

Aer si sanatate in Bucuresti

Raport exploratoriu cu metodologia, formulele, rezultatele si limitele. Versiune bazata exclusiv pe artefacte validate.

Rezumat

Acest document prezinta o analiza ecologica exploratorie a seriilor anuale de poluare si a indicatorilor medicali agregati. El nu estimeaza expunerea individuala si nu demonstreaza cauzalitate. Analiza principala foloseste statiile B-1 pana la B-6 din Bucuresti, selectie fixata inainte de rezultate.

Date si unitatea de analiza

Poluantii PM10, PM2.5 si NO2 provin din trei exporturi de mediu pentru 2014-2024. Indicatorii HOSP, FADM si MORT provin din fisierul medical pentru 2014-2023. Unitatea de observatie este anul calendaristic. Pentru fiecare pereche se pastreaza numai anii in care ambele serii au valoare.

Agregare

Instrumente -> medie pe statie si zi -> medie zilnica a statiilor selectate -> medie anuala. Valorile lipsa sunt excluse si contabilizate, nu sunt inlocuite cu zero. Numarul de statii si acoperirea lor pot varia intre ani.

Formule

Media anuala: $\bar{x} = (1 / D) \sum d$, unde D este numarul zilelor cu media retelei disponibila.

Covarianta de esantion: $Cov(X,Y) = \sum (x_i - \bar{x})(y_i - \bar{y}) / (n - 1)$.

Corelatia Pearson: $r = Cov(X,Y) / (s_X s_Y)$. Testul Pearson este bilateral si foloseste numai anii comuni.

Benjamini-Hochberg: dupa sortarea $p_{(k)}$, pragul este $k \times 0,05 / 171$. q este cel mai mic nivel FDR la care o pereche ar fi declarata.

Rezultate principale

Pentru PM10 x MORT_03 in selectia principala: $r = 0,8275669459285317$, $p = 0,0031235328001662866$, $q = 0,12604150168385167$, $n = 10$ si covarianta = 218,4736783417537. Valoarea este descriptiva; nu trece pragul BH de $q < 0,05$.

Selectie	Teste Pearson	p < 0,05 nominal	q < 0,05
B-1 pana la B-6 (principala)	171	33	0
Bucuresti toate statiile (sensibilitate)	171	-	3

De ce corectia BH nu declara o descoperire

Sunt 171 teste: 19 randuri medicale x 3 indicatori x 3 poluanti. La pragul nominal 0,05, hazardul ar produce in medie 8,55 rezultate sub 0,05 daca toate ipotezele nule ar fi adevarate. Cea mai mica valoare p observata este 0,0018855, mai mare decat pragul BH de rang 1 (0,0002924). De aceea, analiza principala are 0 din 171 la $q < 0,05$. Aceasta nu inseamna ca nu exista legatura; seriile de 7-10 ani au putere foarte mica dupa corectie.

Limite

Doar 7 din 30 de statii au coordonate documentate. Pentru restul exista cel mult identificator, nume sau adresa textuala. Datele medicale nu furnizeaza numitori de populatie, cauza confirmata a decesului, diagnostic, stadiu, histologie, rezidenta, varsta, sex ori deduplicare intre grupe. Nu sunt disponibile covariabilele necesare pentru controlul confuziei.

Lag-uri si analiza bayesiana

Nu se calculeaza lag-uri: un lag anual ar reduce seriile deja scurte, iar o intarziere aleasa dupa rezultate ar fi arbitrara. Latența cancerului este mai lunga si mai heterogena decat un simplu lag anual. O analiza bayesiana poate fi doar exploratorie acum: distributii posterioare pentru r sau o panta standardizata cu priors slabe. Nu poate transforma datele agregate incomplete intr-o estimare cauzala sau individuala.

Concluzie

Acest raport documenteaza o asociere descriptiva intre serii agregate. Pentru un studiu etiologic sunt necesare date RNMCA complete cu QA/QC si coordonate, date medicale cu definitii si numitori, covariabile predefinite si protocol analitic stabilit inaintea analizelor.

Trasabilitate

Manifest verificare: 41 verificari trecute; moment UTC: 2026-09-17T19:07:19.660194+00:00. Detaliile de fisiere, hash-uri si verificari sunt in anexa tehnica separata.