a 24 ipotesi; dai risultati è emerso che il 7,11% dei valori p (n. 34 su 478) era inferiore al livello di significatività dello 0,05 mentre dopo l'aggiustamento secondo il criterio di Bonferroni le ipotesi s.s. sono risultate due (0,4%) (tabella 4-5). Utilizzando l'approccio del False Discovery Rate e del Positive False Discovery Rate, le ipotesi con un valore q inferiore alla soglia dello 0,05 sono risultate tre (tabella 5).

Il comune di Monte Vidon Corrado risulta avere un eccesso di mortalità per malattie ischemiche del cuore particolarmente alto e statisticamente significativo, anche dopo l'aggiustamento per confronti multipli secondo tutti i metodi di correzione utilizzati.

Tenendo conto che il calcolo e l'interpretazione del valore q sono basati sulla lista dei valori-p ordinati in maniera crescente (che di conseguenza sono ordinati in maniera crescente anche per valore q), il valore q di 0,0459 osservato per il comune di Recanati indica che delle ipotesi rifiutate fino a quel livello bisogna attendersi il 4,59% di risultati falsi positivi: in sostanza tra le tre ipotesi rifiutate, meno di una è un risultato falso positivo ( $0,0459 \cdot 3 = 0,1377$ ). Accettando un cut-off più alto per il valore q, come ad esempio la soglia pari a 0,10, si può affermare che tra le cinque ipotesi che mostrano un valore q inferiore o uguale a 0,10 il 7,75% è un risultato falso positivo: tra le cinque ipotesi rifiutate meno di una è un risultato falso positivo ( $0,0775 \cdot 5 = 0,3875$ ). A differenza del valore p, l'interpretazione del valore q è quindi sempre vincolata alla lista ordinata dei test che mostrano un valore q minore del cut-off scelto.

Per tutti i risultati osservati sul campione in studio, il valore q e il valore-p aggiustato secondo l'approccio del False Discovery Rate risultano simili poiché la probabilità di rilevare almeno un risultato statisticamente significativo è molto alta (prossima all'unità), circostanza molto frequente in studi descrittivi di epidemiologia ambientale, come la presente indagine.

**Tabella 4. Regione Marche. Decessi 2006-2009 per Malattie ischemiche del cuore (I20-I25): rapporti standardizzati classici e bayesiani con valore  $p < 0,05$ .**

Comune	Gen.	Oss.	Pop.	Att.	SMR %	LI 95% SMR	LS 95% SMR	BMR	2,5% IC	97,5 % IC
San Benedetto del Tronto	F	80	99.243	135	59	47	74	63	51	76
Monte Vidon Corrado	F	15	1.673	3	<b>488</b>	273	805	<b>197</b>	130	323
Recanati	F	39	44.416	67	58	41	80	72	58	88
Matelica	F	60	21.042	37	<b>161</b>	123	207	<b>140</b>	113	173
Porto Recanati	F	14	23.028	31	45	24	75	76	54	101
Recanati	M	41	40.962	66	62	45	85	81	67	97
Apecchio	M	17	4.108	7	<b>239</b>	139	383	131	97	183
Pieve Torina	F	14	2.869	5	<b>262</b>	143	440	<b>148</b>	108	212
Montecalvo in Foglia	M	14	5.267	6	<b>248</b>	135	416	125	92	176
Ripatransone	F	5	9.011	15	34	11	79	67	49	88
Esanatoglia	F	17	4.523	8	<b>211</b>	123	337	<b>156</b>	106	228
Pieve Torina	M	13	2.891	6	<b>231</b>	123	395	117	90	156
Morrovalle	M	15	19.986	28	54	30	88	78	61	97
Montelupone	M	3	7.066	11	29	6	83	78	58	99
Montefano	F	4	7.180	12	33	9	85	73	51	100
Fermo	M	153	72.900	125	<b>122</b>	104	143	110	95	125
Agugliano	M	20	8.979	11	<b>180</b>	110	278	118	92	157
Sassocorvaro	F	18	6.952	10	<b>186</b>	110	293	114	83	154
Falerone	F	23	7.109	13	<b>171</b>	108	257	<b>141</b>	106	186
Camerino	F	41	14.744	28	<b>146</b>	105	199	124	98	157
Apecchio	F	17	4.227	9	<b>186</b>	108	298	137	97	195
Appignano del Tronto	M	1	3.852	6	16	0	88	83	60	111
Gradara	M	2	8.848	8	25	3	90	82	57	113
Urbino	F	68	31.618	51	<b>133</b>	103	168	114	95	136
Montegiorgio	F	34	14.128	23	<b>151</b>	104	210	<b>129</b>	102	163
Acquacanina	F	4	243	1	<b>424</b>	116	1.086	121	65	211
Camerano	F	28	14.326	18	<b>154</b>	103	223	119	89	159
Monte Roberto	M	13	5.671	7	<b>195</b>	104	333	115	88	155
Jesi	F	159	83.631	134	<b>119</b>	101	138	<b>115</b>	100	132
Fiordimonte	F	5	435	2	<b>327</b>	106	764	129	89	192
Montefano	M	4	6.989	11	38	10	97	87	66	111
Cagli	F	47	18.631	34	<b>138</b>	101	183	<b>129</b>	105	160
Monte Urano	M	13	16.502	22	58	31	100	86	65	109
Cupra Marittima	F	7	10.885	14	49	20	100	62	40	89

**Tabella 5. Regione Marche. Decessi 2006-2009 per Malattie ischemiche del cuore (I20-I25): valori p grezzi e aggiustati per la molteplicità dei test (solo test con  $p < 0,05$ ).**

<b>Comune</b>	<b>Gen.</b>	<b>Valore p</b>	<b>Bonferroni</b>	<b>FDR</b>	<b>Valore q</b>
San Benedetto del Tronto	F	0	0,0002	0,0002	0,0002
Monte Vidon Corrado	F	0,000002	0,0009	0,0004	0,0004
Recanati	F	0,000297	0,1421	0,0474	0,0459
Matelica	F	0,000775	0,3707	0,0799	0,0775
Porto Recanati	F	0,000836	0,3996	0,0799	0,0775
Recanati	M	0,001517	0,725	0,1208	0,1171
Apecchio	M	0,002259	1	0,1527	0,148
Pieve Torina	F	0,002556	1	0,1527	0,148
Montecalvo in Foglia	M	0,004309	1	0,2289	0,2218
Ripatransone	F	0,006487	1	0,3101	0,3006
Esanatoglia	F	0,008119	1	0,3528	0,342
Pieve Torina	M	0,010661	1	0,4029	0,3905
Morrovalle	M	0,010957	1	0,4029	0,3905
Montelupone	M	0,014227	1	0,4556	0,4416
Montefano	F	0,014296	1	0,4556	0,4416
Fermo	M	0,017688	1	0,5284	0,5122
Agugliano	M	0,020613	1	0,5477	0,5309
Sassocorvaro	F	0,021716	1	0,5477	0,5309
Falerone	F	0,022073	1	0,5477	0,5309
Camerino	F	0,024765	1	0,5477	0,5309
Apecchio	F	0,025402	1	0,5477	0,5309
Appignano del Tronto	M	0,02627	1	0,5477	0,5309
Gradara	M	0,026353	1	0,5477	0,5309
Urbino	F	0,028157	1	0,5608	0,5436
Montegiorgio	F	0,029634	1	0,5666	0,5492
Acquacanina	F	0,031358	1	0,5765	0,5588
Camerano	F	0,037892	1	0,6151	0,5962
Monte Roberto	M	0,038725	1	0,6151	0,5962
Jesi	F	0,039738	1	0,6151	0,5962
Fiordimonte	F	0,039797	1	0,6151	0,5962
Montefano	M	0,041014	1	0,6151	0,5962
Cagli	F	0,041175	1	0,6151	0,5962
Monte Urano	M	0,049206	1	0,6878	0,6667
Cupra Marittima	F	0,050408	1	0,6878	0,6667

---

## 4. DISCUSSIONE E CONCLUSIONI

---

Nell'ambito dell'epidemiologia c'è una crescente preoccupazione per gli studi che evidenziano la presenza di un rischio sanitario non riprodotto da ulteriori indagini. Il grado di incertezza insito nel processo inferenziale dà origine alla reale possibilità di sbagliare, ma il rischio di commettere errori aumenta quando le ipotesi da verificare su uno stesso campione di dati sono molteplici. Per ridurre il margine di errore, la metodologia scientifica suggerisce, tra l'altro, l'uso di metodi statistici che riducono il rischio di avere dei risultati falsi positivi. Sulla base dei soggetti deceduti nei comuni della regione Marche che presentavano nella causa di decesso le malattie ischemiche del cuore (classificazione ICDX I20-I25) sono stati confrontati alcuni metodi statistici per il controllo dell'errore dei falsi positivi quali il metodo di Bonferroni e quello del tasso di false scoperte positivo (pFDR) proposto da Storey.

Su un totale di 478 ipotesi valutate sul campione in studio, il 7% (n. 34) mostrava un rischio di decesso statisticamente significativo al livello di confidenza del 5%; dopo l'aggiustamento secondo il criterio di Bonferroni le ipotesi risultate s.s. (in eccesso o in difetto) sono diminuite drasticamente, passando da 34 a 2, mentre secondo l'approccio del valore q le ipotesi s.s. sono risultate 3. Nello specifico è il comune di Monte Vidon Corrado a mostrare un eccesso di mortalità per le malattie ischemiche del cuore molto più alto rispetto alla media regionale (SMR% = 488), in maniera statisticamente significativa anche dopo l'aggiustamento per la molteplicità dei test secondo tutti i metodi di correzione utilizzati; questo risultato pone in evidenza una situazione sanitaria critica di un comune marchigiano che andrebbe ulteriormente esaminata al fine di individuare, con studi ad hoc, i possibili fattori eziologici o gli eventuali vizi nelle registrazioni connessi all'aumento del rischio di morire per malattie ischemiche del cuore.

Negli studi analitici, laddove un piano di analisi prevede la valutazione di confronti tra sottogruppi dello stesso campione di dati, risulta necessario controllare il rischio di commettere errori falsi positivi utilizzando l'approccio del tasso di errore globale dell'esperimento (FWER, es. correzione di Bonferroni) oppure quello del tasso di false scoperte (FDR, es. valore q).

Per studi di tipo esplorativo, in particolare per quanto concerne le indagini basate sui dati provenienti dai database di tipo amministrativo (es. schede di dimissione ospedaliera, schede di morte, prescrizioni farmaceutiche, ecc.) e quindi raccolti con una finalità diversa dalla verifica di un'ipotesi epidemiologica predefinita, la correzione per la molteplicità delle ipotesi in studio non è strettamente necessaria (14). In questo tipo di indagini sono invece prioritari i criteri di correzione delle fonti di distorsioni legate alla limitata accuratezza dei sistemi informatizzati anche in termini di completezza, alla qualità della compilazione e codifica delle informazioni cliniche e al limitato controllo dei fattori di confondimento.

I risultati delle analisi esplorative corrette per la molteplicità dei test sono tuttavia utili per descrivere ulteriormente un fenomeno ed eventualmente essere un elemento aggiuntivo per la formulazione di ipotesi epidemiologiche da valutare in un successivo studio analitico (14).

Un caso particolare di necessità di effettuare la correzione per i confronti multipli, almeno per singola patologia indagata, potrebbe essere quello individuato nelle indagini epidemiologiche descrittive previste dal Decreto 24 aprile 2013 dei Ministeri della Salute e

dell'Ambiente "Criteri metodologici utili per la redazione del rapporto di valutazione del danno sanitario (VDS)" che in caso di evidenza di eccessi di eventi patologici impongono l'effettuazione di analisi più approfondite. La correzione per confronti multipli permetterebbe di essere più selettivi e di concentrare l'attenzione solo sugli eventi sicuramente associabili alle esposizioni e che presentano maggiore indicazione d'impatto.

---

## BIBLIOGRAFIA

---

1. Storey JD, A direct approach to false discovery rates, 2002, *J.R.Statist. Soc. B*, 64, Part 3, pp.479-498.
2. Šidák, Z. (1967), "Rectangular Confidence Regions for the Means of Multivariate Normal Distributions," *Journal of the American Statistical Association*, 62, 626–633.
3. Hommel, G. (1988), "A Comparison of Two Modified Bonferroni Procedures," *Biometrika*, 75, 383–386.
4. Hochberg, Y. (1988), "A Sharper Bonferroni Procedure for Multiple Significance Testing," *Biometrika*, 75, 800–803.
5. Westfall, P. H. and Young, S. S. (1993), *Resampling-Based Multiple Testing: Examples and Methods for p-Value Adjustment*, New York: John Wiley & Sons.
6. Brown, C. C. and Fears, T. R. (1981), "Exact Significance Levels for Multiple Binomial Testing with Application to Carcinogenicity Screens," *Biometrics*, 37, 763–774.
7. Heyse, J. and Rom, D. (1988), "Adjusting for Multiplicity of Statistical Tests in the Analysis of Carcinogenicity Studies," *Biometrical Journal*, 30, 883–896.
8. Benjamini Y, Hochberg Y, Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing, *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 57, No. 1 (1995), pp. 289-300.
9. Storey JD, The positive false discovery rate: a Bayesian interpretation and the q-value, *The Annals of Statistics*, 2003, Vol.31, No. 6, pp. 2013-2035.
10. Catelan D, Biggeri A, Barbone F, Test multipli e analisi di sottogruppo (cosa c'è di sbagliato nel trovare sempre qualcosa di significativo), 2011, *Epidemiologia & Prevenzione*, Anno 35 (2), marzo-aprile, pp.150-154.
11. Minelli G, Manno V, D'ottavi S,M. La mortalità in Italia nell'anno 2009. Istituto Superiore di Sanità; Istituto Nazionale di Statistica; 2012.
12. Minelli G, Manno V, D'ottavi S,M. La mortalità in Italia nel 2008. Istituto Superiore di Sanità; Istituto Nazionale di Statistica; 2011.
13. Bartolacci S, Di Biagio K, Baldini M, Simeoni, Mariottini M, Mortalità per causa nella regione Marche, Agenzia Regionale Per la Protezione Ambientale delle Marche - Servizio di epidemiologia ambientale, Gen 2013.
14. Bender R., Lange S., Adjusting for multiple testing – when and how?, *Journal of Clinical Epidemiology*, Vol.54 (2001), pp. 343-349.
15. Bart Ostro et al, Long-Term Exposure to Constituents of Fine Particulate Air Pollution and Mortality: Results from the California Teachers Study, *Environ Health Perspect.* 2010 March; 118(3): 363–369.
16. Jeremy P. Langrish et al, Reducing Personal Exposure to Particulate Air Pollution Improves Cardiovascular Health in Patients with Coronary Heart Disease, *Environ Health Perspect.* 2012 March; 120(3): 367–372.
17. Jennifer K Mann et al, Air pollution and hospital admissions for ischemic heart disease in persons with congestive heart failure or arrhythmia, *Environ Health Perspect.* 2002 December; 110(12): 1247–1252.