
GEOREFERENZIAZIONE DELLE RESIDENZE: UN CASO DIDATTICO IN EPIDEMIOLOGIA AMBIENTALE

AGENZIA PER LA PROTEZIONE AMBIENTALE DELLE MARCHE - OSSERVATORIO
EPIDEMIOLOGICO AMBIENTALE*
AGENZIA SANITARIA UNICA REGIONALE - AREA VASTA N. 4 FERMO - DIPARTIMENTO DI
PREVENZIONE - SERVIZIO IGIENE E SANITA' PUBBLICA**



Settembre 2014

A cura di:

Marco Baldini*
Silvia Bartolacci*
Katuscia Di Biagio*
Rossana Anna Belfiglio**
Thomas Valerio Simeoni*
Mauro Mariottini*
Giuseppe Ciarrocchi**

SOMMARIO

SOMMARIO	2
ABSTRACT	3
1. INTRODUZIONE.....	4
2. MATERIALI E METODI	4
3. RISULTATI.....	5
4. DISCUSSIONE E CONCLUSIONI.....	9
BIBLIOGRAFIA.....	10

ABSTRACT

Obiettivi

Applicando nuove tecniche di ricerca statistica alle normali rilevazioni di sorveglianza epidemiologica è stato riscontrato un eccesso di mortalità per cause cardiovascolari particolarmente evidente dal punto della significatività statistica in un piccolo comune della regione Marche. Più in particolare, il ricorso a tecniche di indagine spaziale ha evidenziato un anomalo aggregato di casi in una piccola località all'interno del comune.

Ad una analisi più attenta, però, l'osservazione è risultata falsamente positiva.

Il caso si presta ad una presentazione didattica sulle procedure di epidemiologia spaziale e sulla necessaria attenzione nello svolgimento delle stesse.

Dati e metodi

Un risultato di 20 decessi nel periodo 2006-2009 per le cause ICD X I20 - I25 tra i residenti nell'area in studio ha spinto gli autori a verificare l'esistenza di aggregati spaziali di casi e di possibili cause ambientali.

Le informazioni sulle ultime residenze dei deceduti sono state raccolte tramite l'amministrazione comunale di competenza, mentre si è ricorsi ai dati censuari ISTAT 2011 per quelle sulla popolazione residente.

Per la ricerca dei cluster è stato utilizzato il software Satscan 9.3.

Risultati

Un apparente cluster spaziale particolarmente significativo centrato su una sezione censuaria è risultato, ad una analisi più attenta dei dati della georeferenziazione delle residenze, inesistente per alcune imprecisioni nella tecnica di raccolta delle misure di georeferenziazione delle residenze dei deceduti.

L'attribuzione delle sezioni censuarie a partire da una geolocalizzazione dei numeri civici è particolarmente delicata, ed è sufficiente l'errore di pochi metri per ottenere risultati completamente erronei.

Conclusioni

Il caso in studio, a nostro avviso, si presta per importanti osservazioni sulla cautela che bisogna mettere in atto in questo tipo di rilievi, nonché ad interessanti valutazioni didattiche.

KEYWORDS : EPIDEMIOLOGIA SPAZIALE, RESIDENZE, GEOREFERENZIAZIONE

1. INTRODUZIONE

Nel corso dell'attività routinaria di sorveglianza dell'occorrenza degli eventi sanitari nella Regione Marche, si è applicato il test del q-value (1) per valutare gli eccessi particolarmente evidenti di alcune cause di decesso nel periodo 2006-2009.

Il test, particolarmente selettivo, ha messo in evidenza - tra i gruppi di patologia indagati - un solo eccesso che superava i criteri del q-value e in particolare l'eccesso di mortalità per cause cardiovascolari (ICD X I20 - I25) (q-value=0,0004) tra i residenti in un comune marchigiano.

Non risultando cause ambientali di particolare interesse nel comune indagato, si è deciso di chiedere informazioni in merito a possibili fattori di rischio locali sia al Dipartimento ARPAM della provincia che al Servizio di Igiene e Sanità Pubblica dell'Area Vasta dell'ASUR di riferimento. Esposizioni a particolari e specifici rischi nel territorio comunale sono risultate sconosciute.

Con la collaborazione dell'ASUR e dell'Amministrazione comunale si è deciso di studiare la distribuzione spaziale e temporale del fenomeno e di verificare la presenza di aggregati di casi (cluster) statisticamente significativi.

Per motivi di tutela delle informazioni personali, ed in particolare dei dati sensibili che interessano in presente studio, non sono stati riportati nel testo dati identificativi né alcune immagini dei territori compresi nell'area di riferimento; per il medesimo motivo sono stati inoltre modificati alcuni dati originali.

2. MATERIALI E METODI

Il Servizio di Igiene e Sanità Pubblica dell'Area Vasta dell'ASUR ha fornito i dati (in forma anonima) dei 20 decessi registrati nel periodo 2006-2009 per le cause in studio (ICD X I20 - I25), includendo i relativi indirizzi dell'ultima residenza nell'area interessata. Il Comune competente ha successivamente provveduto alla georeferenziazione delle residenze stesse.

Con un software GIS (Mapinfo Pro 11.5.4) si è provveduto ad attribuire ad ogni indirizzo l'unità censuaria di riferimento sulla base delle mappe GIS fornite da ISTAT e relative al censimento della popolazione 2011. Dalla stessa fonte è stata ricavata la popolazione divisa per genere per ogni unità censuaria (tabella 1).

La ricerca dei cluster spaziali e temporali è stata effettuata utilizzando la metodologia del cluster detection basata sulla statistica Scan di Kulldorff, che permette di localizzare aggregazioni di sezioni che presentano eccessi di mortalità statisticamente significativi (2) (3). Un p-value associato alla statistica Scan inferiore a 0,05 indica un risultato statisticamente significativo, ovvero che nell'area costituita dalle sezioni del cluster il fenomeno della mortalità, rappresentato dal rischio relativo del cluster, è superiore/inferiore a quello medio comunale non per effetto del caso.

L'analisi statistica è stata effettuata con il software Satscan 9.3.

Tabella 1. Popolazione residente nell'area in studio per sezione di censimento e genere (ISTAT, censimento 2011)

<u>Sez</u>	<u>Anno</u>	<u>Pop</u>	<u>Genere</u>
a	2011	95	M
a	2011	105	F
b	2011	17	M
b	2011	22	F
c	2011	8	M
c	2011	7	F
d	2011	8	M
d	2011	13	F
e	2011	24	M
e	2011	22	F
f	2011	77	M
f	2011	75	F
g	2011	25	M
g	2011	29	F
h	2011	38	M
h	2011	47	F
i	2011	14	M
i	2011	9	F
l	2011	6	M
l	2011	9	F
m	2011	20	M
m	2011	20	F
n	2011	8	M
n	2011	10	F
o	2011	34	M
o	2011	27	F
p	2011	5	M
p	2011	3	F

3. RISULTATI

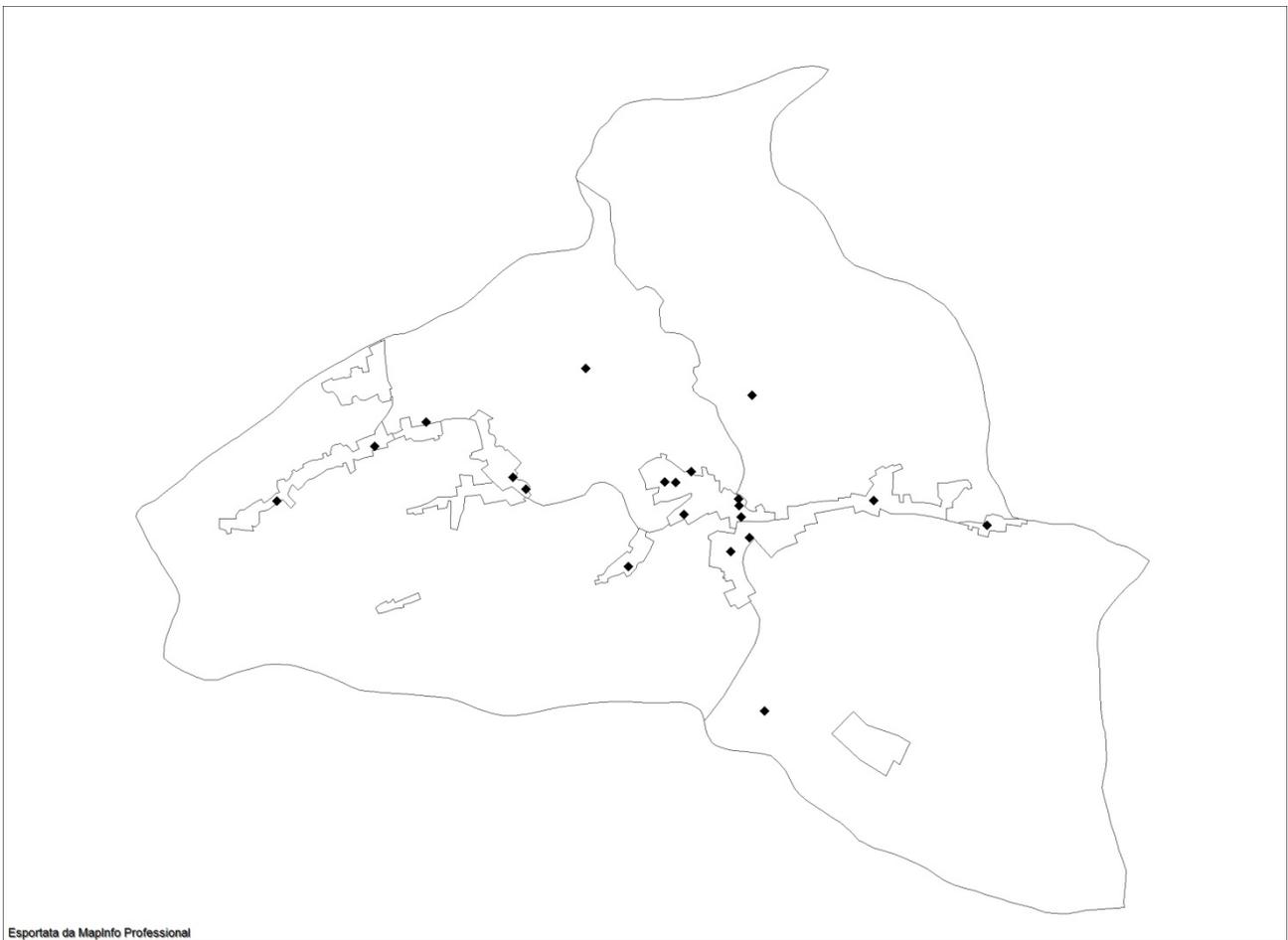
I decessi per le cause ICD X I20 - I25 tra i residenti nell'area in studio nel periodo 2006-2009 sono rappresentati in tabella 2. La sezione di censimento è stata determinata riportando le coordinate geografiche dell'ultima residenza sulla mappa ISTAT delle sezioni censuarie (figura 1).

Tabella 2. Decessi per causa ICD X I20 - I25 tra i residenti nell'area in studio nel periodo 2006-2009

<u>Genere</u>	<u>Anno</u>	<u>Sez. cens.</u>
F	2009	a
F	2009	a
F	2006	a
M	2008	a
M	2006	a
F	2008	b

F	2006	f
M	2008	f
M	2007	f
F	2009	g
F	2006	g
F	2006	l
F	2006	n
F	2006	o
F	2009	p
F	2008	p
M	2006	p

Figura 1. Mappa delle sezioni di censimento nell'area in studio e georeferenziazione delle residenze dei deceduti



La ricerca dei cluster spaziali (4) con Satscan ha evidenziato un cluster statisticamente significativo ($p\text{-value}=0,016$) con 3 decessi e con un rischio relativo pari a 19,54 centrato e comprendente una sola sezione.

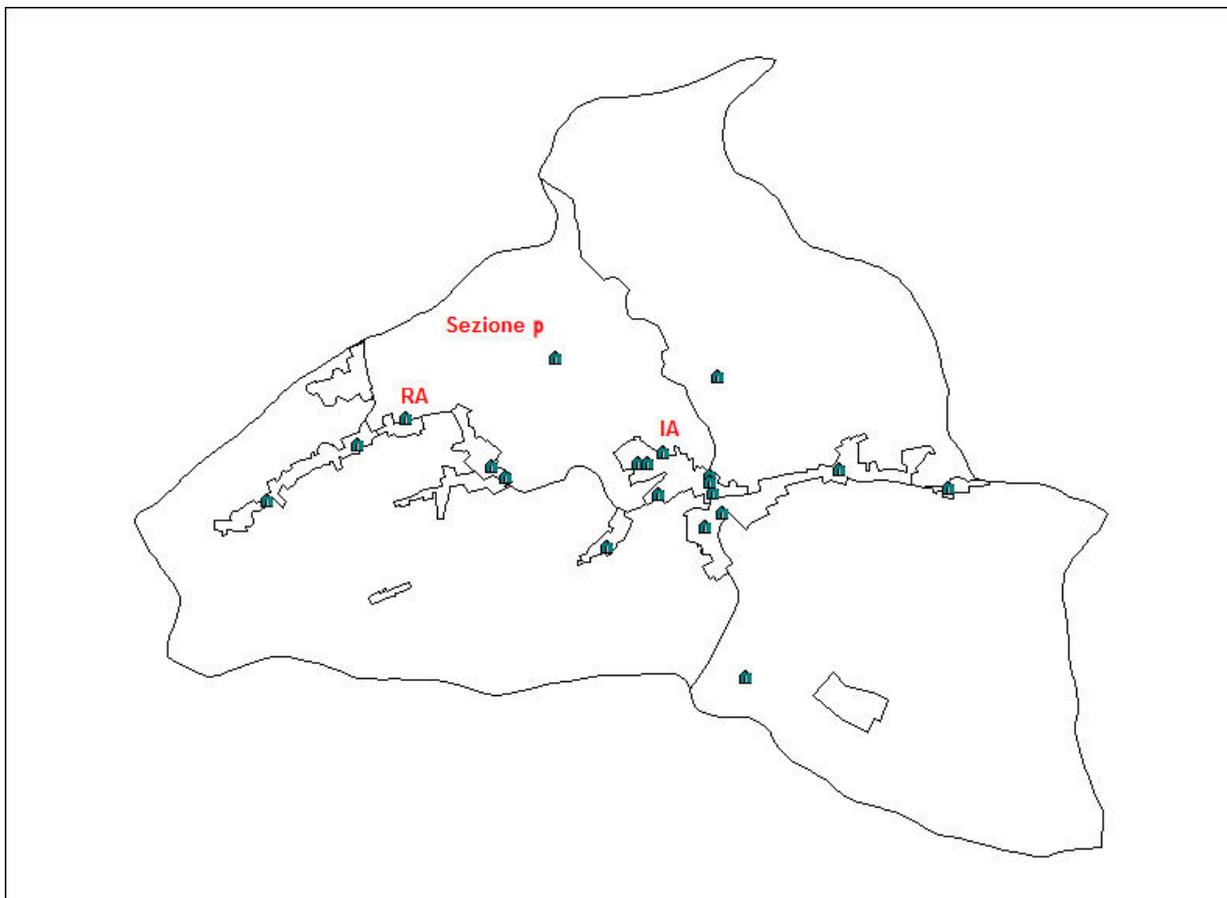
La ricerca dei cluster temporali con Satscan ha evidenziato un cluster statisticamente significativo nell'anno 2006 ($p\text{-value}=0,017$) con 11 decessi e con un rischio relativo pari a 3,67.

La ricerca dei cluster combinati spazio-temporali ha confermato i risultati precedenti.

Il risultato ampiamente significativo/suggestivo ha tuttavia dato luogo ad alcuni dubbi, in particolare per il fatto che nella sezione del cluster non sembrano esserci particolari pressioni ambientali che giustifichino una così alta proporzione di decessi per cause cardiovascolari sulla scarsa popolazione residente.

Si è proceduto quindi ad un esame più accurato delle mappe delle sezioni censuarie e delle georeferenziazioni delle residenze, evidenziando due misure sospette (figura 3).

Figura 2. Georeferenziazioni sospette nella sezione sede del cluster.



Sovrapponendo con l'applicativo GIS le carte delle sezioni censuarie, delle residenze georeferenziate e della mappa aerea del territorio ricavata da "Google Earth" e zoomando sufficientemente le due aree in esame, si è constatata l'imprecisa georeferenziazione delle residenze nella sezione censuaria per l'inesistenza di fabbricati nelle aree indicate dalla georeferenziazione collocate appena all'interno dei confini nella stessa sezione (figure 4 e 5).

Figura 3. Errata georeferenziazione della residenza "RA"

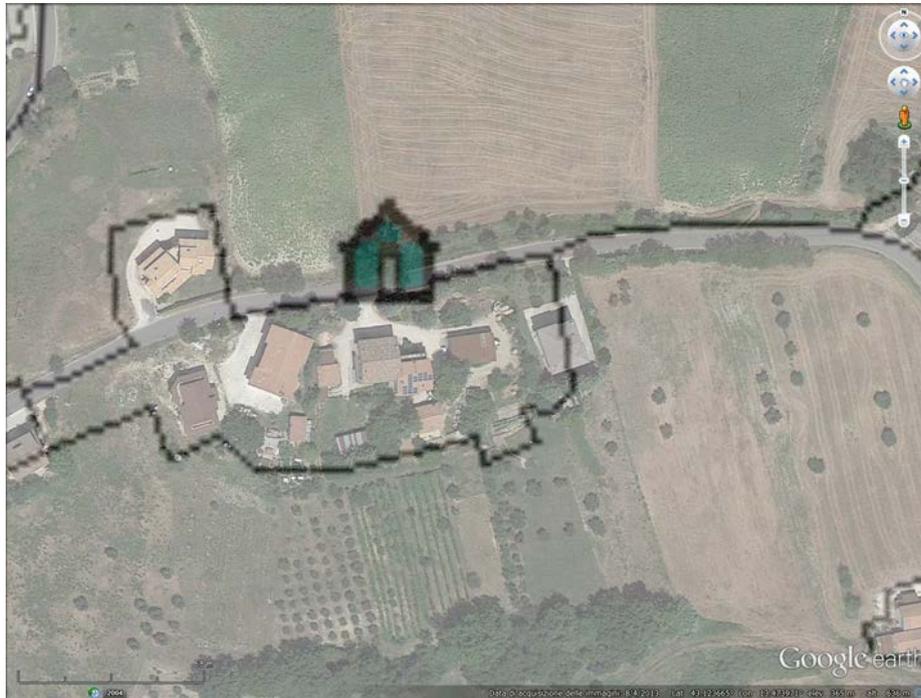


Figura 4. Errata georeferenziazione della residenza "IA"



A seguito di tale accertamento due casi sono stati spostati in altre due sezioni.

Dopo la correzione dei dati è stata calcolata nuovamente la statistica Scan per la ricerca dei cluster; tale ricerca ha naturalmente riconfermato il riscontro del cluster temporale ma non di quello spaziale.

4. DISCUSSIONE E CONCLUSIONI

In seguito al rilevamento di un eccesso di mortalità per malattie ischemiche del cuore osservato in un comune marchigiano, rispetto ai dati riferiti all'intera regione, è stato effettuato un approfondimento per indagare la presenza di eventuali aggregazioni di decessi all'interno di una data area comunale.

Dall'analisi dei cluster è risultato un aggregato spaziale statisticamente significativo in una sezione censuaria, ma l'assenza di pressioni ambientali o fattori di rischio locali ha reso difficilmente interpretabile il risultato.

Un controllo aggiuntivo sui dati ha rilevato errori nel metodo di georeferenziazione delle residenze di due decessi nella sezione, che sono quindi stati spostati in altre sezioni censuarie. L'analisi dei cluster sui nuovi dati aggiornati dopo la correzione della georeferenziazione non ha riconfermato il risultato statisticamente significativo della prima analisi.

Le valutazioni di epidemiologia spaziale sono molto importanti per lo studio della distribuzione spaziale degli eventi sanitari e in particolare per ipotizzare l'esistenza di effetti sanitari in concomitanza di rilevate pressioni ambientali.

In questi studi è indispensabile avere informazioni sulle residenze dei "casi" e sulla quantificazione e strutturazione della popolazione residente nell'area.

Negli archivi dei dati sanitari correnti non sono raccolte notizie residenziali nel dettaglio necessario per detti studi, e per avere quelle relative ai denominatori di popolazione residente necessita spesso utilizzare dati correnti quali i dati censuari. Questi ultimi, tuttavia, non consentono dal punto di vista cartografico di definire con lo stesso strumento anche la localizzazione degli eventi al numeratore.

Nell'attesa del completamento e nella speranza di poter poi accedere all'Archivio nazionale degli stradari e dei numeri civici (ANSC) dell'Agenzia delle Entrate e dell'ISTAT (5), si devono utilizzare metodologie diverse per la geolocalizzazione dei casi e della popolazione residente nell'area in studio.

Poiché la georeferenziazione dei numeri civici non è ancora disponibile in tutti i comuni italiani, talora si deve ricorrere alla geolocalizzazione manuale delle residenze con i possibili errori legati alle misure e agli strumenti GPS utilizzati e ai punti di accesso per le misure stesse.

Questo può comportare imprecisioni come quelle descritte nel caso in studio in cui è probabile che i confini delle proprietà georiferite corrispondano ai confini dell'unità censuaria più vicina (es. bordo stradale).

Per ovviare al problema sarebbe necessario il controllo sulle carte delle singole residenze che, se possibile negli studi di piccole dimensioni come quello presentato, è pressoché impossibile in maniera sistematica in studi di maggiori dimensioni che richiederebbero il controllo manuale di migliaia di punti.

Anche in quest'ultimo caso, così come in presenza di bassa numerosità degli eventi o di riscontro di rischi relativi ai limiti della significatività statistica, il Servizio di Epidemiologia Ambientale effettua comunque controlli a campione.

Le valutazioni di epidemiologia spaziale non hanno in questa ricerca portato contributi apprezzabili alla spiegazione dell'eccesso di mortalità per malattie cardiovascolari rilevato nel piccolo comune marchigiano, ma hanno riaffermato la necessità di porre la massima attenzione in particolare sulla qualità e consistenza dei dati, al fine di evitare falsi esiti in questo tipo di studio dove gli errori possono influenzare i risultati delle analisi in maniera sostanziale soprattutto quando il campione in studio è di ridotte dimensioni.

BIBLIOGRAFIA

1. Storey J. A direct approach to false discovery rates. *J.R.Statist. Soc. B.* 2002; 64, Part 3: p. 479-498.
2. Kulldorff M. A spatial scan statistic. *Communications in Statistics: Theory and Methods.* 1997;(26).
3. Kulldorff M, et al. An elliptic spatial scan statistic. *Statistics in Medicine.* 2006;(25).
4. Wartenberg D. Investigation of disease clusters. In Baker D, Nieuwenhuijsen MJ. *environmental epidemiology.* Oxford: Oxford University Press; 2008. p. 196-206.
5. Agenzia delle entrate. [Online].; 2014 [cited 2014 9 2]. Available from: <http://www.agenziaentrate.gov.it/wps/content/Nsilib/Nsi/Home/CosaDeviFare/Consultare+dati+catastali+e+ipotecari/Scambio+dati+catastali+e+cartografici+con+enti+o+PA/Portale+per+i+Comuni/Servizi+portale+per+i+Comuni/Servizi+portale+dei+comuni/toponomastica/>.