



BIG data per la valutazione degli Effetti sanitari dell'inquinamento atmosferico nella Popolazione Italiana

Progetto BIGEPI

OBIETTIVO SPECIFICO 1:

Valutare gli effetti acuti dell'esposizione ambientale su tutto il territorio nazionale

**Attività 1.4. Analisi di dati per la valutazione degli effetti acuti delle esposizioni
ambientali sulla mortalità causa-specifica**

PROTOCOLLO

Ultima versione: 30/09/2021

Descrizione attività:

- a) Raccolta dei dati giornalieri di mortalità per cause specifiche (non-accidentali, cardiovascolari, respiratorie, neurologiche, mentali) per ogni comune del territorio italiano nel periodo 2006-2015
- b) Costruzione delle serie giornaliere comunali dei dati di concentrazione di PM₁₀, PM_{2.5}, NO₂ ed O₃, stimati a partire da modelli di “machine learning” a risoluzione 1-km²
- c) Costruzione delle serie giornaliere comunali dei dati di temperatura dell’aria stimati a partire da modelli di regressione lineare ad effetti misti a risoluzione 1-km²
- d) Definizione degli aspetti operativi per la valutazione degli effetti degli inquinanti atmosferici e della temperatura dell’aria sulla mortalità causa-specifica per tutto il territorio nazionale e per categorie individuali/ecologiche differenti quali classe di età, sesso e tipologia di comune (urbano, suburbano, rurale)

Modalità di esecuzione attività

- a) Raccolta dei dati nazionali di mortalità (fonte ISTAT) relativi al periodo 2006-2015. Per ogni individuo, informazioni disponibili sull’età, il sesso, il comune di residenza, il comune di decesso, la data di decesso, e la causa di morte (in codice International Classification of Diseases-10th Version [ICD-10]). A partire da questi dati, costruzione delle serie giornaliere comunali di decesso per le cause naturali, cardiovascolari, respiratorie, metaboliche, diabete, neurologiche e psichiatriche. Le stesse serie giornaliere sono state anche prodotte separatamente per classe di età (0-64, 65-74, 75-84, 85+ anni) e per sesso. Infine, ogni comune è stato classificato in base all’indice “DEGURBA” (Degree of Urbanization, fonte EUROSTAT) in “area metropolitana e città maggiore”, “città minore”, e “area sub-urbana e rurale”.
- b) Per ognuno degli 8,092 comuni italiani (al Censimento 2011) sono state costruite le serie giornaliere di concentrazioni di inquinanti atmosferici (valori medi giornalieri), relativamente ai periodi 2006-2015 (PM₁₀) e 2013-2015 (PM_{2.5}, NO₂, O₃). In particolare, il grigliato fisso 1x1-km degli inquinanti è stato intersecato allo shapefile dei comuni italiani, e per ogni comune sono state calcolate le concentrazioni giornaliere di inquinanti mediando sulle celle di intersezione. Al fine di ottenere delle medie pesate per popolazione, sono stati utilizzati due pesi nel calcolo delle medie: peso proporzionale all’area di intersezione (maggior peso alle celle interamente contenute nel comune, minore peso a quelle sul confine), e peso proporzionale alla popolazione residente nella cella (maggior peso alle celle con più popolazione residente)
- c) In modo del tutto speculare a quanto presentato al precedente punto b), sono state calcolate le medie giornaliere di temperatura dell’aria per ogni comune italiano ed ogni giorno del periodo 2006-2015
- d) Le analisi statistiche dei dati, finalizzate alla valutazione degli effetti acuti di inquinamento e temperatura su mortalità causa-specifica, sono tuttora in corso. Si sta adottando il seguente **protocollo operativo**:

Esiti sanitari

Conte giornaliero di decessi per cause:

- Naturali: A00-R99
 - o Cardiovascolari: I00-I99
 - Cardiache: I00-I52
 - Ischemiche del cuore: I20-I25
 - Cerebrovascolari: I60-I69
 - o Respiratorie: J00-J99
 - o Mentali: F00-F99
 - o Nervose: G00-G99
 - o Metaboliche: E00-E99
 - Diabete: E10-E14

Esposizioni ambientali

- Inquinanti atmosferici:
 - o Concentrazioni giornaliere di PM10 per il periodo 2006-2015
 - o Concentrazioni giornaliere di PM2.5, NO2 ed O3 per il periodo 2013-2015
- Temperatura: temperatura dell'aria media giornaliera per il periodo 2006-2015

Altri confondenti

Vengono definite altre variabili giornaliere ai fini dell'aggiustamento del confondimento nelle analisi epidemiologiche. Esse includono: a) le epidemie influenzali (su base regionale e settimanale); b) decrementi estivi di popolazione; c) festività; d) trend temporali di lungo periodo e stagionali.

Metodi statistici

La relazione tra gli esiti sanitari identificati e le variabili di esposizione viene analizzata attraverso modelli di regressione multipla di Poisson con sovra-dispersione, secondo la seguente formula:

$$\log E [Y_t] = \beta_0 + f(E_{sp_t}) + \text{confondenti}_t$$

dove $E[Y_t]$ rappresenta il valore atteso del conteggio di decessi (causa-specifico) nel giorno t , E_{sp_t} rappresenta l'esposizione di interesse al giorno t (concentrazione di inquinante o temperatura dell'aria), f è un'opportuna funzione dell'esposizione (termine lineare, spline, ecc.), e "confondenti_t" include una lista estesa di fattori di confondimento al giorno t , come descritto di seguito.

Fattori di confondimento

I fattori di confondimento sono scelti sulla base di analisi esplorative ed evidenze di letteratura. Includono:

- *Trend temporale*. Fattore finalizzato a rimuovere il confondimento legato ai trend di lungo periodo e stagionali. Esso viene modellato con una interazione a 4 livelli tra anno, mese, giorno della settimana e comune (analogo al case-crossover "time-stratified")
- *Temperatura (nell'analisi degli inquinanti)*. Termine finalizzato a rimuovere il confondimento legato agli effetti avversi delle temperature basse invernali e delle temperature elevate estive. A tal proposito viene modellato con curve non lineari: due spline separate per la temperatura al lag 0-1 e la temperatura al lag 1-6 (come da articoli EPIAIR e MED-PARTICLES)
- *Inquinante (nell'analisi delle temperature)*. Termine finalizzato a rimuovere il confondimento legato agli effetti avversi dell'inquinamento atmosferico. A tal proposito

l'inquinante scelto è il PM10 (l'unico disponibile per l'intero periodo 2006-2015), che viene modellato con un termine lineare al lag 0-5.

- *Epidemie influenzali, decrementi estivi di popolazione e festività*: vengono modellati con variabili indicatrici

Struttura lag della variabile di esposizione

Nel caso dell'inquinamento atmosferico, viene condotta una analisi esplorativa utilizzando lag scelti a priori per rappresentare effetti immediati (lag 0-1), ritardati (lag 2-5) e prolungati (lag 0-5), e tali variabili vengono messe nel modello come termini lineari. Le stime di associazione vengono scalate su 10 µg/m³ ed espresse come incrementi percentuali di rischio (e relativi intervalli di confidenza al 95%). A fini di analisi di sensibilità, le stime vengono riportate anche per incrementi interquartili (IQR) degli inquinanti.

Modificazione di effetto

Tutti gli aspetti precedentemente descritti verranno replicati per stimare effetti separatamente per: classi di età (0-64, 65-74 e 75+), genere, tipologia del comune (rurale/sub-urbano, città minori, città maggiori/aree metropolitane).

Tempistica di esecuzione attività:

Il presente protocollo e la costruzione del dataset sono stati finalizzati a Ottobre 2021. Le analisi dei dati sono tutt'ora in corso, in fase avanzata di elaborazione.