ISSN 2282-0930 • Epidemiology Biostatistics and Public Health - XIII th SISMEC Congress - Vol. 20 Suppl. 1

Comparative Evaluation of Linear, Log-Concentration, and SCHIF Exposure–Response Functions for Estimating Attributable Deaths from Air Pollution: A Simulation Study

Smimmo Annafrancesca[(1)], Schiattarella Paola[(1)], Speranza Teresa[(1)], Di Gennaro Piergiacomo[(1)], Nicolao Giovanni[(1)], Fordellone Mario[(1)], Simeon Vittorio[(1)], Stafoggia Massimo[(2)], Ranzi Andrea[(3)], Bonvicini Laura[(4)], Broccoli Serena[(5)], Caranci Nicola[(6)], Zengarini Nicolás[(7)], Signoriello Simona[(1)], Chiodini Paolo[(1)]

(1) Unità di Statistica Medica, Università degli Studi della Campania "Luigi Vanvitelli"
(2) Dipartimento Epidemiologia del S.S.R. - ASL Roma1 Regione Lazio
(3) Unità Epidemiologia Ambientale, Arpae – Agenzia Prevenzione Ambiente Energia dell'Emilia-Romagna
(4) Servizio di Epidemiologia – Azienda USL di Reggio Emilia
(5) Settore Prevenzione collettiva e Sanità pubblica, Regione Emilia-Romagna
(6) Settore innovazione nei Servizi sanitari e sociali, Regione Emilia-Romagna
(7) S.C. Servizio Sovrazonale di Epidemiologia ASL TO3, Regione Piemonte

CORRESPONDING AUTHOR: *Smimmo Annafrancesca, annafrancesca.smimmo@unicampania.it*

INTRODUCTION

World Health Organization (WHO) has estimated that 4.2 million premature deaths were attributable to ambient (outdoor) air pollution in 2019 [1,2]. The effectiveness of regulatory actions aimed to improve air quality are frequently evaluated by forecasting the specific impacts that these measures will have on public health outcomes, such as reductions in hospital admissions or morbidity and mortality rates, following their implementation. Mainly, the regulatory actions act on long-term exposure to pollution. As consequence, the most common measure in this context is the calculus of attributable deaths (ADs) to levels of air pollution. ΔDs, can be expressed mathematically using the following equation:

$$\Delta D = M_0 \left(1 - \frac{1}{R(\beta, \Delta z)}\right) \times pop$$

where M_0 is the baseline mortality rate, Δz is the predicted or observed change in ambient concentrations (z - z_f), *pop* is the size of the target population, and $R(\beta, \Delta z)$ is the relative risk function of a vector of unknown parameters β. In this context, $R(\beta, \Delta z)$ denotes the ratio of the probability of an adverse event occurring over a fixed period for a population exposed to z compared to the probability if the same population were instead exposed to z_f and it is also called exposure-response function. Generally, it is used ERF as linear shape. Nevertheless, in the last years, several studies have shown that this association is better presented as non-linear form [3,4]. One of the first proposal for non-linear ERF consisted of considering the log of pollutant's concentration, but in recent year new non-linear ERFs were proposed. In particular, Shape Constrained Health Impact Function (SCHIF) proposed by Nasari et al. in 2016, has the specific aim to be used for health impact assessment [5]. This model is stated on a sigmoidal relative risk shape that appropriately describe the hypothetical bound between air pollution and health's risk.

OBJECTIVE

The aim of this work is to compare these three ERFs (linear, log-linear and SCHIF) in the calculus of ADs through a simulation study, when the relationship between air pollution and mortality is supposed to have a sigmoidal shape.

METHODS

Firstly, a large cohort study (n = 100,000) was generated. For each subject two covariates were considered $PM_{2.5}$ ~ N(15, 2.5) and age ~ N(52, 5), aiming to reproduce a realistic scenario [4,5]. Survival times were generated using a parametric model based on an exponential distribution for both censored and uncensored observations. For the censored observations, a risk of λ_c = 0.00035 was set. For the uncensored observations, a risk of λ_T = 0.0012 was used, along with the following coefficients: β = 0.03 for $PM_{2.5}$ and γ = 0.07 for age. The coefficient β was weighted by a logistic

DOI: 10.54103/2282-0930/29532

weighting function to obtain a sigmoidal shape. Eleven scenarios were evaluated, varying the location parameter (μ) from the 5th to the 50th percentile of the distribution of z and one based on pure logarithm shape. Therefore, ADs were calculated for each scenario, using five different ERFs: linear (L), log-concentration (Log), optimal SCHIF (O), ensemble of best three SCHIF (E3B) and ensemble of all models SCHIF (EA). The European population aged 30+ was set as population; the counterfactual scenario was based on the WHO 2021 Air Quality Guidelines ($PM_{2.5}$ < 5 μg/m³ per year) and the estimated β from the five different ERFs for $PM_{2.5}$ was considered. The percentage change of each ERF from the attributable deaths simulated (considered as true) was reported as the median of the 1.000 replications.

RESULTS

We observed that the performance of each ERF varies over every considered setting. O and E3B models appear to be more stable than other ERFs (Δ% 0 to 15 and Δ% 5 to 17, respectively) in settings based on a sigmoidal shape. If the curvature point in sigmoidal shapes happens at lower concentrations, the L model remains an acceptable approximation of the true shape, Δ% -39 to 9 (setting 5 to 20). SCHIF O and E3B in these setting overestimate deaths, but no more than 15%. In the same scenarios the L and Log model reaches 84% and 174% of overestimates in ADs, respectively. On the other hand, the setting based on pure logarithmic shape both O and E3B models have worse performance compared to L and Log models. EA model underestimates ADs in each setting.

CONCLUSIONS

This simulation study highlights the importance of selecting an appropriate ERF when estimating ADs to air pollution. When the true risk relationship is sigmoidal, SCHIF models, O and E3O, provide more accurate estimates than L and Log models, particularly when the curve inflects at higher pollutant concentrations, while it lacks accuracy in detecting correctly ADs in case of pure logarithmic shape. Linear model remain reasonable in context when the shape inflects at low concentrations or in pure logarithmic shape. In conclusion, the adoption of non-linear models represents a significant advancement toward more accurate health impact assessments of air pollution, but more complex scenarios should be evaluated to give robustness to SCHIF results.

ACKNOWLEDGMENTS

This project was carried out with the technical support and funding of the Italian Ministry of Health – PNC PREV-A-2022-12376981 Aria outdoor e salute: un atlante integrato a supporto delle decisioni e della ricerca.

REFERENCES

1. WHO global air quality guidelines: particulate matter ($PM_{2.5}$ and PM_{10}), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. https://www.who.int/publications/i/item/9789240034228 (last accessed May 2024)
2. Burnett R., Chen H., Szyszkowicz M., et al. Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. Proc Natl Acad Sci U S A. 2018 Sep 18;115(38):9592-9597.
3. Burnett R., Pope C. 3rd, Ezzati M., et al. An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure. Environ Health Perspect. 2014 Apr;122(4):397-403.
4. Nasari M.M., Szyszkowicz M., Chen H. et al. A class of non-linear exposure-response models suitable for health impact assessment applicable to large cohort studies of ambient air pollution. Air Qual Atmos Health, 2016, 9, 961–972.
5. Cesaroni G., Forastiere F., Stafoggia M., et al. Long term exposure to ambient air pollution and incidence of acute coronary events: prospective cohort study and meta-analysis in 11 European cohorts from the ESCAPE. Project BMJ, 2014; 348 :f7412

DOI: 10.54103/2282-0930/29532

Table 1. Attributable deaths are reported for each setting, along with percentage change with respect to simulated (true) deaths (first column). Linear (L), logarithmic (Log), optimal SCHIF (O), best three ensemble SCHIF (E3B), all models' ensemble SCHIF (EA)

Setting	AD	L	Δ(%)	Log	Δ(%)	O	Δ(%)	E3B	Δ(%)	EA	Δ(%)
μ=5	237718	144545	-39%	223818	-6%	238701	0%	249085	5%	137025	-42%
μ=10	226152	186440	-18%	284495	26%	245591	9%	249121	10%	145139	-36%
μ=15	214894	207796	-3%	317079	48%	236554	10%	243606	13%	139837	-35%
μ=20	204116	223069	9%	338523	66%	234393	15%	236927	16%	142806	-30%
μ=25	193489	234385	21%	353349	83%	214574	11%	219462	13%	128863	-33%
μ=30	183023	241854	32%	363895	99%	209881	15%	212982	16%	127838	-30%
μ=35	172879	248262	44%	373684	116%	198342	15%	201901	17%	119199	-31%
μ=40	162825	254148	56%	380610	134%	182839	12%	184422	13%	108112	-34%
μ=45	152727	258234	69%	385618	152%	165088	8%	166401	9%	94293	-38%
μ=50	142272	261652	84%	389836	174%	158699	12%	159890	12%	90244	-37%
Log(z)	102441	63474	-38%	97603	-5%	45762	-55%	46169	-55%	18849	-82%

DOI: 10.54103/2282-0930/29532