



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

# **Progetto BIGEPI**

### **OBIETTIVO SPECIFICO 3:**

**Valutare gli effetti cronici dell'esposizione ambientale negli studi longitudinali  
italiani**

**Deliverable 3.1**

**Attività 3.1, 3.2**

**Protocollo operativo per la raccolta e l'analisi di dati per la valutazione degli effetti cronici delle esposizioni ambientali sulla mortalità causa-specifica e sull'incidenza di patologie cardiovascolari in 6 studi longitudinali italiani**

---

*Ultima versione: 09/03/2022*

### Descrizione attività:

- a) Protocollo di raccolta dei dati per la costruzione dei 6 studi longitudinali (di Roma, Torino, Bologna, Brindisi, Taranto e Siracusa)
- b) Protocollo per l'assegnazione delle variabili di esposizione ambientale all'indirizzo di residenza dei membri delle coorti
- c) Definizione degli aspetti operativi per la valutazione degli effetti cronici degli inquinanti atmosferici e della temperatura dell'aria sulla mortalità causa-specifica e sull'incidenza di specifiche patologie cardiovascolari nel totale dei 6 studi longitudinali, e per sottogruppi di popolazione identificati dalle classi di età, dal genere, dallo stato civile, dal livello di istruzione, dall'occupazione, e dallo stato socio-economico di area

### Modalità di esecuzione attività

- a) Per ognuna delle 6 aree in studio si seguiranno i seguenti passaggi operativi:
  - o Linkage tra anagrafe residenti (ove possibile, altrimenti anagrafe sanitaria) e Censimento 2011
  - o Esclusione di soggetti con residenze atipiche (case di riposo, RSA, ecc.), senza tetto, dati mancanti per le variabili individuali e di area
  - o Esclusione di soggetti di età < 30 anni al baseline
  - o Ove possibile, esclusione dei soggetti che risiedono all'indirizzo baseline di residenza da meno di un anno
  - o Raccolta dati individuali su: età, sesso, cittadinanza (italiana si/no), luogo di nascita (nella città dello studio si/no), stato civile, livello di istruzione (da Censimento), occupazione (da Censimento)
  - o Raccolta dati di area su: indice di deprivazione materiale e sociale (per sezione di censimento, anno 2011), prevalenza di disoccupati, prevalenza di scarsamente istruiti, prezzo medio degli appartamenti
  - o Aggiornamento stato in vita (mortalità, emigrazione, ecc.) fino all'ultimo anno utile, possibilmente il 2019 (e comunque non oltre il 2019)
  - o Linkage con dati di mortalità per causa, e con schede di dimissione ospedaliera al fine di definire gli esiti di mortalità causa-specifica e di incidenza di patologie cardiovascolari (ischemiche e cerebrovascolari)
- b) Per ognuno dei 6 studi longitudinali, si assegnerà l'esposizione ai diversi fattori ambientali come segue:
  - o Linkage geografico tra lo shapefile del grigliato 1x1-km con i dati di esposizione, e lo shapefile degli indirizzi
  - o Media annua di PM10 (anno 2011), di PM2.5, di NO2 (anno 2013)
  - o Media stagione calda (Aprile-Settembre) di O3 (anno 2013)
  - o Media e DS annua 2011, media e DS Aprile-Settembre 2011 e media e DS Gennaio-Marzo + Ottobre-Dicembre 2011 della temperatura dell'aria.

- c) Gli effetti cronici degli inquinanti atmosferici e della temperatura sugli esiti sanitari verranno indagati utilizzando modelli multivariati a rischi proporzionali di Cox, con l'età come asse temporale, ed aggiustamento per fattori di confondimento individuali e di area. Ogni esito verrà analizzato separatamente. Ogni esposizione verrà innanzitutto analizzata separatamente, quindi si valuterà la possibilità di implementare modelli a due inquinanti, o modelli che integrino sia la temperatura che un inquinante. Infine, si valuterà la modificazione di effetto per caratteristiche individuali e di area al fine di identificare sottogruppi maggiormente suscettibili. Tutte le analisi verranno condotte all'interno di ciascuno studio longitudinale sulla base di script condivisi. Infine, i risultati coorte-specifici verranno meta-analizzati. I dettagli del **protocollo di analisi** sono riportati di seguito

### Esiti sanitari

Decesso per cause:

- Naturali: 0-799
  - o Cardiovascolari: 390-459
  - o Respiratorie: 460-519

Incidenza di:

- Malattie ischemiche: 410-414
- Malattie cerebrovascolari: 430-438

L'incidenza delle patologie in esame verrà definita escludendo tutti i soggetti che, nei 5 anni precedenti (o periodo inferiore se 5 anni non sono disponibili) abbiano avuto almeno un ricovero per le suddette patologie (casi prevalenti al baseline) e considerando come data dell'evento la data del primo ricovero ospedaliero per la patologia in esame. Se i dati degli anni precedenti il baseline non sono disponibili, si analizzerà il primo ricovero tra tutti i soggetti della coorte (senza esclusione dei prevalenti).

### Esposizioni ambientali

Si veda il precedente punto b).

### Confondenti individuali

Età (asse temporale nel modello di Cox), sesso (termine di stratificazione, vedere successivamente), cittadinanza (italiana si/no), luogo di nascita (nella città si/no), stato civile, livello di istruzione, occupazione (variabili categoriche).

### Confondenti di area

Indice di deprivazione socio-economica (quintili, variabile categorica), tasso di disoccupazione, livello di istruzione, prezzo medio degli appartamenti (quintili, variabile categorica).

## Metodi statistici

La relazione tra gli esiti sanitari identificati e le variabili di esposizione viene analizzata attraverso modelli di regressione multipla di Cox con l'età utilizzata come asse temporale, e livelli crescenti di aggiustamento:

- *Modello 1.* Aggiustato solo per età (asse temporale) e strata (sesso)
- *Modello 2.* Modello aggiustato per tutte le variabili individuali: età (asse temporale), sesso (termine "strata"), cittadinanza, luogo di nascita, stato civile, livello di istruzione, occupazione (variabili categoriche)
- *Modello 3.* Modello 2 + aggiustamento per variabili di area [es. indice di deprivazione materiale e sociale (quintili, variabile categorica)]

### *Modificazione di effetto*

Tutti gli aspetti precedentemente descritti verranno replicati per stimare effetti separatamente per: classi di età (30-64, 65-74 e 75+), genere, livello di istruzione (4 livelli), occupazione (3 livelli) e indice di deprivazione materiale e sociale (quintili).

### *Modelli a due esposizioni*

Si valuterà l'opportunità di inserire nello stesso modello:

- PM10 e NO2
- PM2.5 e NO2
- PM10 e O3
- PM2.5 e O3
- NO2 e O3
- Temperatura media annua e PM10
- Temperatura media annua e PM2.5
- Temperatura media annua e NO2
- Temperatura media annua e O3

### *Meta-analisi*

I risultati di tutti i modelli sopra descritti verranno quindi raccolti in una meta-analisi ad effetti fissi al fine di stimare associazioni "pooled" dei 6 studi longitudinali

### *Tempistica di esecuzione attività:*

Il presente protocollo e la costruzione di 3 dataset (Roma, Torino e Bologna) sono stati finalizzati a Ottobre 2021. Gli altri 3 dataset (Brindisi, Taranto e Siracusa) verranno ragionevolmente finalizzati entro Marzo 2022. Le analisi dei dati sono partite a Febbraio per la città di Roma, e verranno ragionevolmente finalizzate entro Giugno 2022.