GHEAS · RESEARCH PAPER 01

Ambient PM2.5 Exposure and Life Expectancy Across Countries, 1990-2023

Cross-sectional and two-way fixed-effects analyses of World Development Indicators

Authors	GHEAS Research Team
Affiliation	Global Health Evaluation & Assessment System

Manuscript	Version 1.0 | 20 July 2026

Abstract

Background: Long-term exposure to fine particulate matter (PM2.5) is an established public-health concern, but country-level comparisons of PM2.5 and life expectancy are vulnerable to socioeconomic, demographic, health-system, and temporal confounding. Methods: We harmonized World Development Indicators for PM2.5, life expectancy at birth, gross domestic product per capita, urbanization, population aged 65 years or older, and health expenditure. Descriptive analyses used 6,800 country-year observations from 200 countries or economies during 1990-2023. The 2023 complete-case cross-section included 184 countries. Longitudinal models used an unbalanced panel of 4,406 observations from 186 countries during 2000-2023, with country and year fixed effects and country-clustered standard errors. Results: In 2023, PM2.5 and life expectancy were negatively correlated ($r = -0.279$). The unadjusted cross-sectional estimate was -1.15 years of life expectancy per 10 μg/m³ higher PM2.5 (95% CI -2.25 to -0.05), but the estimate attenuated by 94.7% after adjustment for income, urbanization, age structure, health expenditure, and region (-0.06 years; 95% CI -0.57 to 0.45). In the pre-specified two-way fixed-effects model with time-varying controls, the estimate was 0.047 years per 10 μg/m³ (95% CI -0.217 to 0.312; $p = 0.727$). One-year-lag, three-year-average, balanced-panel, and pandemic-year sensitivity analyses were similarly inconclusive. An independent dummy-variable implementation reproduced the primary coefficient and clustered standard error within 0.0001. Conclusions: The raw country-level association between PM2.5 and life expectancy was largely explained by development context and common time trends. These ecological data do not identify an independent or causal effect on life expectancy; they also do not show that PM2.5 is harmless. More specific health outcomes and stronger exposure and causal designs are needed.

Keywords: air pollution; PM2.5; life expectancy; ecological study; fixed effects; confounding; World Development Indicators

Research highlights

- The raw 2023 country-level correlation was negative ($r = -0.279$).
- Cross-sectional adjustment attenuated the PM2.5 coefficient by 94.7%.
- The primary two-way fixed-effects estimate was +0.047 years per 10 μg/m³ (95% CI -0.217 to +0.312).
- Lag, averaging, balanced-panel, and pandemic-year checks were similarly inconclusive.
- The study does not show that PM2.5 is harmless; it shows that this ecological design cannot isolate its independent effect on life expectancy.

1. Introduction

Fine particulate matter with an aerodynamic diameter of 2.5 μm or less (PM2.5) can penetrate deep into the respiratory tract and is associated with cardiovascular, respiratory, and other adverse outcomes. The World Health Organization's 2021 air-quality guidelines set an annual PM2.5 guideline level of 5 μg/m³ and emphasize that health effects occur at concentrations below earlier guideline levels [1]. Large cohort, causal-modeling, and burden-estimation studies have also linked long-term PM2.5 exposure to mortality and reduced life expectancy [2-6].

A global comparison of national average PM2.5 and life expectancy addresses a different question from an individual-level cohort study. Countries with lower pollution may also have higher incomes, stronger health systems, older populations, different urban structures, and different trajectories of economic and epidemiological change. A simple cross-country correlation can therefore combine pollution effects with development context, while a pooled time series can mix within-country change with global trends. Ecological averages also cannot be used to infer individual risk [7].

The first GHEAS study was designed both as an empirical investigation and as a test of a reproducible evidence workflow. We asked: (1) how strongly were national PM2.5 exposure and life expectancy associated in the latest complete cross-section; (2) how did pre-specified adjustment for economic, demographic, health-system, and regional factors change that association; and (3) did within-country changes remain associated after country and year fixed effects, time-varying controls, alternative exposure timing, and sensitivity analyses? We expected the crude association to be negative, but treated the fully adjusted two-way fixed-effects result as the primary longitudinal test.

2. Methods

2.1 Study design and data sources

We conducted an ecological observational study using a country-year panel. All indicators were retrieved from the World Bank's World Development Indicators API and stored as raw, versioned responses in the project repository [8]. World Bank economy metadata were used to identify countries or economies and to exclude regional, income-group, and other aggregate records. The core PM2.5-life-expectancy dataset covers 1990-2023; the multivariable panel begins in 2000 because health-expenditure coverage is the limiting series.

The analysis unit was a country-year. Records were joined by ISO3 economy code and calendar year. No interpolation or imputation was used. A country-year entered a given model only when every variable required by that model was observed. This produced a complete 200-country core dataset for descriptive analyses and an unbalanced multivariable panel for fixed-effects analyses.

2.2 Variables

The outcome was life expectancy at birth, total, in years (World Bank code SP.DYN.LE00.IN). The main exposure was annual mean PM2.5 concentration in µg/m³ (EN.ATM.PM25.MC.M3), scaled so coefficients represent a 10 µg/m³ difference. Pre-specified controls were real GDP per capita at purchasing-power parity (NY.GDP.PCAP.PP.KD; natural log), urban population share (SP.URB.TOTL.IN.ZS; per 10 percentage points), population aged 65 years or older (SP.POP.65UP.TO.ZS; per 5 percentage points), and current health expenditure per capita at purchasing-power parity (SH.XPD.CHEX.PP.CD; natural log). World Bank region indicators were added to the final cross-sectional model.

The control set was selected before examining multivariable coefficients because these variables capture major differences in national development, settlement, demographic structure, and health-system resources. They should not be interpreted as a complete causal adjustment set: some may be confounders, proxies, or variables on more complex pathways.

2.3 Samples and descriptive analysis

The core sample contained 6,800 matched observations for 200 countries or economies across all 34 years from 1990 through 2023. We summarized annual means, the 2023 cross-sectional correlation, the pooled country-year correlation, income-group correlations, and long changes. The 2023 multivariable sample was chosen by the pre-recorded rule of using the latest year with at least 150 complete observations across all pre-specified variables; 184 countries met that rule.

The panel sample required all six indicators in the same country-year. It contained 4,406 observations for 186 countries over 2000-2023. The median country contributed 24 years (range 11-24); 177 countries had all 24 years. The within-country standard deviation of PM2.5 was 3.406 µg/m³.

2.4 Cross-sectional models

We estimated ordinary least-squares models for 2023 in a pre-specified sequence: M0, PM2.5 only; M1, plus log GDP per capita; M2, plus urbanization and age structure; M3, plus log health expenditure; and M4, plus World Bank region fixed effects. Heteroskedasticity-consistent HC3 standard errors were used. The PM2.5 coefficient is reported in years of life expectancy per 10 µg/m³ higher exposure. The principal cross-sectional diagnostic was attenuation of the PM2.5 coefficient and stability of its direction and interval across the sequence, not a binary significance threshold.

2.5 Panel models

The primary panel specification was: $LE_{it} = \beta PM_{it} + \theta' X_{it} + \alpha_i + \lambda_t + \varepsilon_{it}$, where α_i denotes country fixed effects, λ_t denotes year fixed effects, and X_{it} contains the four time-varying controls. Country fixed effects absorb time-invariant national characteristics; year fixed effects absorb shocks and trends common to all countries. Coefficients were estimated after iterative alternating-projection demeaning. Standard errors were clustered by country with a CR1 finite-sample correction [9]. Two-sided p values and 95% confidence intervals used a large-sample normal approximation.

P1 included country fixed effects only. P2 added year fixed effects. The primary model P3 added all controls. P4 replaced current PM2.5 with a one-year lag, and P5 used the current and previous two years' mean PM2.5. Pre-specified sensitivity models restricted analysis to the 177 countries with complete 24-year panels (S1) and excluded 2020-2021 (S2). These models estimate conditional within-country associations and do not, by themselves, establish causality.

2.6 Reproducibility, verification, and evidence grading

All raw API responses, processed datasets, analysis programs, and machine-readable model outputs are retained in the repository. The primary two-way fixed-effects model was independently re-estimated with a full dummy-variable ordinary least-squares design in NumPy. Agreement was required within an absolute tolerance of 0.0001 for coefficients and clustered standard errors.

GHEAS evidence grades separate description from causal claims. E1 denotes a descriptive pattern or an association that is not stable after reasonable adjustment; E2 denotes a conclusion supported across multiple models or sensitivity checks. No ecological observational result was eligible for a causal E4 grade. AI-assisted computation and drafting followed recorded rules; numerical outputs were independently checked, and publication remains subject to human scientific review.

Table 1. Indicators and analytical roles

Role	Indicator	Code	Scale in models
Outcome	Life expectancy at birth	SP.DYN.LE00.IN	Years
Exposure	Annual mean PM2.5	EN.ATM.PM25.MC.M3	Per 10 μg/m³
Control	GDP per capita, PPP	NY.GDP.PCAP.PP.KD	Natural log
Control	Urban population share	SP.URB.TOTL.IN.ZS	Per 10 percentage points
Control	Population aged 65+	SP.POP.65UP.TO.ZS	Per 5 percentage points
Control	Health expenditure per capita, PPP	SH.XPD.CHEX.PP.CD	Natural log

Source: World Development Indicators API. The 2023 cross-sectional model additionally includes World Bank region fixed effects.

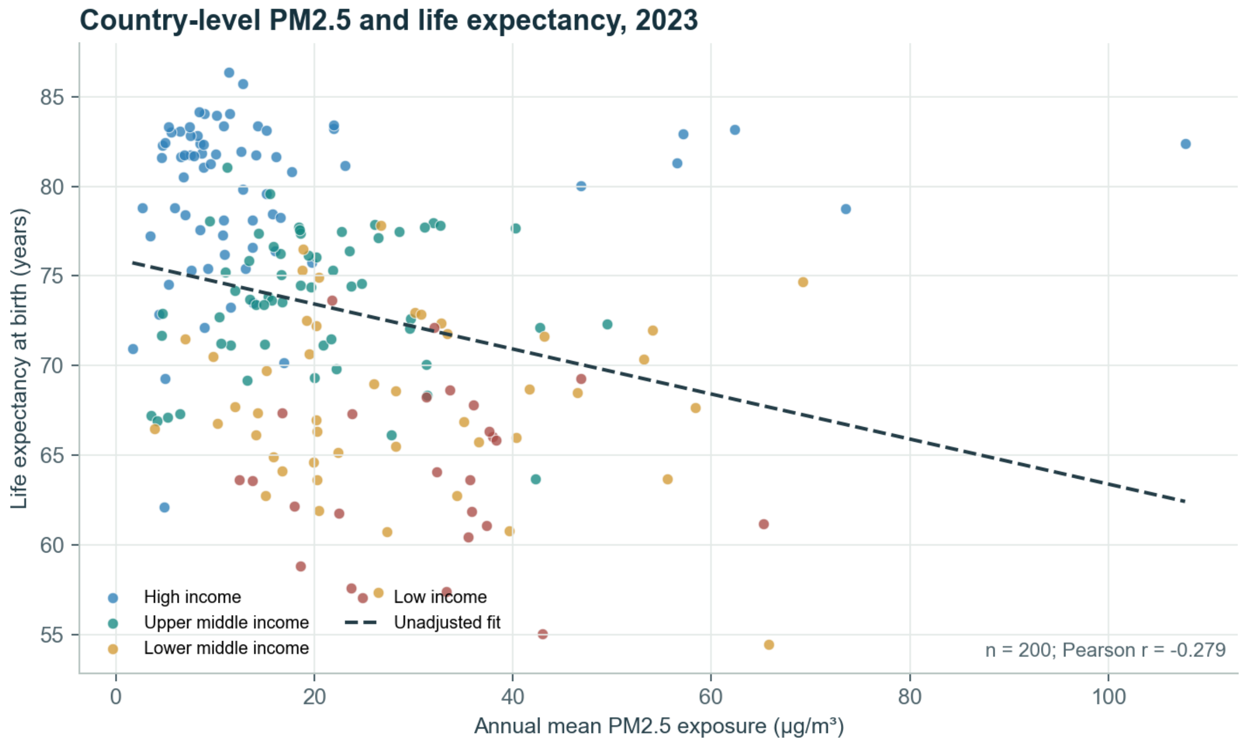
3. Results

3.1 Descriptive patterns

Across the core sample, mean PM2.5 fell by 4.7 μg/m³ between the first and last observed years, while mean life expectancy increased by 8.6 years. The pooled country-year correlation was -0.278, and the 2023 correlation was -0.279. In 2023, the 200-country mean PM2.5 exposure was 21.7 μg/m³ and mean life expectancy was 73.2 years in the core descriptive sample.

The overall 2023 relationship did not persist consistently within World Bank income groups. Correlations were +0.154 in high-income, +0.067 in upper-middle-income, -0.077 in lower-middle-income, and -0.071 in low-income countries. This attenuation and change of direction were early evidence that pooled development differences were an important alternative explanation.

Figure 1. Country-level PM2.5 exposure and life expectancy in 2023



Each point represents one country or economy in the 200-country core sample. The dashed line is the unadjusted ordinary least-squares fit. Colors show World Bank income groups; the pooled ecological relationship should not be interpreted as an individual-level effect.

3.2 Cross-sectional adjustment

The 2023 complete-case sample included 184 countries. Mean PM2.5 was 22.23 µg/m³ (SD 15.87), and mean life expectancy was 73.09 years (SD 7.11). In M0, 10 µg/m³ higher PM2.5 was associated with 1.15 fewer years of life expectancy (95% CI -2.25 to -0.05; $p = 0.041$). Adding log GDP per capita reduced the estimate to -0.20 years (95% CI -0.60 to 0.21), and R^2 increased from 0.066 to 0.746.

The coefficient changed direction after adding demography and urbanization, remained small after adding health expenditure, and returned to -0.06 years in the region-adjusted M4 model (95% CI -0.57 to 0.45; $p = 0.816$). The absolute coefficient attenuated by 94.7% from M0 to M4. Leave-one-region-out estimates from the M3 control set ranged from -0.22 to +0.59 years, further showing that the adjusted country-level estimate was not directionally stable.

Table 2. Descriptive statistics for the 2023 complete-case sample (n = 184)

Variable	N	Mean	SD	Median	Min	Max
PM2.5 (µg/m³)	184	22.23	15.87	18.14	3.52	107.76
Life expectancy (years)	184	73.09	7.11	73.60	54.46	85.71
GDP per capita PPP	184	25451.12	25752.10	16355.01	1033.01	130296.84
Urban population (%)	184	60.64	22.14	62.21	15.06	100.00
Population age 65+ (%)	184	9.60	6.85	6.76	1.57	29.56
Health spending per capita PPP	184	2064.43	2444.33	1110.73	28.26	13473.19

Table 3. Sequential 2023 cross-sectional models

Model	Adjustment	Estimate	HC3 SE	95% CI	p	R^2
M0	Unadjusted	-1.15	0.56	-2.25 to -0.05	0.0411	0.066
M1	+ income	-0.20	0.21	-0.60 to +0.21	0.3434	0.746
M2	+ demography & urbanization	+0.15	0.25	-0.33 to +0.63	0.5432	0.770
M3	+ health resources	+0.24	0.24	-0.23 to +0.71	0.3188	0.772
M4	+ region fixed effects	-0.06	0.26	-0.57 to +0.45	0.8155	0.819

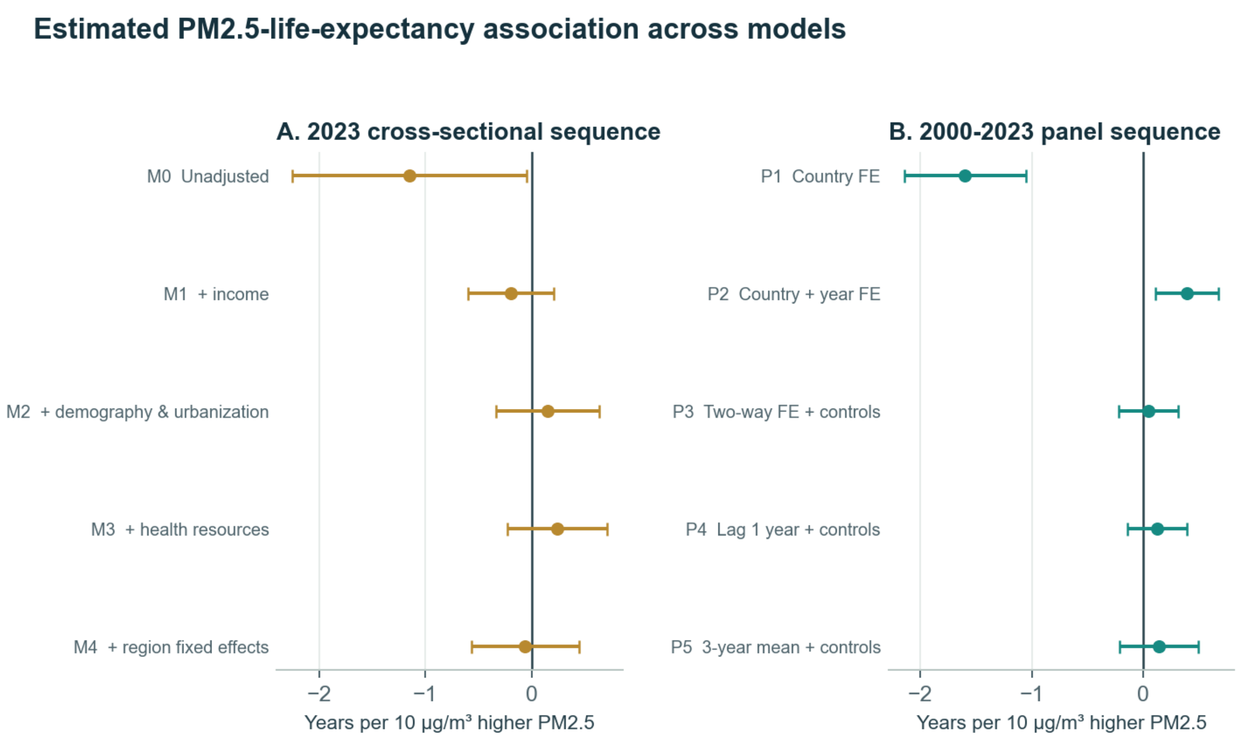
Estimate unit: years of life expectancy per 10 μg/m³ higher PM2.5. M4 includes World Bank region fixed effects. All models use the same 184-country complete-case sample.

3.3 Fixed-effects panel analysis

Country fixed effects alone produced a negative estimate of -1.599 years per 10 μg/m³ (P1; 95% CI -2.146 to -1.052). Adding year fixed effects reversed the estimate to +0.391 years (P2; 95% CI +0.107 to +0.675), demonstrating that global common time trends materially confounded the country-only model.

In the primary P3 model with country and year fixed effects plus time-varying controls, the PM2.5 estimate was +0.047 years per 10 μg/m³ (95% CI -0.217 to +0.312; $p = 0.727$). The one-year-lag estimate was +0.125 years (95% CI -0.143 to +0.393), and the three-year-average estimate was +0.141 years (95% CI -0.212 to +0.493). None of these estimates was distinguishable from zero, and the intervals excluded effects as large as the crude cross-sectional estimate in either direction.

Figure 2. PM2.5 estimates across cross-sectional and panel model sequences



Points are coefficient estimates and bars are 95% confidence intervals. Cross-sectional intervals use HC3 standard errors; panel intervals use country-clustered CR1 standard errors.

Table 4. Fixed-effects panel models, 2000-2023

Model	Specification	Estimate	Cluster SE	95% CI	p	Within R^2
P1	Country FE	-1.599	0.279	-2.146 to -1.052	0.0000	0.048
P2	Country + year FE	+0.391	0.145	+0.107 to +0.674	0.0070	0.005
P3	Two-way FE + controls	+0.047	0.135	-0.217 to +0.312	0.7271	0.134
P4	Lag 1 year + controls	+0.125	0.137	-0.143 to +0.393	0.3615	0.135
P5	3-year mean + controls	+0.141	0.180	-0.212 to +0.493	0.4341	0.135

All panel models include 4,406 observations from 186 countries. P3 is the pre-specified primary model. P4 uses one-year-lag PM2.5; P5 uses the current and previous two years' mean.

Table 5. Complete coefficient table for the primary P3 model

Variable	Estimate	Cluster SE	95% CI	p
PM2.5 current per 10 μg/m³	+0.047	0.135	-0.217 to +0.312	0.7271
Log GDP per capita PPP	+2.079	0.713	+0.682 to +3.476	0.0035
Urban population per 10 pp	+0.647	0.367	-0.073 to +1.368	0.0782
Age 65+ per 5 pp	-2.865	0.506	-3.857 to -1.874	0.0000
Log health spending per capita PPP	+0.074	0.417	-0.743 to +0.891	0.8599

3.4 Sensitivity analysis and numerical verification

Restricting the P3 specification to 177 complete-panel countries produced an estimate of +0.066 years (95% CI -0.202 to +0.333). Excluding 2020-2021 produced +0.034 years (95% CI -0.211 to +0.279). These checks did not change the substantive conclusion.

The independent dummy-variable implementation used 4,406 observations, 186 country clusters, 24 years, and a full-rank 214-column design matrix. Its PM2.5 coefficient differed from the primary implementation by 0.00000567, and its clustered standard error differed by 0.00003353. The largest absolute difference across all P3 coefficients was 0.0000211 and across clustered standard errors was 0.0000356; both were below the 0.0001 tolerance.

Table 6. Pre-specified panel sensitivity analyses

Model	Sample	N	Countries	Estimate	95% CI	p
S1	Balanced 24-year countries	4248	177	+0.066	-0.202 to +0.333	0.6302
S2	Exclude 2020–2021	4034	186	+0.034	-0.211 to +0.279	0.7861

4. Discussion

4.1 Principal findings

A clear negative country-level association between PM2.5 and life expectancy was visible before adjustment. Most of that association disappeared when development context was added to the 2023 cross-section. In the longitudinal analysis, the sign reversal from country-only to country-and-year fixed effects showed that shared time trends were also consequential. The primary controlled two-way fixed-effects estimate was close to zero, imprecise around small positive and negative values, and stable in being inconclusive across exposure-timing and sample sensitivity analyses.

The evidence therefore supports two distinct statements. First, a descriptive global pattern exists: countries with higher PM2.5 tended to have lower life expectancy in the raw 2023 comparison. Second, the present country-level data and models could not isolate an independent PM2.5-life-expectancy association after accounting for measured development factors, stable country characteristics, and common year effects. The first statement is graded E1; the reproducible finding that adequate adjustment removes confirmation of the crude association is graded E2.

4.2 Relation to prior evidence

These results should not be interpreted as contradicting cohort, causal-modeling, or burden-of-disease research. Pope and colleagues reported that reductions in US metropolitan PM2.5 were associated with gains in life expectancy [2]. Apte and colleagues used concentration-response and life-table methods to estimate global life-expectancy loss attributable to ambient PM2.5 [3]. Burnett and colleagues developed a global exposure-mortality function from multiple cohorts [4], while Schwartz and colleagues estimated life-expectancy effects with individual-level Medicare data and causal modeling [5]. Such designs use finer exposure, mortality, or individual covariate information and target different estimands from the national-average conditional association examined here.

Our null or near-null adjusted estimates are compatible with several mechanisms: strong confounding by national development, limited within-country exposure variation, measurement error in modeled national PM2.5 averages, mismatch between annual exposure and the long induction periods of chronic disease, and the insensitivity of all-cause life expectancy to pollution-specific changes. A coefficient near zero in this design is not evidence that the underlying biological effect is zero.

4.3 Strengths

The study used a transparent, fully versioned data pipeline; an explicit sample-formation rule; a pre-specified model sequence; robust or clustered uncertainty estimates; multiple exposure-timing and sample checks; and a separate numerical implementation of the primary estimator. Reporting includes the strong crude association, its attenuation, the sign reversal caused by year effects, and the inconclusive primary estimate, avoiding selective emphasis on a preferred model.

4.4 Limitations

The analysis is ecological. National averages conceal within-country spatial and socioeconomic heterogeneity and cannot support individual-level inference. Exposure may include model-based estimates and measurement error. Life expectancy is a broad summary shaped by many causes and changes slowly; disease-specific mortality may be more responsive to pollution. The control set omits education, smoking, conflict, nutrition, policy, climate, and other time-varying factors. Conversely, some included controls could lie on causal pathways, so the fully adjusted coefficient is not necessarily a total causal effect.

Two-way fixed effects remove time-invariant country differences and common year shocks but do not eliminate unmeasured time-varying confounding, reverse causation, differential trends, or exposure measurement error. Annual current, one-year-lag, and three-year-average exposure definitions may still be too short for chronic outcomes. The panel is unbalanced, although the balanced-panel analysis was similar. Standard errors were clustered by country but not modeled for spatial cross-country dependence. Finally, World Bank source revisions can change historical values, so future updates should be treated as new data versions rather than assumed to reproduce this release exactly.

4.5 Implications and next research step

For GHEAS, the finding is methodologically useful: a compelling unadjusted global pattern did not survive stronger adjustment, and the reporting system preserved that reversal. The next study should prioritize outcomes with closer mechanistic links to PM2.5, including ischemic heart disease, stroke, lung cancer, chronic obstructive pulmonary disease, respiratory mortality, and all-cause mortality. Stronger work should also use subnational or gridded exposure, longer distributed lags, smoking and education covariates, population weighting, spatial dependence checks, and clearly articulated causal identification strategies.

5. Conclusions

Across countries, higher PM2.5 exposure was associated with lower life expectancy in unadjusted data. The estimate attenuated by approximately 95% in the fully adjusted 2023 model, and the primary two-way fixed-effects panel estimate was small and statistically inconclusive. The most defensible conclusion is not that PM2.5 has no health effect, but that national-average World Development Indicators do not isolate an independent effect on life expectancy under the present design. The study demonstrates why cross-country environmental-health correlations require explicit treatment of development context, common time trends, ecological limits, and model uncertainty.

Declarations

Data availability

The raw World Bank API responses, processed country-year datasets, data inventory, and country lists are stored in the GHEAS project repository. Indicator codes and access dates are reported in this manuscript. Public dissemination should preserve the World Bank's source attribution and applicable terms.

Code availability

Reproducible JavaScript programs for data construction, cross-sectional models, and fixed-effects models are located in analysis/. An independent NumPy verification program and machine-readable JSON outputs are also included.

Ethics

The study used publicly available aggregate country-level data and involved no individual participants, identifiable private information, intervention, or specimen collection. Institutional review board review was not required for this analysis.

Funding

No external funding was received for this study.

Competing interests

The authors declare no competing interests.

Author and AI contribution statement

The GHEAS Research Team defined the research logic, interpretation boundaries, and release criteria. AI-assisted workflows performed documented data checks, model execution, comparison, and drafting. The principal fixed-effects calculation was independently reproduced. Human scientific review is required before journal submission or public claims beyond this version.

References

1. World Health Organization. WHO global air quality guidelines: particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. Geneva: WHO; 2021. ISBN 978-92-4-003422-8. https://iris.who.int/handle/10665/345329
2. Pope CA III, Ezzati M, Dockery DW. Fine-particulate air pollution and life expectancy in the United States. N Engl J Med. 2009;360(4):376-386. doi:10.1056/NEJMsa0805646.
3. Apte JS, Brauer M, Cohen AJ, Ezzati M, Pope CA III. Ambient PM2.5 reduces global and regional life expectancy. Environ Sci Technol Lett. 2018;5(9):546-551. doi:10.1021/acs.estlett.8b00360.
4. 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 USA. 2018;115(38):9592-9597. doi:10.1073/pnas.1803222115.
5. Schwartz JD, Wang Y, Kloog I, Yitshak-Sade M, Dominici F, Zanobetti A. Estimating the effects of PM2.5 on life expectancy using causal modeling methods. Environ Health Perspect. 2018;126(12):127002. doi:10.1289/EHP3130.
6. Lelieveld J, Pozzer A, Pöschl U, Fnais M, Haines A, Münzel T. Loss of life expectancy from air pollution compared to other risk factors: a worldwide perspective. Cardiovasc Res. 2020;116(11):1910-1917. doi:10.1093/cvr/cvaa025.
7. Greenland S, Robins J. Invited commentary: ecologic studies - biases, misconceptions, and counterexamples. Am J Epidemiol. 1994;139(8):747-760. doi:10.1093/oxfordjournals.aje.a117069.
8. World Bank. World Development Indicators. Indicator codes EN.ATM.PM25.MC.M3, SP.DYN.LE00.IN, NY.GDP.PCAP.PP.KD, SP.URB.TOTL.IN.ZS, SP.POP.65UP.TO.ZS, and SH.XPD.CHEX.PP.CD. API data retrieved 13 July 2026. https://data.worldbank.org/
9. Cameron AC, Miller DL. A practitioner's guide to cluster-robust inference. J Hum Resour. 2015;50(2):317-372. doi:10.3368/jhr.50.2.317.

Appendix A. Reproducibility record

Artifact	Repository path
Core descriptive data	data/processed/gheas_country_year_pm25_life_expectancy.csv
Multivariable panel data	data/processed/gheas_multivariable_country_year.csv
Cross-sectional analysis	analysis/analyze_multivariable.mjs
Panel analysis	analysis/analyze_panel.mjs
Independent verification	analysis/verify_panel_numpy.py
Machine-readable cross-sectional output	analysis/multivariable_analysis.json
Machine-readable panel output	analysis/panel_analysis.json
Verification output	analysis/panel_verification.json

Raw World Bank responses are stored under data/raw/world_bank/. The data inventory records indicator coverage and source update dates.