

Stima della variabilità spaziale dell'esposizione a inquinanti aerodispersi in ambito urbano

Estimate of the spatial variability of exposure to airborne pollutants in urban areas

Andrea Ranzi

Epidemiol Prev 2012; 36(5): 287-295

RIASSUNTO

Numerosi studi italiani e internazionali hanno indagato gli effetti a breve termine dell'inquinamento atmosferico in diverse aree geografiche, mentre le conoscenze sull'associazione tra esposizione a lungo termine a sostanze inquinanti e mortalità per cause naturali e cardiorespiratorie sono ad oggi non completamente esaustive. La variabilità spaziale degli inquinanti su scala intraurbana è stata analizzata con diverse tecniche modellistiche. Obiettivo di questo lavoro è la presentazione dell'approccio *Land Use Regression (LUR) model* che sta avendo applicazioni notevoli per la stima e la rappresentazione della variabilità intraurbana dell'inquinamento atmosferico ai fini di studi epidemiologici. A partire dalle esperienze disponibili in letteratura, viene qui presentato un inquadramento metodologico della procedura, evidenziandone vantaggi e limiti, soprattutto in relazione all'applicabilità dei modelli nella valutazione dell'esposizione di popolazioni residenti in aree urbane.

Parole chiave: modello Land Use Regression (LUR), variabilità spaziale dell'inquinamento atmosferico, GIS, esposizione della popolazione

ABSTRACT

Several Italian and international studies investigated the short-term effects of air pollution in different geographical areas, while knowledge on the association between long-term exposure to pollutants and mortality due to natural causes and cardio are fully comprehensive at present. The spatial variability of pollutants within urban areas was analyzed with different modelling techniques. The aim of the present work is an introduction to the Land Use Regression (LUR) model, a promising approach to the estimation and representation of intraurban air pollution for epidemiological studies. Starting from the experiences available in the literature, a methodological framework of the procedure is presented, highlighting the advantages and limitations, especially in relation to the applicability of these models to exposure assessment of populations living in urban areas.

Keywords: Land Use Regression (LUR) model, spatial variability of air pollution, GIS, population exposure

INTRODUZIONE

L'inquinamento urbano rappresenta un fattore di rischio per la salute umana con effetti a breve e a lungo termine. Nelle ultime due decadi sono stati condotti numerosi studi sugli effetti a breve termine in diverse aree geografiche (si ricordano per la realtà italiana il progetto MISA^{1,2} e il più recente EpiAir),³ mentre lo sviluppo di studi sugli effetti a lungo termine è più recente, e le conoscenze sull'associazione tra esposizione a lungo termine a sostanze inquinanti e mortalità per cause naturali e cardiorespiratorie sono ancora oggi considerate non completamente esaustive. Un'analisi dei lavori sugli effetti cronici che l'inquinamento atmosferico ha sulla salute rivela l'importanza degli studi effettuati sulle due maggiori coorti esistenti: la coorte dell'*American Cancer Society* (ACS)⁴⁻⁸ e lo studio delle sei città americane.⁹ I risultati raggiunti in questi studi, relativi all'associazione fra inquinamento e mortalità generale e causa-specifica, ha influenzato le decisioni legislative riguardo agli standard di qualità dell'aria; le stime ottenute sono anche state utilizzate dall'OMS come base per la stima della mortalità attribuibile a inquinamento atmosferico.¹⁰

In questi studi l'attenzione è stata posta sulla variabilità nell'esposizione tra diverse regioni, città o aree estese. L'analisi è stata condotta confrontando i livelli di esposizione di diverse città (o area più vasta), assegnando a tutti i soggetti di una città lo stesso valore di esposizione valutato tramite l'analisi delle centraline fisse di monitoraggio. Tuttavia è noto come l'esposizione all'inquinamento atmosferico possa essere diversa dal punto di vista spaziale all'interno di una città¹¹⁻¹⁵ e queste variazioni possono anche essere connesse a differenze nelle caratteristiche delle persone esposte (per esempio gradienti socioeconomici). Infatti gli effetti sulla salute possono essere maggiori intorno alle sorgenti di inquinamento, e la stima di questi effetti è attenuata quando si utilizzano le concentrazioni medie per l'intera comunità. Teoricamente, il classico errore di misura dell'esposizione indotto dall'utilizzo delle centraline fisse distorce i risultati verso il valore nullo.¹⁶

In anni più recenti, l'esposizione cronica a differenti gradienti di inquinamento su aree più limitate (cioè intraurbane) è stata associata a eventi cardiaci e respiratori.¹⁷ Un numero crescente di studi ha evidenziato gradienti di esposizione derivanti dalla vicinanza ad arterie stradali¹⁸ e ha dimostrato che vi sono impatti negativi a livello locale sulla salute respiratoria nei bambini.^{19,20} Per meglio caratterizzare la variabilità spaziale degli inquinanti su scala intraurbana sono state utilizzate diverse tecniche, dai metodi di interpolazione, come il *kriging*,²¹ ai più recenti metodi di regressione basati su informazioni sull'utilizzo del territorio (LUR, *Land Use Regression*), di cui il primo esempio risale a un lavoro inglese del 1997.²² Inoltre, misure *proxy* più semplici, come la distanza dalle strade, si sono rivelate molto utili in alcuni studi sugli effetti respiratori, specie nei bambini.^{19,20,23,24} I modelli LUR, oggetto del presente contributo, stimano le concentrazioni di inquinanti in una data posizione

basandosi sul traffico circostante, sull'uso del suolo e altri indicatori correlati (per esempio, i percorsi di origine). Una serie di modelli di questo tipo sono stati sviluppati in Europa e in Nord America, con un notevole successo.²⁵⁻²⁷ E' stata del resto valutata la maggior efficacia dei modelli LUR rispetto a tecniche geostatistiche, quali il *kriging*, nel rappresentare variazioni locali di inquinamento atmosferico.²⁸ Successivamente, sono comparsi nella letteratura internazionale numerosi altri studi con utilizzo di questi modelli, qui non riportati. In Italia, l'unica esperienza disponibile è quella condotta a Roma²⁹ che ha portato allo sviluppo di un modello della concentrazione di NO₂ nella città e a tre applicazioni successive sull'infarto del miocardio,³⁰ i disturbi respiratori negli adulti,³¹ e la funzione polmonare nei bambini.³² Obiettivo del presente lavoro è descrivere il modello LUR e la sua applicabilità all'interno di studi epidemiologici sugli effetti a lungo termine dell'inquinamento sulla salute umana.

L'INQUINAMENTO IN AMBITO URBANO

Senza entrare nei dettagli dei fenomeni fisico-chimici che governano l'inquinamento in ambito urbano, è opportuno accennare, ai fini delle considerazioni modellistiche oggetto del presente contributo, ai fattori che caratterizzano le concentrazioni di inquinanti in area urbana. Come schematizzato in figura 1, a seconda dello scenario urbano si possono identificare più componenti responsabili della concentrazione complessiva di un inquinante. Appare evidente come vi siano componenti che incidono sulla misura quantitativa a prescindere dalla posizione di misura all'interno della città, altre invece che sono fortemente legate al punto di misurazione. Una componente fissa, invariante rispetto allo scenario puntuale, è definita come back-

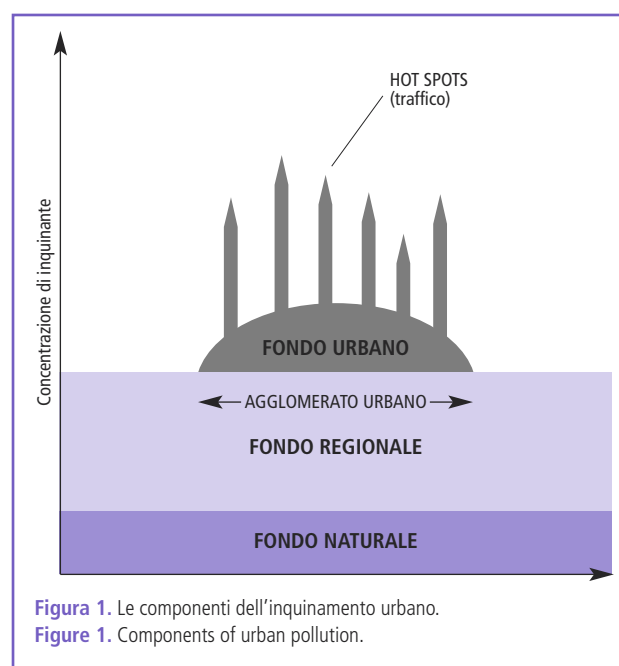


Figura 1. Le componenti dell'inquinamento urbano.
Figure 1. Components of urban pollution.

ground regionale (comprendendo spesso con questo termine anche un valore di base di background emisferico/naturale). La presenza di urbanizzazione porta una componente aggiuntiva di background urbano. In presenza poi di arterie di traffico si hanno i cosiddetti *hot spot*, valori di concentrazione elevati molto localizzati spazialmente, il cui contributo è determinato dal traffico presente su quell'arteria.

Obiettivo della stima dell'entità dell'inquinamento in ambito urbano è la valutazione dell'eterogeneità spaziale presente. Esperienze di monitoraggio e indagini a scala suburbana hanno dimostrato come sovente la variabilità intraurbana sia maggiore di quella fra diverse città, specialmente per alcuni inquinanti strettamente legati alle emissioni da traffico,^{33,34} con importanti ripercussioni nella valutazione dell'esposizione a inquinamento atmosferico nelle indagini epidemiologiche.

MODELLI PER LA VALUTAZIONE DELL'ESPOSIZIONE A INQUINAMENTO URBANO

Vi sono differenti approcci alla modellizzazione dell'inquinamento. Il recente rapido sviluppo dei sistemi informativi geografici (GIS) ha orientato la preferenza su alcuni modelli che fanno un utilizzo sostanziale di dati georeferiti. Seguendo una classificazione proposta da Jerrett et al,²⁸ si possono schematicamente riconoscere 5 approcci:

- proximity models;
- geostatistical interpolators;
- land use regression models;
- dispersion models;
- hybrid models.

Il primo approccio assume la vicinanza di una persona a una sorgente di inquinamento come *proxy* dell'esposizione. E'

stato utilizzato in numerosi studi epidemiologici che associano eventi sanitari alla distanza da una strada, combinando le informazioni della tipologia di strada e/o densità di traffico. In uno studio sui veterani USA, un incremento di sibili persistenti è stato associato alla residenza entro 50 metri dalle principali arterie stradali,³⁵ mentre uno studio tedesco ha associato la residenza in prossimità di strade ad alto traffico a sintomi respiratori, bronchite cronica e rinite allergica.³⁶ Lo studio di coorte SAPALDIA (*Swiss Cohort Study on Air Pollution and Lung Diseases in Adults*) ha trovato un aumento di rischio per disturbi respiratori cronici (catarro e sibili) nei non fumatori residenti entro 20 metri da strade ad alto traffico.³⁷

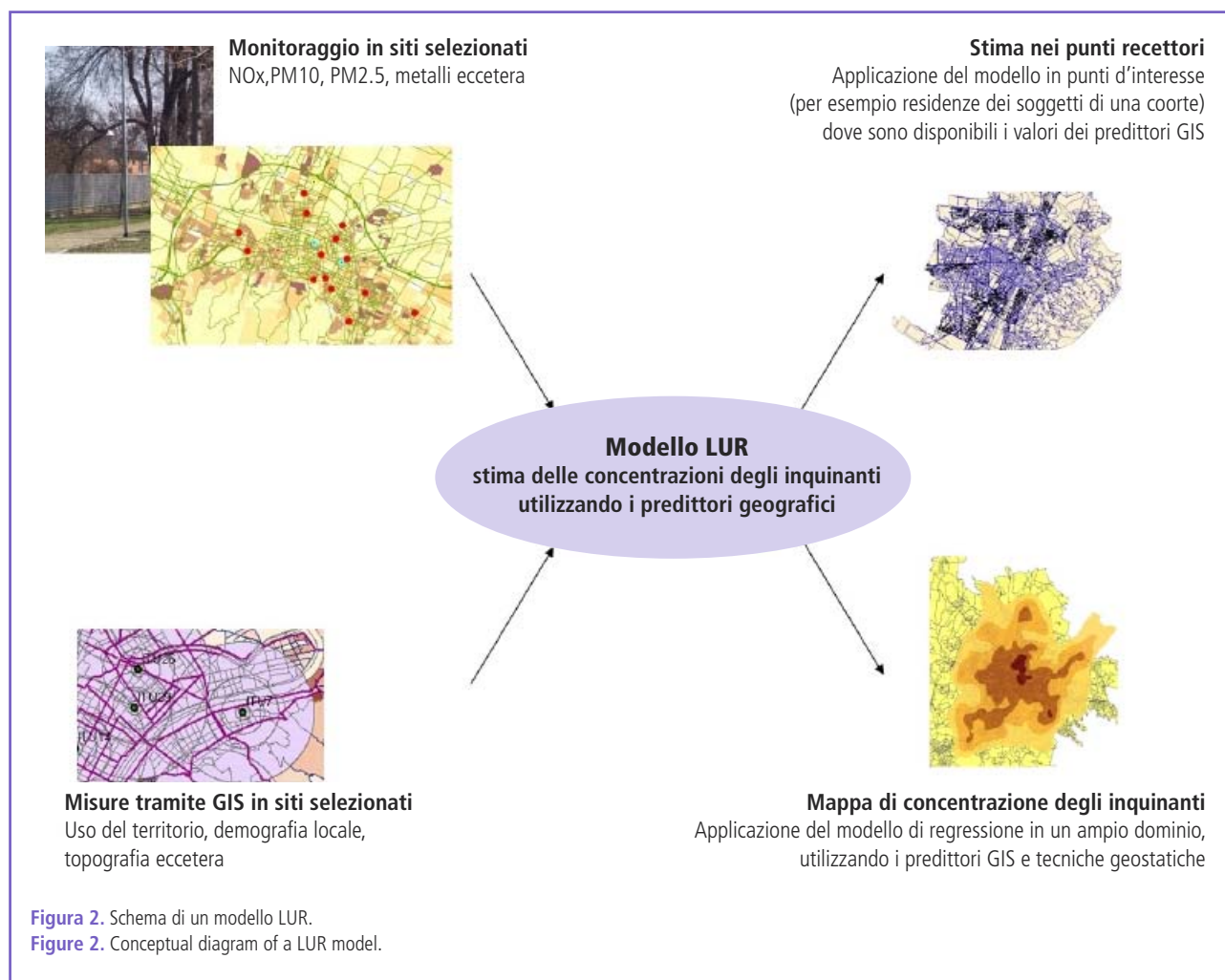
I modelli di interpolazione geostatistica partono da misurazioni ambientali dell'inquinante (per esempio i valori misurati dalle centraline presenti nell'area in studio) e stimano il valore dell'inquinante in punti diversi da quelli misurati e identificati tramite tecniche geostatistiche in modo da creare una superficie continua della concentrazione del contaminante ambientale; la tecnica più usata, anche in applicazioni epidemiologiche, è il *kriging*. Tali applicazioni hanno utilizzato questo metodo di valutazione dell'esposizione in studi sulla mortalità per tutte le cause e causa-specifica^{38,39} e per i disturbi respiratori.^{40,41}

I modelli di dispersione simulano il processo fisico che porta dalle emissioni da parte delle sorgenti inquinanti considerate alle concentrazioni in aria. Utilizzano dati sulle emissioni, dati meteorologici e informazioni relative alla topografia dell'area in studio per stimare le concentrazioni di inquinanti nello spazio, assumendo un processo deterministico alla base della dispersione in aria. Trovano numerose applicazioni in studi sull'inquinamento da sorgenti puntiformi (per esempio inceneritori o industrie), laddove la meteorologia e le condizioni dell'area rendono

MODELLO	UTILIZZO IN STUDI EPIDEMIOLOGICI	VANTAGGI	LIMITI
Proximity models	■ Malattie respiratorie, in particolare sintomi asmatici nei bambini, tumore polmonare e ictus	■ Economici ■ Utili come analisi esplorative	■ Solitamente riferiti a una sola sorgente ■ Possibilità di errori elevati in casi di complessità meteorologica e topografica ■ passibili di recall bias
Geostatistici	■ Studi di mortalità e malattie respiratorie	■ Uso di misure oggettive di inquinamento ■ Misura quantitativa dell'esposizione utile per calcolo di relazioni dose-risposta	■ Numerosità e disposizione centraline spesso non ottimale ■ Costoso per misurazioni ad hoc ■ Assume che la variabilità dell'inquinante dipenda solo da distanza tra i siti
Dispersione	■ Studi multicentrici sugli effetti dell'inquinamento ■ Effetti di sorgenti puntuali quali inceneritori	■ Stima delle variazioni spaziali e temporali dell'inquinamento senza bisogno di misurazioni ad hoc ■ Applicabilità a diverse risoluzioni spaziali (urbana, regionale eccetera)	■ Costoso ■ Necessita di molte informazioni in input ■ Fortemente legato alla qualità dei dati di input
LUR	■ Studi condotti in città europee a livello intraurbano	■ Basso costo (soprattutto per NO₂) ■ Capacità di cogliere variazioni su piccola scala	■ Area-specifico (la generalizzabilità è legata all'omogeneità e all'uso del territorio) ■ Performance e costi legati all'inquinante modellizzato (ideale con NO₂)

Tabella 1. Diversi approcci modellistici per la stima dell'inquinamento in ambito urbano: vantaggi e svantaggi nell'utilizzo in studi epidemiologici.

Table 1. Different modelling approaches to estimate air pollution in urban areas: advantages and limits in their use in epidemiological studies.



l'approssimazione basata sulla distanza fonte di potenziali misclassificazioni dell'esposizione.^{42,43} Numerosi studi epidemiologici utilizzano questi modelli per una valutazione dell'esposizione da sorgenti multiple (puntuali, lineari, diffuse).⁴⁴⁻⁴⁶

Ai modelli *Land Use Regression* (LUR) è interamente dedicato il paragrafo successivo. L'ultimo approccio (*hybrid models*) consiste, invece, nell'associare a uno dei modelli precedenti informazioni relative a monitoraggi personali o diari di attività finalizzata a fornire misure di esposizione personale.

La tabella 1 schematizza i principali vantaggi e limiti nell'utilizzo di tali modelli ai fini della valutazione dell'esposizione in studi epidemiologici.

Come ogni classificazione, anche quella proposta non è sicuramente esaustiva di tutti gli approcci possibili. Un metodo che unisce la filosofia dei modelli di interpolazione geostatistica con l'utilizzo di covariate che correlano con i dati misurati (tipica dell'approccio LUR) è il *co-kriging*, che implementa una procedura di *kriging* con l'utilizzo di covariate geografiche per migliorare la stima dell'inquinamento in presenza di alta variabilità nella distribuzione dell'inquinante.

Assieme ai metodi di regressione, questi sembrano essere gli approcci più promettenti nella modellizzazione spaziale dell'inquinamento atmosferico.⁴⁷

IL MODELLO LAND USE REGRESSION (LUR)

Fra gli approcci modellistici utilizzati per la valutazione dell'inquinamento atmosferico a livello intraurbano, i modelli LUR hanno conosciuto una crescente popolarità nel mondo dell'epidemiologia ambientale in questi ultimi anni per il loro utilizzo ai fini di valutazione dell'esposizione.

Questo approccio combina il monitoraggio dell'inquinamento atmosferico in un limitato numero di siti allo sviluppo di modelli stocastici che utilizzano predittori, solitamente ottenute attraverso sistemi GIS. I modelli LUR utilizzano in sostanza informazioni derivanti da campagne di monitoraggio e analisi geografiche. Per ogni sito di campionamento attivo devono essere disponibili le seguenti informazioni:

- coordinate geografiche;
- media annua di concentrazione dell'inquinante misurato;
- set di valori dei predittori.

Le concentrazioni medie e i valori dei potenziali predittori costituiscono il *dataset* per l'implementazione del modello di regressione lineare. Le equazioni di regressione ottenute vengono poi utilizzate per predire le concentrazioni dell'inquinante in siti non campionati, per i quali siano stati calcolati i valori dei potenziali predittori (figura 2). Come indicato in una recente *review*,⁴⁸ questa tecnica è stata utilizzata con successo in un buon numero di studi, che hanno mostrato come i modelli LUR possano spiegare gran parte della variabilità spaziale nelle concentrazioni di inquinanti.

DATI DI MONITORAGGIO

Solo un limitato numero di studi LUR ha utilizzato i dati provenienti dalle centraline fisse di monitoraggio della qualità dell'aria, mentre nella maggior parte delle applicazioni di questi modelli è stata programmata una campagna di monitoraggio specifica per lo sviluppo del modello stesso. I motivi di questa scelta sono molteplici e sostanzialmente legati alla buona performance del modello da sviluppare: le reti di monitoraggio fisse, infatti, non sono sufficientemente dense all'interno delle realtà urbane da permettere una soddisfacente implementazione di modelli LUR. Inoltre il ricercatore può definire l'ubicazione dei siti di monitoraggio in modo da coprire un'ampia gamma di situazioni differenti di inquinamento (per esempio traffico o background). Gli svantaggi di questo approccio rispetto all'utilizzo delle centraline fisse stanno nei maggiori costi operativi e organizzativi e nella limitata durata dei tempi di monitoraggio, che in genere possono variare dalle 1 alle 4 campagne di 7-14 giorni ciascuna. E' buona norma, comunque, localizzare alcuni campionatori nei siti delle stazioni di monitoraggio fisse, sia ai fini di una validazione delle misure campionarie sia per una possibile relazione del modello con andamenti temporali dell'inquinamento dell'area.

Non ci sono in letteratura metodi rigorosi per definire la numerosità necessaria dei punti di monitoraggio, in funzione delle caratteristiche dell'area e degli obiettivi dello studio. C'è un'ampia variabilità di situazioni presenti in letteratura, da 20 a 100 siti. Alcune indagini sperimentali in questo senso hanno mostrato come il superamento di una numerosità minima utile per lo sviluppo di modelli affidabili sia funzione della densità abitativa, della dimensione dell'area e della presenza o meno di importanti sorgenti di inquinamento industriali.^{49,50}

Anche l'ottimale allocazione delle centraline segue solitamente metodi non formali, finalizzati a massimizzare i contrasti per le variabili che a priori sono considerate come probabili importanti predittori nel modello, tenendo comunque anche conto della distribuzione geografica della popolazione in studio. Per esempio, nello studio TRAPCA, che indagava l'esposizione di una coorte di nati in tre aree dell'Olanda, si è deciso di localizzare 28 dei 40 siti di monitoraggio in punti di background urbano e regionale e 12 in postazioni da traffico, sovrarappresentando così i siti da traffico rispetto

alla distribuzione della popolazione, che si stimava abitasse vicino a strade ad alto traffico per il 5-10% del totale.⁵¹ Esiste un metodo formale per l'ottimale allocazione delle centraline, in funzione della distribuzione della popolazione e della presunta variabilità spaziale dell'inquinamento:²⁶ a partire da un numero definito di centraline, la densità delle postazioni aumenta all'aumentare della popolazione residente e dell'eterogeneità dell'inquinamento.

POTENZIALI PREDITTORI

Partendo dalla considerazione che i potenziali predittori delle variabilità intraurbane dell'inquinamento sono molto legati alle condizioni locali dell'area in studio, si possono comunque identificare nei diversi lavori di letteratura alcune variabili di uso comune, come riportato in tabella 2.⁴⁸

Partendo da un ampio spettro di potenziali predittori, mediante opportune procedure di selezione, viene individuato un sottoinsieme di variabili al fine di ottenere un modello che abbia buone performance e non sia eccessivamente complesso. I valori dei predittori vengono calcolati in un intorno circolare del sito di campionamento, utilizzando le funzioni di buffer tipiche dei sistemi GIS. La dimensione del buffer è fondamentale e può variare a seconda dello studio o della variabile indagata. Idealmente, entro il buffer deve cadere tutto il pattern di dispersione conosciuta dell'inquinamento dovuto alla sorgente in esame. Le caratteristiche peculiari del territorio e delle dinamiche locali dei fenomeni, quali la meteorologia, sono influenti nella potenzialità predittiva di una variabile all'interno di un buffer piuttosto che un altro. Solitamente vengono testati buffer di diverse dimensioni per ogni variabile considerata, partendo da alcune considerazioni a priori sulla natura di ciascun potenziale predittore. Vi sono, per esempio, diversi studi che hanno valutato il decadimento esponenziale dell'inquinamento da traffico con l'aumentare della distanza dalla strada. Una distanza di 100 metri per le strade principali urbane e di 500 metri per le grosse autostrade pare sufficiente a comprendere l'inquinamento da traffico. Se poi si prendono in considerazione le città europee, l'estrema compattezza della configurazione urbana riduce la gran parte della variabilità spaziale dell'in-

TRAFFICO	POPOLAZIONE	USO DEL SUOLO	CARATTERISTICHE DEL TERRITORIO
Lunghezza strade	Densità abitativa	Utilizzo prevalente (urbano, industriale, aree verdi, ...)	Altitudine
Distanza	Abitazioni	Zona della città	Localizzazione geografica
Intensità traffico			Distanza dal mare Zona climatica

Tabella 2. Principali predittori utilizzati nello sviluppo di modelli LUR.

Table 2. Key predictors used in LUR model development.

quinamento dovuto a una singola strada. Non è possibile fare le stesse considerazioni per altri potenziali predittori, legati all'uso del territorio e alla vicinanza o meno a fattori influenti dal punto di vista meteo-climatico (per esempio distanza dal mare), per i quali vengono considerati solitamente anche buffer di raggio dell'ordine dei chilometri.

Questo aspetto apre il problema della precisione della misura legata alla georeferenziazione dei siti, soprattutto quando si considera un fattore di pressione la cui influenza ha un decadimento esponenziale con la distanza, quali le arterie di traffico. Se, infatti, si può ipotizzare che una grande variabilità spaziale si verifichi nell'arco di poche decine di metri, è opportuno disporre di sistemi di georeferenziazione delle arterie stradali e dei civici residenziali più accurati possibile. Questo porta a considerare l'utilizzo di sistemi a elevata accessibilità ed economicità, quali *Google Earth* o *Bing*, utili solo per iniziali indagini esplorative dell'area.

Altra possibile fonte di errore, per il quale non si trova un'analisi in letteratura sui potenziali problemi associati, è la qualità ed eterogeneità dei dati sui flussi di traffico, che sono spesso problematici in termini di accessibilità e di confrontabilità fra realtà comunali differenti, perché raccolti o stimati secondo standard e metodologie diverse a seconda del luogo geografico e del *know-how* della società che li produce.

SVILUPPO DEL MODELLO LUR

La maggior parte degli studi usa tecniche di regressione lineare per sviluppare i modelli LUR, della forma:

$$Z(s) = \alpha + \beta_1 x_{1s} + \beta_2 x_{2s} + \dots + \varepsilon$$

dove

$Z(s)$ = concentrazione dell'inquinante nel sito s

α = intercetta del modello

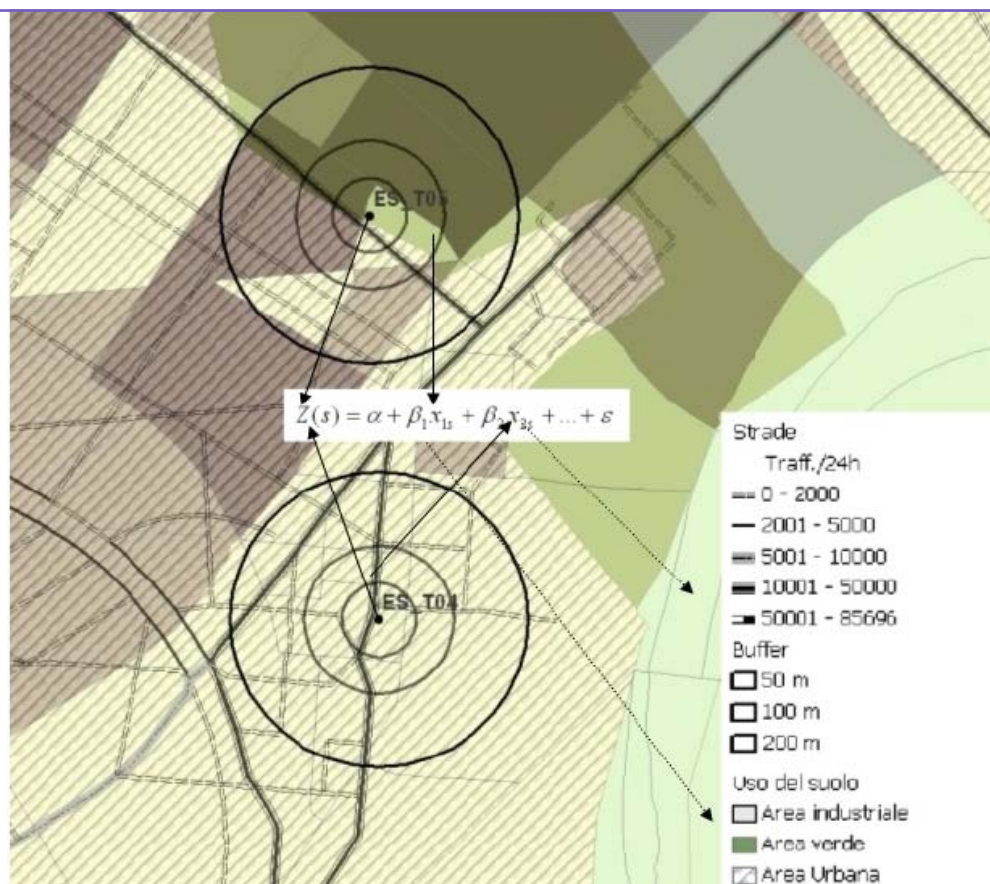
x_{1s} = predittori valutati in ciascun sito s .

Un esempio di tale struttura e di ricavo delle informazioni sui predittori tramite GIS è riportato in figura 3.

Tecniche di selezione automatica dei predittori (per esempio *forward*, *backward stepwise*) permettono di passare da un ampio *dataset* iniziale di potenziali predittori a un modello "parsimonioso", che massimizzi la percentuale di variabilità spiegata (R^2). La struttura composta delle componenti legate all'inquinamento urbano descritte in precedenza ha anche suggerito lo sviluppo di modelli a differente grado di risoluzione spaziale; dapprima vengono valutate le variabili legate all'inquinamento di background, seguite in un secondo momento dai predittori da traffico e legati a scala più locale. Il vantaggio di questo approccio è di introdurre una conoscenza teorica

Figura 3. Elementi di un modello LUR: siti di monitoraggio e predittori calcolati all'interno dei buffer.

Figure 3. Elements of a LUR model: monitoring sites and predictors calculated within buffers.



a priori riguardo ai processi che governano la variabilità spaziale dell'inquinamento, aumentando così l'esportabilità del modello ad altre realtà. In più, si riduce la probabilità di sviluppare modelli poco realistici contenenti predittori autocorrelati. La maggiore limitazione risiede nella necessità di disporre di un numero sufficiente di siti di monitoraggio a differenti scale spaziali, risultando quindi più adatto a studi su aree geografiche ampie, per le quali le variazioni su scala regionale sono significative. In alternativa, si può ragionevolmente assumere che la costante del modello di regressione rappresenti la concentrazione regionale di background (in questo caso considerata invariante).

ANALISI DEL MODELLO

Un aspetto di potenziale violazione degli assunti della regressione lineare è l'indipendenza delle osservazioni. La classica misura indiretta è la valutazione di autocorrelazione dei residui, che, data la struttura dei dati raccolti, è di tipo spaziale. Adottando il punto di vista stocastico, ogni punto dello spazio non ha un solo valore per una proprietà, ma un intero insieme di valori. Il set di valori reali (misurati) dell'inquinante (Z) può considerarsi come una variabile regionalizzata, poiché il suo valore è fortemente condizionato dalla posizione spaziale. Detta $Z(x)$ la variabile regionalizzata nella posizione spaziale x , si può scrivere:

$$Z(x) = R(x) + \alpha$$

dove

$R(x)$ = componente regionalizzata;

α = componente aleatoria (casuale) che tiene conto degli effetti locali.

Esistono strumenti di analisi per la valutazione dell'autocorrelazione spaziale, quali:

- il semivariogramma (plot della semivarianza verso la distanza fra i punti), che permette di valutare gli eventuali pattern di autocorrelazione;
- il *kriging*, che nella sua versione *universal kriging* permette di verificare la precisione (come varianza di stima) in siti dove non si sono effettuate misure sperimentali;
- l'indice I di Moran, una misura quantitativa dell'autocorrelazione spaziale.

Aspetto cruciale del processo di regressione è la validazione del modello, per la quale in letteratura vengono utilizzati diversi approcci. Fra i più diffusi troviamo uno dei metodi di *cross-validation*, il metodo *leave-one-out*, nel quale il modello viene stimato su $n-1$ siti (lasciandone fuori uno alla volta), mantenendo le stesse covariate per tutte le n volte in cui viene stimato; le concentrazioni previste vengono confrontate con quelle misurate nel sito escluso. Il processo viene ripetuto n volte, l'errore di stima totale viene calcolato come media degli

n errori e il livello complessivo di *fitting* fra osservati e predetti su tutti i siti fornisce una misura di performance del modello.

Un metodo alternativo e di diretta implementazione consiste nel dividere il *dataset* originale in due parti: una parte delle osservazioni costituisce il training, la rimanente costituirà il *test set*. Quest'ultimo servirà per verificare la validità del modello tra dati predetti e osservati.

VANTAGGI, LIMITI E POSSIBILI SVILUPPI

I modelli LUR sono stati applicati per creare mappe di NO_2 , NO_x , $\text{PM}_{2.5}$, carbonio elementare, VOCs.

L' NO_2 è l'inquinante più utilizzato per questi modelli per ragioni di economicità e di caratteristiche di dispersione a scala locale, più adatte alla modellizzazione tramite LUR. Per questo inquinante le performance dei modelli, intese come variabilità spiegata (R^2), si attesta intorno allo 0,6-0,7. Le differenze nel valore di R^2 riscontrabili nei diversi studi sono riconducibili alla variabilità nelle misure di concentrazione, all'approccio modellistico, alla qualità dei predittori utilizzati e alla complessità del contesto urbano, soprattutto a livello topografico.

Alcuni studi hanno confrontato i risultati dei modelli LUR con altri metodi più comuni, quali tecniche geostatistiche o modelli di dispersione. In generale, questi modelli hanno fornito performance migliori dei metodi geostatistici e comparabili a quelli dei modelli di dispersione.⁴⁸

A fronte di simili caratteristiche, vanno evidenziati anche alcuni limiti e considerazioni, in relazione soprattutto all'utilizzo in studi epidemiologici:

- la difficoltà nel separare l'impatto di alcuni inquinanti principali: modelli fatti sulla medesima area per NO_2 e $\text{PM}_{2.5}$ hanno mostrato correlazioni elevatissime (0,95), a fronte di correlazioni dei dati misurati alte ma inferiori (0,64-0,80);
- la stima è riferita alla concentrazione in aria a livello di vicino residenziale; i predittori non permettono di considerare fattori legati all'*indoor*;
- i modelli di dispersione, essendo modelli di simulazione, possono stimare concentrazioni dovute a una particolare sorgente (per esempio traffico o riscaldamento); i modelli LUR stimano solo le concentrazioni totali;
- le concentrazioni sono riferite a medie annue: per certi esiti (per esempio effetti riproduttivi) è necessario integrare i dati con informazioni che permettano una considerazione sulla variabilità temporale, dell'ordine dei mesi;
- l'utilizzo di alcuni predittori è potenzialmente associato a fattori legati agli esiti di interesse di uno studio epidemiologico: è il caso della densità di popolazione che può risultare associata allo stato socioeconomico; bisogna tenerne conto nella fase di analisi dello studio epidemiologico, al fine di evitare una distorsione nell'associazione con la malattia di interesse;
- un aspetto messo in luce da un recente articolo⁵² riguarda

la relazione fra la riduzione di errore nella stima dell'esposizione e il miglioramento nelle stime degli effetti sanitari, mostrando, attraverso uno studio di simulazione, che questo assunto non sempre risulta essere vero.

Infine, è opportuno sottolineare brevemente i possibili sviluppi nell'implementazione dei modelli LUR, soprattutto i miglioramenti utili ai fini di un utilizzo in studi epidemiologici.

L'utilizzo di funzioni più flessibili nel modello di regressione può permettere l'introduzione strutture spaziali nelle covariate e studiarne un'associazione anche non lineare. Altro aspetto, già oggetto di approfondimenti e lavori in letteratura, è l'integrazione dei modelli LUR con altri, quali i modelli di dispersione. Le ipotesi ad oggi avanzate sono di un utilizzo dei modelli di dispersione a integrazione del LUR in siti poco caratterizzati da situazioni rappresentate dai predittori utilizzati (per esempio

ampie zone rurali), di un uso a fini di validazione del modello LUR e della possibilità di utilizzare l'informazione dei modelli di dispersione come covariata all'interno di questo.

Un ultimo aspetto, forse più importante ai fini epidemiologici, riguarda l'introduzione nel modello di covariate che permettano di tenere conto della componente temporale, per poter applicare i modelli anche a studi sugli effetti a breve termine dell'inquinamento.

Conflitti di interesse: nessuno

Note: Il presente lavoro è tratto dalla tesina di Master universitario di II livello in epidemiologia dell'Università degli Studi di Torino – Biennio 2009-2010. Si ringraziano Francesco Forastiere, Paolo Lauriola, Rossella Miglio ed Enrica Migliore per il loro contributo alla stesura dell'elaborato, in qualità di relatori o tutor della tesina di Master.

BIBLIOGRAFIA

1. Biggeri A, Bellini P, Terracini B; Italian MISA-Group. Meta-analysis of the Italian studies on short-term effects of air pollution. *Epidemiol Prev* 2001;25(2) Suppl:1-71.
2. Biggeri A, Bellini P, Terracini B. Meta-analysis of the Italian studies on short-term effects of air pollution – MISA 1996-2002. *Epidemiol Prev* 2004;28(4-5) Suppl:4-100.
3. Berti G, Galassi C, Faustini A, Forastiere F. EPIAIR – Air pollution and health, epidemiological surveillance and prevention. *Epidemiol Prev* 2009;33(6) Suppl:1-143.
4. Pope CA, Thun MJ, Namboodiri MM et al. Particulate air pollution as a predictor of mortality in a prospective study of US adults. *Am J Respir Crit Care Med* 1995;151(3 Pt 1):669-74.
5. Krewski D, Burnett RT, Goldberg MS et al. *Reanalysis of the Harvard Six Cities and the American Cancer Society Study of Particulate Air Pollution and Mortality, Phase II: Sensitivity Analysis*. Cambridge (MA), Health Effects Institute, 2000.
6. Jerrett M, Burnett RT, Willis A et al. Spatial analysis of the air pollution-mortality relationship in the context of ecologic confounders. *J Toxicol Environ Health A* 2003;66(16-19):1735-77.
7. Pope CA, Burnett RT, Thun MJ et al. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *JAMA* 2002;287(9):1132-41.
8. Pope CA, Burnett RT, Thurston GD et al. Cardiovascular mortality and long-term exposure to particulate air pollution: epidemiological evidence of general pathophysiological pathways of disease. *Circulation* 2004;109(1):71-7.
9. Dockery DW, Pope CA, Xu X et al. An association between air pollution and mortality in six U.S. cities. *N Engl J Med* 1993;329(24):1753-9.
10. WHO. *Strategy on Air Quality and Health*. Geneva, World Health Organization, 2001.
11. Briggs DJ, de Hoogh C, Gulliver J et al. A regression-based method for mapping traffic-related air pollution: application and testing in four contrasting urban environments. *Sci Total Environ* 2000;253(1-3):151-67.
12. Zhu Y, Hinds WC, Kim S, Sioutas C. Study of ultrafine particles near a major highway with heavy-duty diesel traffic. *Atmos Environ* 2002;36(27):4323-35.
13. Jerrett M, Burnett RT, Kanaroglou P et al. A GIS – environmental justice analysis of particulate air pollution in Hamilton, Canada. *Environment and Planning A* 2001;33(6):955-73.
14. Brauer M, Hoek G, van Vliet P et al. Estimating long-term average particulate air pollution concentrations: application of traffic indicators and geographic information systems. *Epidemiology* 2003;14(2):228-39.
15. Brunekreef B, Holgate S. Air pollution and health. *Lancet* 2002;360(9341):1233-42.
16. Thomas DC, Stram D, Dwyer J. Exposure measurement error: influence on exposure-disease. Relationships and methods of correction. *Annu Rev Public Health* 1993;14:69-93.
17. Künzli N, Jerrett M, Mack WJ et al. Ambient air pollution and atherosclerosis in Los Angeles. *Environ Health Perspect* 2005;113(2):201-6.
18. Hoek G, Brunekreef B, Goldbohm S, Fischer P, van den Brandt PA. Association between mortality and indicators of traffic-related air pollution in the Netherlands: a cohort study. *Lancet* 2002;360(9341):1203-9.
19. van Vliet P, Knappe M, de Hartog J, Janssen N, Harssema H, Brunekreef B. Motor vehicle exhaust and chronic respiratory symptoms in children living near freeways. *Environ Res* 1997;74(2):122-32.
20. Zmirou D, Gauvin S, Pin I et al. Traffic related air pollution and incidence of childhood asthma: results of the Vesta case-control study. *J Epidemiol Comm Health* 2004;58(1):18-23.
21. Schaub J, Iversen T, Pedersen U. Comparison of measurements and model results for airborne sulfur and nitrogen components with Kriging. *Atmos Environ* 1993;27(6):831-44.
22. Briggs DJ, Collins S, Elliott P et al. Mapping urban air pollution using GIS: a regression-based approach. *Int J Geogr Inform Sci* 1997;11(7):699-718.
23. English P, Neutra R, Scalf R, Sullivan M, Waller L, Zhu L. Examining associations between childhood asthma and traffic flow using a geographic information system. *Environ Health Perspect* 1999;107(9):761-7.
24. Lin S, Munsie JP, Hwang SA, Fitzgerald E, Cayo MR. Childhood asthma hospitalization and residential exposure to state route traffic. *Environ Res* 2002;88:73-81.
25. Hoek G, Fischer P, Van Den Brandt P, Goldbohm S, Brunekreef B. Estimation of long-term average exposure to outdoor air pollution for a cohort study on mortality. *J Expo Anal Environ Epidemiol* 2001;11(6):459-69.
26. Kanaroglou PS, Jerrett M, Morrison J et al. Establishing an air pollution monitoring network for intra-urban population exposure assessment: a location-allocation approach. *Atmos Environ* 2005;39(13):2399-409.
27. Gilbert NL, Goldberg MS, Beckerman B, Brook JR, Jerrett M. Assessing spatial variability of ambient nitrogen dioxide in Montreal,

- Canada, with a land-use regression model. *J Air Waste Manage Assoc* 2005;55(6):1059-63.
28. Jerrett M, Arain A, Kanaroglou P et al. A review and evaluation of intra-urban air pollution exposure models. *J Expo Anal Environ Epidemiol* 2005;15(2):185-204.
29. Rosenlund M, Forastiere F, Stafoggia M et al. Comparison of regression models with land-use and emissions data to predict the spatial distribution of traffic-related air pollution in Rome. *J Expo Sci Environ Epidemiol* 2008;18(2):192-9.
30. Rosenlund M, Picciotto S, Forastiere F, Stafoggia M, Perucci CA. Traffic-related air pollution in relation to incidence and prognosis of coronary heart disease. *Epidemiology* 2008;19(1):121-8.
31. Cesaroni G, Badaloni C, Porta D, Forastiere F, Perucci CA. Comparison between various indices of exposure to traffic-related air pollution and their impact on respiratory health in adults. *Occup Environ Med* 2008;65(10):683-90.
32. Rosenlund M, Forastiere F, Porta D, De Sario M, Badaloni C, Perucci CA. Traffic-related air pollution in relation to respiratory symptoms, allergic sensitisation and lung function in schoolchildren. *Thorax* 2009;64(7):573-80.
33. Gilliland F, Avol E, Kinney P et al. Air pollution exposure assessment for epidemiologic studies of pregnant women and children: lessons learned from the Centers for Children's Environmental Health and Disease Prevention Research. *Environ Health Perspect* 2005;113(10):1447-54.
34. Stein AF, Isakov V, Godowitch J, Draxler RR. A hybrid modeling approach to resolve pollutant concentrations in urban area. *Atmos Environ* 2007;41(40):9410-26.
35. Garshick E, Laden F, Hart JE, Caron A. Residence near a major road and respiratory symptoms in U.S. Veterans. *Epidemiology* 2003;14:728-36.
36. Heinrich J, Topp R, Gehring U, Thefeld W. Traffic at residential address, respiratory health, and atopy in adults: the National German Health Survey 1998. *Environ Res* 2005;98:240-9.
37. Bayer-Oglesby L, Schindler C, Hazenkamp-von Arx ME et al; SAPAL-DIA team. Living near main streets and respiratory symptoms in adults: the Swiss Cohort Study on Air Pollution and Lung Diseases in Adults. *Am J Epidemiol* 2006;164(12):1190-8.
38. Abbey DE, Nishino N, McDonnell WF et al. Long-term inhalable particles and other air pollutants related to mortality in nonsmokers. *Am J Respir Crit Care Med* 1999;159(2):373-82.
39. Finkelstein MM, Jerrett M, DeLuca P et al. Relation between income, air pollution and mortality: a cohort study. *CMAJ* 2003;169:397-402.
40. Pikhart H, Bobak M, Gorynski P et al. Outdoor sulphur dioxide and respiratory symptoms in Czech and Polish school children: a small-area study (SAVIAH). Small-Area Variation in Air Pollution and Health. *Int Arch Occup Environ Health* 2001;74(8):574-8.
41. Mulholland JA, Butler AJ, Wilkinson JG, Russell AG, Tolbert PE. Temporal and spatial distributions of ozone in Atlanta: regulatory and epidemiologic implications. *J Air Waste Manage Assoc* 1998;48:418-26.
42. Floret N, Mauny F, Challier B, Arveux P, Cahn JY, Viel JF. Dioxin emissions from a solid waste incinerator and risk of non-Hodgkin lymphoma. *Epidemiology* 2003;14(4):392-8.
43. Viel JF, Daniau C, Gorja S et al. Risk for non Hodgkin's lymphoma in the vicinity of French municipal solid waste incinerators. *Environ Health* 2008;7:51.
44. Bartonova A, Clench-Aas J, Gram F et al. Air pollution exposure monitoring and estimation. Part V. Traffic exposure in adults. *J Environ Monit* 1999;1(4):337-40.
45. Nyberg F, Gustavsson P, Jarup L et al. Urban air pollution and lung cancer in Stockholm. *Epidemiology* 2000;11(5):487-95.
46. Bellander T, Berglund N, Gustavsson P et al. Using geographic information systems to assess individual historical exposure to air pollution from traffic and house heating in Stockholm. *Environ Health Perspect* 2001;109(6):633-9.
47. Nieuwenhuijsen MJ, Brunekreef B. Environmental Exposure Assessment. In: Baker D, Nieuwenhuijsen MJ (eds.) *Environmental Epidemiology – Study methods and Application*. New York, Oxford University Press, 2008.
48. Hoek G, Beelen R, de Hoogh K et al. A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmos Environ* 2008;42(33):7561-78.
49. Madsen C, Lodrup-Carlsen KC, Hoek G et al. Modeling the intra-urban variability of outdoor traffic pollution in Oslo, Norway – A GA² LEN project. *Atmos Environ* 2007;41:7500-11.
50. Jerrett M, Arain MA, Kanaroglou P et al. Modeling the intraurban variability of ambient traffic pollution in Toronto, Canada. *J Toxicol Environ Health Part A* 2007;70(3-4):200-12.
51. Brauer M, Hoek G, van Vliet P et al. Estimating long-term average particulate air pollution concentrations: application of traffic indicators and geographic information systems. *Epidemiology* 2003;14(2):228-39.
52. Szpiro AA, Paciorek CJ, Sheppard L. Does more accurate exposure prediction necessarily improve health effect estimates? *Epidemiol* 2011;22(5):680-5.

ABBONAMENTI 2012

A CIASCUNO IL SUO

	E&P on line + Suppl on-line	E&P on line + Suppl on line + versione cartacea	E&P on line + versione cartacea + supplementi cartacei
PRIVATI ITALIA			
1 anno	70 euro	80 euro	95 euro
2 anni	130 euro	150 euro	180 euro
3 anni	185 euro	210 euro	255 euro
ENTI ITALIA AD ACCESSO UNICO			
ENTI ITALIA AD ACCESSO MULTIPLO: ABBONAMENTI DA CONCORDARE CON L'EDITORE			
1 anno	145 euro	155 euro	170 euro
2 anni	270 euro	280 euro	310 euro
3 anni	385 euro	395 euro	440 euro
ENTI ESTERO			
1 anno	165 euro	175 euro	195 euro
2 anni	290 euro	310 euro	350 euro
3 anni	405 euro	425 euro	475 euro
PRIVATI ESTERO			
1 anno	85 euro	95 euro	115 euro
2 anni	160 euro	180 euro	220 euro
3 anni	230 euro	260 euro	320 euro

PROMOZIONI 2012

- **Per giovani epidemiologi:** abbonamento on line a 45 euro per gli under 30.
- **Per generosi epidemiologi già abbonati a E&P:** regala un abbonamento a E&P per il 2012.
Costa solo 50 euro per l'edizione on line e 60 euro per avere anche il cartaceo.
Ovviamente, l'abbonamento sarà accompagnato da un biglietto che svelerà l'identità del donatore per fare una gran bella figura e nello stesso tempo aiutare E&P.
- **Per epidemiologi "contagiosi":** se ti piace E&P e fai sottoscrivere due nuovi abbonamenti a chi non conosce la rivista o non è più abbonato da almeno due anni, il tuo abbonamento o il tuo rinnovo è gratuito.

EPIDEMIOLOGIA & PREVENZIONE

Modalità di abbonamento per il 2012

data Abbonamento annuo a partire dal primo numero raggiungibile:

☐ **Tipo di abbonamento** ☐ **euro**

Modalità di pagamento:

☐ **Versamento:** a mezzo conto corrente postale n. 55195440
intestato a Inferenze scrl, via Ricciarelli 29, 20148 Milano
(allegare la ricevuta di versamento alla richiesta di abbonamento)

☐ **Assegno:** intestato a Inferenze scrl

☐ **Bonifico bancario:** UGF BANCA, piazza Wagner 8, 20145 Milano
IBAN IT 53 P 03127 01600 0000 0000 3681 intestato a Inferenze scrl,
via Ricciarelli 29, 20148 Milano (allegare la contabile alla richiesta di abbonamento)

☐ **PayPal:** sul sito www.epiprev.it

☐ **Carta di credito:** ☐ American Express ☐ Carta Si ☐ Master Card ☐ Eurocard ☐ VISA

cognome e nome

azienda

indirizzo

cap località prov.

tel. fax. e-mail

numero scadenza / / firma

cod. CV2 _ _ _ (ultime tre cifre stampate sul retro della carta, per una garanzia di sicurezza in più)

Compilare e inviare a Inferenze - via Ricciarelli 29, 20148 Milano; e-mail abbonamenti@inferenze.it o per fax allo 02 48706089



Epidemiologia&Prevenzione in versione "classica"
di 64 pagine, con peer-review,
indicizzazione in medline e impact factor
è quella che trovate **integralmente on-line**

Da questa versione vengono estratte 32 pagine a **stampa**
(per gli affezionati della carta)



EPdiMezzo è la nuova parte di EP dedicata ai dibattiti,
esce a cavallo tra un numero e l'altro di E&P,
viene pubblicata solo on-line,
non è impaginata (non trovate i PDF), non è sottoposta
a peer-review e non viene indicizzata,
ma avvia i dibattiti che poi troveranno posto
anche su E&P, vi permette di lasciare commenti
in tempo reale e favorisce il dialogo anche con soggetti
diversi dagli epidemiologi



EP Channel

IL CALENDARIO DI *ep*

GENNAIO	FEBBRAIO	MARZO	APRILE	MAGGIO	GIUGNO	LUGLIO	AGOSTO	SETTEMBRE	OTTOBRE	NOVEMBRE	DICEMBRE
	1/2011		2/2011		3/2011			4/2011		5/2011	
E&P DI MEZZO 	E&P ON LINE 64 pp 	E&P DI MEZZO 	E&P ON LINE 64 pp 	E&P DI MEZZO 	E&P ON LINE 64 pp 	E&P DI MEZZO 		E&P ON LINE 64 pp 	E&P DI MEZZO 	E&P ON LINE 64 pp 	
	 E&P STAMPA 32 pp		 E&P STAMPA 32 pp		 E&P STAMPA 32 pp			 E&P STAMPA 32 pp		 E&P STAMPA 32 pp	