Environmental Exposure Assessment and Biomarker-Based Prediction of Acute Respiratory Health Deterioration in Vulnerable Populations Using Multi-Source Air Quality Data

Naoise O'Kane
Department of Computer Science, University of North Texas, Denton, TX, USA.
naoise540@unt.edu

Abstract

Acute respiratory deterioration among vulnerable populations is shaped by short-term environmental exposures and by individual biological susceptibility, yet exposure assessment and biomarker-based prediction are usually developed in separate analytic traditions. This manuscript develops a methodological framework that couples multi-source air quality data with biomarker dynamics to predict near-term respiratory deterioration and to support uncertainty-aware triage. The exposure layer estimates personal exposure from ambient concentration fields, activity-location weights, and microenvironment infiltration factors; concentration fields are fused from monitors, satellite aerosol optical depth, land-use and meteorological predictors. The biomarker layer represents short-lag inflammatory and endothelial responses as a state-space process driven by exposure and clinical covariates. The prediction layer uses an interpretable regularized risk model with explicit exposure-by-biomarker interactions, a group-weighted objective, and ensemble uncertainty. Decision thresholds are derived from explicit false-positive and false-negative costs rather than from a default probability cutoff. Analytical assessment shows that model value depends less on classifier complexity than on exposure misclassification, temporal alignment between exposure and biomarker measurement, calibration for vulnerable strata, and accounting for label delay. The framework provides a coherent basis for evaluating whether biomarker-informed prediction adds actionable information beyond exposure and clinical covariates, while preserving the distinction between proposed methodology and empirical validation. Its significance lies in linking exposure science, inflammatory mechanisms, and predictive decision-making without collapsing them into a single black-box score.

Keywords

environmental exposure assessment; multi-source air quality fusion; respiratory risk prediction; biomarkers; vulnerable populations; uncertainty-aware machine learning

1. Introduction

Ambient air pollution remains one of the most consequential environmental contributors to respiratory morbidity, and current air quality guidance emphasizes fine particulate matter, ozone, nitrogen dioxide, sulfur dioxide, and carbon monoxide as priority pollutants [1]. Global burden estimates indicate that ambient particulate pollution continues to account for a substantial share of premature morbidity and mortality [2], while long-term cohort evidence has established fine particulate exposure as a risk factor for cardiopulmonary mortality [3]. These population-level findings do not imply uniform individual risk. Susceptibility is concentrated among people with pre-existing respiratory disease, older adults, children, and

those with cardiovascular or metabolic comorbidity, as well as among socially disadvantaged groups with limited access to care and poorer housing conditions [4]. The analytical framework for adverse air pollution effects therefore needs to distinguish between exposure contrasts that are sufficient to shift population risk and short-term changes that are clinically meaningful for a vulnerable individual [5].

Mechanistic evidence provides a basis for connecting short-term exposure to acute respiratory deterioration. Inhaled pollutants can provoke oxidative stress, airway inflammation, endothelial activation, and autonomic dysfunction, with downstream effects on respiratory symptoms and cardiopulmonary function [6]. Cardiovascular and inflammatory pathways are not incidental to respiratory risk; they help explain why particulate exposure can precipitate decompensation in people with limited physiological reserve [7]. Controlled and observational studies have linked air pollution exposure with systemic inflammatory markers and endothelial dysfunction [8]. In chronic obstructive pulmonary disease, short-term particulate exposures have been examined alongside biomarkers of systemic inflammation and endothelial activation [9], which supports the plausibility of a biomarker-informed predictive layer while also underscoring the need for careful temporal alignment between exposure, biomarker measurement, and clinical outcome.

The exposure side of the problem has also become more computationally tractable. High-resolution spatiotemporal models can estimate ambient particulate concentrations by combining ground monitors with satellite retrievals, meteorology, and land-use information [10]. Random forest and related ensemble models have improved fine-scale concentration estimation in settings where monitor coverage is sparse [11]. Machine learning is now widely used in geoscientific and remote-sensing settings to fuse heterogeneous predictors with irregular spatial support [12]. In clinical prediction, machine learning can integrate multimodal covariates and improve risk stratification when validation is carefully designed [13]. However, algorithmic risk models can reproduce or amplify inequity when the outcome label, access to care, or feature measurement differs systematically across groups [14]. A respiratory prediction model for vulnerable populations must therefore treat exposure assessment, biomarker interpretation, and decision thresholds as linked design choices rather than independent modules.

This manuscript addresses that integration problem. The central contribution is a transparent prediction framework in which multi-source air quality data are converted into personal exposure estimates, short-lag biomarker dynamics are modeled as a state process, and acute respiratory deterioration is predicted with an interpretable risk model that includes exposure-by-biomarker interactions, group-aware regularization, and cost-sensitive decision rules. The framework is methodological and analytical: it specifies the variables, mechanisms, assumptions, and validation logic needed to test whether biomarker-informed prediction adds actionable information beyond exposure and clinical covariates. It does not report a completed field study or claim empirical performance. Its purpose is to make the inferential dependencies explicit so that future evaluation can distinguish genuine predictive signal from artifacts of exposure misclassification, temporal misalignment, and inequitable label generation.

2. Exposure and Biomarker Framework

2.1 Personal exposure and vulnerable-population stratification

Vulnerable populations should not be treated as a residual category defined only by age or diagnosis. For respiratory deterioration, vulnerability reflects the joint presence of elevated exposure, limited physiological reserve, and contextual factors that shape whether a short-term exposure becomes an acute care event. Satellite aerosol optical depth can improve spatial coverage where monitors are sparse [15], and cohort evidence shows that long-term exposure gradients are associated with mortality even at relatively low concentrations [16]. Metabolic comorbidity such as diabetes has been associated with increased susceptibility to particulate health effects [17], and reviews of epidemiological evidence reinforce the importance of exposure timing, concentration-response shape, and susceptible subgroups [18]. These findings motivate an exposure estimate that varies across individuals and days rather than assigning a single city-level concentration to everyone.

Let C(s(k), t) denote the estimated ambient concentration of a pollutant or pollutant mixture at location s(k) on day t, where k indexes the microenvironments occupied by individual i. Let w(i, k, t) be the fraction of time individual i spends in microenvironment k on day t, and let f(k) be the infiltration factor that maps ambient concentration to microenvironmental concentration. Personal exposure is represented as

$$E_{i,t} = \sum_k w_{i,k,t} \cdot f_k \cdot C_{sk,t}$$

where the sum runs over k = 1, ..., K, w(i, k, t) is nonnegative and sums to one across microenvironments, f(k) lies between zero and one for indoor environments and equals one for outdoor exposure, and C(s(k), t) is the fused ambient concentration at the relevant location. The expression is deliberately simple: it makes explicit that personal exposure error can arise from concentration estimation, activity weighting, and infiltration assumptions. It also prevents a common analytical error in which a model with strong ambient concentration prediction is assumed to have equally strong personal exposure prediction.

Stratification should be based on clinical diagnosis, recent exacerbation history, age, medication use, and social context, but these variables should not be collapsed into a single uninterpretable susceptibility score. A defensible approach is to define prespecified vulnerable strata, fit the prediction model with group-aware regularization, and report calibration and decision consequences for each stratum. The purpose is not to build separate models for every subgroup, which would fragment information, but to ensure that the shared model does not perform well only for the easiest-to-measure or most frequently observed group.

2.2 Multi-source concentration fusion

Multi-source air quality data differ in spatial support, temporal resolution, error structure, and missingness. Regulatory monitors provide relatively accurate point measurements but sparse coverage; satellite retrievals provide broad spatial coverage but are affected by cloud cover, retrieval error, and vertical distribution; land-use and meteorological predictors explain persistent spatial gradients but may not capture transient emissions. A practical fusion estimator combines source-specific concentration predictions through weights that reflect their conditional accuracy. Let $\hat{C}(s, t)$ be the fused concentration, and let $\hat{C}(m)(s, t)$ be the prediction from source or model m. The fusion layer is

$$C_{s,t} = \sum_m \omega_m \cdot C_{s,t}^{(m)}$$

where m = 1, ..., M indexes monitor-based interpolation, satellite-informed estimation, land-use regression, or meteorological adjustment outputs, and ω(m) are nonnegative weights

that sum to one. The weights should be estimated or specified using out-of-sample performance within spatial blocks and time periods, not random cross-validation alone. If monitor data are scarce, the weights can be regularized toward equal weighting or toward a physically motivated source hierarchy. The fused estimate should carry an uncertainty estimate because downstream risk prediction depends on exposure uncertainty, not only on the point estimate.

The choice of fusion model should be justified relative to credible alternatives. A purely mechanistic chemical transport model offers interpretability and spatial continuity but may require emission inventories that are unavailable or delayed. A purely data-driven ensemble can capture complex interactions but may overfit monitor-specific patterns and perform poorly in underrepresented neighborhoods. A hybrid approach that uses machine learning for residual correction while retaining physically interpretable spatial and meteorological terms is often preferable for regulatory and clinical translation [10]. The key design condition is that the fusion model be evaluated on spatial holdouts and time-forward partitions that resemble deployment, since random splits can overstate performance when nearby monitors share unmodeled local emissions [11].

2.3 Biomarker dynamics and temporal alignment

Biomarkers can inform prediction only if their dynamics are modeled with appropriate lag structure. Inflammatory markers such as C-reactive protein, interleukin-6, fibrinogen, and endothelial adhesion molecules respond to exposure over hours to days, and the relevant lag may differ by marker and clinical phenotype [8]. A state-space formulation is useful because it separates the latent biological state from measurement error and because it can accommodate irregular clinical sampling. Let B(i, t) be a vector of biomarkers for individual i on day t, E(i, t) be the personal exposure estimate, and Z(i, t) be clinical and demographic covariates. A linear state transition is

$$B_{i,t} = \Phi B_{i,t-1} + \Lambda E_{i,t} + \Gamma Z_{i,t} + \eta_{i,t}$$

where Φ is the state-transition matrix, Λ is the exposure-loading vector or matrix, Γ is the covariate-loading matrix, and η(i, t) is a stochastic disturbance term. The equation is a modeling proposal rather than an estimated biological law. Its value is that it forces explicit assumptions about persistence, exposure lag, and covariate adjustment. If biomarker measurements are available only during clinic visits, the model should be interpreted as a partially observed process, with observation times informative about health status. That informativeness can induce selection bias if sicker patients are measured more often.

Temporal alignment is a central failure mode. Exposure on day t should not be paired with a biomarker measured later without accounting for the possibility that the biomarker reflects an earlier exposure or an acute infection. Similarly, an acute respiratory event should not be labeled using a hospitalization date if the biological deterioration began several days earlier. The prediction horizon must be chosen before model fitting, because a model trained to predict same-day deterioration answers a different clinical question from one trained to predict deterioration within seven days. For triage, a seven-day horizon may be more actionable, but it also increases the chance that intervening events dominate the signal. The framework therefore treats horizon, exposure lag, and biomarker lag as explicit design parameters to be prespecified and reported.

3. Prediction Model and Decision Logic

3.1 Risk formulation with exposure-by-biomarker interaction

The prediction target is an acute respiratory deterioration event within a prespecified horizon H, defined operationally as an emergency department visit, hospitalization, unscheduled respiratory care, or a clinically documented escalation in rescue therapy. The risk model should be interpretable enough to show whether exposure and biomarkers contribute independently and jointly. A risk formulation with explicit interaction is

$$logit(p_{i,t}) = \beta_0 + \beta_E\, E_{i,t} + \beta_B^{\,T} B_{i,t} + \beta_Z^{\,T} Z_{i,t} + \sum_j \gamma_j E_{i,t} B_{i,j,t}$$

where p(i, t) is the probability of deterioration within H days, logit(p) = log(p/(1 − p)), E(i, t) is personal exposure, B(i, t) is the biomarker vector with components B(i, j, t), Z(i, t) is a covariate vector including age, diagnosis, medication, previous exacerbations, and socioeconomic indicators, and β and γ are coefficients. The interaction term allows the effect of exposure to vary with biomarker level. A positive γ(j) would imply that the same exposure increment is associated with a larger log-odds of deterioration when biomarker j is elevated; a negative value would imply buffering or a different biological state. This formulation is preferable to a black-box model when the purpose is to test a mechanistic hypothesis, because it separates main effects from interaction and permits partial dependence to be reported.

The model should be compared with three baselines: an exposure-only model, a clinical-covariate-only model, and a biomarker-only model. The comparison answers whether multi-source exposure assessment adds information beyond existing clinical predictors, and whether biomarker measurement adds information beyond exposure and clinical covariates. A model that improves discrimination but worsens calibration for high-risk strata is not necessarily an improvement. The relevant performance measures therefore include discrimination, calibration slope and intercept, net benefit across plausible thresholds, and subgroup-specific false-negative rates.

3.2 Regularized learning objective and group-aware validation

The risk model should be estimated with an objective that reflects both predictive accuracy and the consequences of unequal measurement. A regularized objective with a group penalty is

where θ denotes model parameters, N is the number of person-days, w(i) is a prespecified weight that can upweight vulnerable strata or downweight groups with highly uncertain labels, L is a log-loss or another proper scoring rule, λ controls ridge regularization, μ controls the fairness penalty, FPR(g) is the false-positive rate in group g, and FPR(ref) is a reference false-positive rate or target level. The penalty is not a guarantee of fairness; it is a transparent device for exposing trade-offs. If the penalty is too large, the model may sacrifice sensitivity for the most vulnerable group to equalize false-positive rates. If it is too small, the model may achieve good average performance while missing the patients for whom the intervention is most needed.

Validation must mimic deployment. Random splits can leak information when repeated measures from the same person, nearby locations, or the same seasonal period appear in both training and test sets. Temporal forward-chaining validation, spatial block validation, and patient-level holdouts are all necessary in different ways. The reported result should state which form of validation supports which claim. Internal calibration does not establish transportability to a new city, and a model trained on monitored neighborhoods may not

perform well in communities where monitor coverage is thin and satellite retrievals are less reliable.

3.3 Ensemble uncertainty and cost-sensitive thresholds

Prediction should produce a distribution or at least a calibrated uncertainty estimate, because a point probability near a decision threshold is not the same as a confident classification. Ensemble learning provides a practical approximation. Let $\hat{p}(m)(i, t)$ be the predicted probability from ensemble member m. The ensemble mean and variance can be represented as

$$\hat{p}_{i,t} = (1/M)\sum_m \hat{p}_{i,t}^{(m)}; \ \hat{\sigma}_{i,t}^2 = (1/M)\sum_m (\hat{p}_{i,t}^{(m)} - \hat{p}_{i,t})^2$$

where M is the number of ensemble members or bootstrap models, $\hat{p}(i, t)$ is the ensemble mean risk, and $\sigma^2(i, t)$ is the ensemble variance. The variance is not a full predictive distribution, but it can flag unstable predictions for review. A practical triage rule can use both the mean risk and an upper predictive bound, so that a patient with high uncertainty is not automatically treated as low risk simply because the mean probability is below threshold.

The decision threshold should be derived from consequences rather than fixed at 0.5. Let C_FP be the cost of a false-positive action, such as unnecessary preventive treatment or additional monitoring, and C_FN be the cost of a false-negative event, such as delayed care for a deterioration that could have been prevented. The expected-cost threshold for taking action is

$$Treat\ if\ \hat{p}_{i,t} \geq C_FP/(C_FN + C_FP)$$

where the threshold increases as false-positive costs rise and decreases as false-negative costs rise. The threshold is not a clinical constant; it should be justified for the specific action, population, and care setting. A model that is well calibrated but paired with an inappropriate threshold can still cause harm. For this reason, decision-curve analysis and subgroup-specific net benefit are more informative than discrimination alone.

4. Technical Discussion

4.1 Error modes, sensitivity, and confounding

The dominant error mode is likely to be exposure misclassification rather than classifier architecture. Ambient concentration estimates can be accurate at the monitor but wrong at the residence because of local sources, indoor infiltration, and activity patterns. Satellite-informed estimates improve spatial coverage [15], but retrieval error and vertical mixing can introduce temporally structured error. If exposure error is classical and independent, it tends to attenuate associations; if it is differential with respect to health status, it can create or reverse apparent effects. For example, people with worsening respiratory symptoms may spend more time indoors, reducing personal exposure at exactly the time when ambient concentration is high. The activity-weight formulation in the exposure layer makes this dependence explicit, but it cannot eliminate it without individual-level time-location data.

Biomarker measurement error and confounding present a second challenge. Inflammatory markers vary with infection, medication, time of day, and comorbidity. A model that predicts deterioration from biomarkers alone may be capturing an incipient infection rather than an air pollution effect. The exposure-by-biomarker interaction helps only if the exposure and biomarker are measured with sufficient temporal resolution and if the model is compared with clinical baselines. Large cohort analyses show that exposure contrasts of public health

relevance can be modest in magnitude [19], so small systematic errors in exposure or outcome ascertainment can materially change the estimated risk gradient.

Label generation is a third error mode. Acute respiratory deterioration is often ascertained through claims, emergency department records, or hospital admissions. Access to care, insurance, and neighborhood resources affect whether a deterioration is recorded. If vulnerable groups have less access to care, a model trained on recorded events may underestimate their risk or learn to predict utilization rather than biology. Group-aware validation and explicit reporting of label definitions are therefore not optional refinements; they are part of the causal interpretation of the model.

4.2 Implementation conditions and equity implications

Implementation would require a data pipeline with explicit lineage for monitor data, satellite products, meteorological fields, land-use predictors, clinical covariates, and biomarker measurements. Each source should carry a timestamp, spatial support, and missingness indicator. Imputation should be separated from model fitting so that imputation uncertainty can be propagated. The model should be evaluated in a silent deployment period before any triage use, with prospective monitoring of calibration drift and subgroup performance. A model that performs well in one season may fail in another because of changes in pollution sources, respiratory virus circulation, or care-seeking behavior.

Equity considerations extend beyond the fairness penalty in the learning objective. A biomarker-informed model may be more accurate for patients who already receive frequent biomarker testing and less accurate for patients who do not. If the model is used to allocate preventive outreach, this difference could widen existing disparities. A defensible implementation would require that the model be evaluated with a prespecified protocol for missing biomarkers, that performance be reported for groups defined by age, diagnosis, and socioeconomic position, and that the decision threshold be justified in relation to the capacity and cost of the proposed action. The model should be treated as a decision aid that combines exposure and biomarker information, not as a substitute for clinical judgment.

The framework also clarifies what would count as evidence of added value. A biomarker-informed model would add value if it improved calibration and net benefit relative to an exposure-plus-clinical model in a prospective or temporally held-out evaluation. It would not add value merely by increasing the number of predictors or by achieving a higher area under the curve in a random split. The distinction matters because multi-source air quality data and biomarker panels are expensive to maintain; their justified use depends on whether they change decisions and outcomes, not on whether they produce a more complex risk score.

5. Conclusion

This manuscript has developed a methodological framework for environmental exposure assessment and biomarker-based prediction of acute respiratory deterioration in vulnerable populations using multi-source air quality data. The framework links three normally separate steps: personal exposure estimation from fused ambient concentration fields, state-space modeling of short-lag inflammatory and endothelial biomarkers, and an interpretable risk model with explicit exposure-by-biomarker interaction. It also specifies a group-aware learning objective, ensemble uncertainty, and a cost-sensitive decision rule. The analytical contribution is not an empirical performance claim but a coherent specification of the assumptions, error modes, and validation conditions that determine whether biomarker-informed prediction is clinically useful.

The main implications are cautious but substantive. Multi-source air quality fusion can improve exposure coverage, yet personal exposure remains vulnerable to activity, infiltration, and temporal misalignment. Biomarkers can encode biological susceptibility, yet their interpretation depends on lag structure, measurement error, and confounding by infection and treatment. Predictive performance depends on calibration and decision consequences, not only on discrimination. The most defensible use of such a framework is to test, with prospective and subgroup-specific evaluation, whether exposure and biomarker information improve net benefit for the populations at greatest risk. That standard keeps the model accountable to the environmental and clinical mechanisms it is meant to represent.

References

1. World Health Organization. (2021). WHO global air quality guidelines: Particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. World Health Organization.

2. Cohen, A. J., Brauer, M., Burnett, R., Anderson, H. R., Frostad, J., Estep, K., ... Forouzanfar, M. H. (2017). Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: An analysis of data from the Global Burden of Diseases Study 2015. The Lancet, 389(10082), 1907–1918.

3. Pope, C. A., III, Burnett, R. T., Thun, M. J., Calle, E. E., Krewski, D., Ito, K., & Thurston, G. D. (2002). Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. JAMA, 287(9), 1132–1141.

4. Sacks, J. D., Stanek, L. W., Luben, T. J., Johns, D. O., Buckley, B. J., Brown, J. S., & Ross, M. (2011). Particulate matter-induced health effects: Who is susceptible? Environmental Health Perspectives, 119(4), 446–454.

5. Thurston, G. D., Kipen, H., Annesi-Maesano, I., Balmes, J., Brook, R. D., Cromar, K., ... Brunekreef, B. (2017). A joint ERS/ATS policy statement: What constitutes an adverse health effect of air pollution? An analytical framework. European Respiratory Journal, 49(1), 1600419.

6. Kelly, F. J., & Fussell, J. C. (2015). Air pollution and public health: Emerging hazards and improved understanding of risk. Environmental Geochemistry and Health, 37(4), 631–649.

7. Brook, R. D., Rajagopalan, S., Pope, C. A., III, Brook, J. R., Bhatnagar, A., Diez-Roux, A. V., ... Kaufman, J. D. (2010). Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. Circulation, 121(21), 2331–2378.

8. Chuang, K. J., Chan, C. C., Su, T. C., Lee, C. T., & Tang, C. S. (2007). The effect of urban air pollution on inflammation, oxidative stress, coagulation, and autonomic dysfunction in young adults. American Journal of Respiratory and Critical Care Medicine, 176(4), 370–376.

9. Huang, S., Garshick, E., Vieira, C. L., Grady, S. T., Schwartz, J. D., Coull, B. A., ... & Koutrakis, P. (2020). Short-term exposures to particulate matter gamma radiation activities and biomarkers of systemic inflammation and endothelial activation in COPD patients. Environmental research, 180, 108841.

10. Di, Q., Kloog, I., Koutrakis, P., Lyapustin, A., Wang, Y., & Schwartz, J. (2016). Assessing PM2.5 exposures with high spatiotemporal resolution across the continental United States. Environmental Science & Technology, 50(9), 4712–4721.

11. Hu, X., Belle, J. H., Meng, X., Wildani, A., Waller, L. A., Strickland, M. J., & Liu, Y. (2017). Estimating PM2.5 concentrations in the conterminous United States using the random forest approach. Environmental Science & Technology, 51(12), 6936–6944.

12. Lary, D. J., Alavi, A. H., Gandomi, A. H., & Walker, A. L. (2016). Machine learning in geosciences and remote sensing. Geoscience Frontiers, 7(1), 3–10.

13. Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. New England Journal of Medicine, 380(14), 1347–1358.

14. Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. Science, 366(6464), 447–453.

15. Just, A. C., Wright, R. O., Schwartz, J., Coull, B. A., Baccarelli, A. A., Tellez-Rojo, M. M., ... Kloog, I. (2015). Using high-resolution satellite aerosol optical depth to estimate daily PM2.5 geographical distribution in Mexico City. Environmental Science & Technology, 49(14), 8576–8584.

16. Dockery, D. W., Pope, C. A., Xu, X., Spengler, J. D., Ware, J. H., Fay, M. E., ... Speizer, F. E. (1993). An association between air pollution and mortality in six U.S. cities. New England Journal of Medicine, 329(24), 1753–1759.

17. Zanobetti, A., & Schwartz, J. (2001). Are diabetics more susceptible to the health effects of airborne particles? American Journal of Respiratory and Critical Care Medicine, 164(5), 831–833.

18. Rückerl, R., Schneider, A., Breitner, S., Cyrys, J., & Peters, A. (2011). Health effects of particulate air pollution: A review of epidemiological evidence. Inhalation Toxicology, 23(10), 555–592.

19. Beelen, R., Raaschou-Nielsen, O., Stafoggia, M., Andersen, Z. J., Weinmayr, G., Hoffmann, B., ... Hoek, G. (2014). Effects of long-term exposure to air pollution on natural-cause mortality: An analysis of 22 European cohorts within the multicentre ESCAPE project. The Lancet, 383(9919), 785–795.

20. Hoek, G., Krishnan, R. M., Beelen, R., Peters, A., Ostro, B., Brunekreef, B., & Kaufman, J. D. (2013). Long-term air pollution exposure and cardio-respiratory mortality: A review. Environmental Health, 12(1), 43.