Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Potentially Modifiable National Factors Associated With Global Interstitial Lung Disease Burden and Trajectories to 2050

Jefferson A. Buendía

jefferson.buendia@gmail.com

Universidad de Antioquia

Research Article

Keywords: interstitial lung disease, DALYs, tuberculosis, clean cooking, poverty, universal health coverage

Posted Date: August 29th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10617326/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Rationale. Global studies show that interstitial lung disease (ILD) and pulmonary sarcoidosis are increasing and unequally distributed, but potentially modifiable national correlates of this burden remain poorly characterized. Our aim is to identify potentially modifiable national factors associated with ILD burden and estimate burden trajectories to 2050.

Methods. We linked GBD 2021 estimates with World Bank and WHO indicators for 204 countries, 2000–2021, and directly standardized reconstructed both-sex outcomes. We fitted country and year fixed-effects models with country-clustered standard errors and policy-relevant lags, then projected empirically shrunk country trends to 2030, 2040, and 2050.

Results. Four potentially modifiable national factors were associated with ILD burden. Higher tuberculosis incidence (+ 68/100,000) was associated with a 1.00% higher DALY rate at 3 years (95% CI, 0.48–1.53%); greater clean-cooking access (+ 4 percentage points) with a 1.00% lower rate at 10 years (0.32–1.68% lower); higher societal poverty (+ 4.1 percentage points) with a 1.52% higher rate at 5 years (0.04–3.04%); and greater universal health coverage (+ 5.1 points) with a 1.90% lower rate at 1 year (0.32–3.46% lower). Tuberculosis associations reflected partial structural confounding rather than an independent causal effect. Global DALYs rose from 2.09 to 4.04 million (2000–2021), reaching a projected 9.53 million by 2050 under continuation of 2010–2021 trends.

Conclusions. Tuberculosis burden, household energy, poverty, and health-service coverage were the principal modifiable national correlates of ILD burden. Their distinct mechanisms support an integrated agenda of infectious-disease control, clean energy, social protection, and equitable respiratory care, though ecological associations do not establish intervention effects.

Introduction

The global burden of interstitial lung disease (ILD) and pulmonary sarcoidosis has risen markedly over the past three decades. Between 1990 and 2019, incident cases increased by 118.6%, deaths by 166.6%, and DALYs by 122.9% (1). Age-period-cohort analyses indicate that population ageing has amplified this increase while adverse period and birth-cohort patterns have sustained underlying risk, with further growth projected through 2030 (2). Updated GBD 2021 analyses estimated 4.04 million DALYs in 2021—169.3% more than in 1990—and showed that the age-standardized DALY rate increased from 37.1 to 47.6 per 100,000 while the absolute inequality gap widened from 19.6 to 53.4 DALYs per 100,000 (3). Over the same period, incident cases rose from 157,441 to 390,267, and age-standardized incidence, mortality, and DALY rates increased annually by 0.61%, 1.32%, and 0.83%, respectively; projections suggest that incidence and DALY rates may continue rising through 2050 (4). Thus, the available evidence consistently identifies a growing and increasingly unequal burden, but international comparisons remain affected by major differences in registries, diagnostic capacity, and representation of low- and middle-income countries (5).

The remaining knowledge gap concerns specific, potentially modifiable factors that could influence ILD burden through different mechanisms. Pulmonary tuberculosis can produce parenchymal fibrosis and volume loss that resemble or contribute to chronic interstitial and fibrotic lung damage; in a Malawian cohort, 60.7% remained symptomatic after treatment, 34.2% had abnormal spirometry, 44.2% had bronchiectasis, and 9.4% had a destroyed lobe (6–8). Inhaled pollutants provide a more direct link to ILD initiation and progression through repeated alveolar epithelial injury and profibrotic responses: among 433,738 UK Biobank participants, each interquartile-range increase in PM2.5 was associated with 9% higher incident idiopathic pulmonary fibrosis, and combined high PM2.5 exposure and genetic susceptibility conferred a 3.65-fold risk (9). Occupational dusts and fumes represent another preventable source of interstitial injury; 44% of patients with idiopathic pulmonary fibrosis reported relevant exposure, with a pooled association of approximately 1.8 and an estimated population-attributable fraction of 21% (10). Poverty may increase these infectious, household, and occupational exposures, while limited health-service coverage can delay high-resolution CT, multidisciplinary diagnosis, treatment, and rehabilitation; delayed tertiary referral has independently predicted higher mortality, and disparities have been documented across the ILD care pathway (11, 12). These connected but distinct pathways make each domain relevant to ILD burden, yet their relative contribution to national DALY trajectories has not been evaluated jointly.

We therefore conducted a global longitudinal study designed to answer two questions: which potentially modifiable national factors are associated with changes in ILD burden, and what burden could occur by 2030, 2040, and 2050 if recent country trends continue? We linked annual GBD 2021 outcomes with infectious, environmental, social, occupational, and health-system indicators in 204 countries, compared each country with itself over time, and applied policy-relevant lag periods. This distinction matters clinically and for public health: preventing parenchymal injury, recognizing post-tuberculosis lung disease, reducing household and occupational exposures, protecting socially vulnerable populations, and ensuring timely multidisciplinary diagnosis may complement disease-specific pharmacotherapy.

Methods

Study design and data sources

We conducted an ecological longitudinal analysis in which the unit of observation was the country-year. The outcome panel included 204 countries and territories observed annually from 2000 through 2021 (4,488 country-years). GBD 2021 supplied age-, sex-, location-, year-, measure-, and metric-specific estimates for the composite cause “interstitial lung disease and pulmonary sarcoidosis” (13). World Development Indicators supplied tuberculosis incidence, clean-cooking access, societal poverty, vulnerable employment, PM2.5, industrial employment, coal-generated electricity, smoking, UHC service coverage, health expenditure, out-of-pocket spending, and demographic indicators (14). WHO sources supplied additional mortality, air-quality, health-system, and diagnostic-device indicators (15). We did not interpolate, simulate, or extrapolate missing exposure values; each model used complete country-years. Reporting followed GATHER (16).

Definitions and units of the principal policy indicators

Tuberculosis incidence was the estimated number of new and relapsed tuberculosis episodes per 100,000 population per year. Clean-cooking access was the percentage of the total population primarily using clean fuels or technologies for cooking; electricity, liquefied petroleum gas, natural gas, biogas, solar, and alcohol are classified as clean, whereas kerosene and solid polluting fuels are excluded. Societal poverty was the percentage of the population living below the World Bank societal poverty line, a country-year threshold defined as the greater of the international minimum and a relative line linked to median national income or consumption; it therefore captures deprivation relative to the resources available in each society rather than only extreme poverty. UHC service coverage was a unitless 0–100 index, with higher values indicating better coverage, calculated as the geometric mean of 14 tracer indicators across reproductive, maternal, newborn and child health, infectious diseases, noncommunicable diseases, and service capacity and access (14, 15). For indicators expressed as percentages, a percentage-point change is an absolute difference between two proportions: an increase from 40% to 44% is 4 percentage points, equivalent to a 10% relative increase. A one-point increase in the UHC index means movement, for example, from 60 to 61 on its 0–100 scale and is not a percentage-point change. The reported exposure increments—67.7 tuberculosis cases per 100,000, 4.0 percentage points of clean-cooking access, 4.1 percentage points of societal poverty, and 5.1 UHC-index points—were the typical within-country changes used to translate model coefficients into interpretable original units.

Both-sex outcomes and direct standardization

For country i, year t, age group a, and sex s, we summed male and female numbers. When both numbers and rates were available, we reconstructed the sex-specific population denominator and then calculated the both-sex age-specific rate:

$$N_{ita} = \sum_s N_{its} ; \quad P_{its} = N_{its} \div (r_{its} / 100{,}000) ; \quad r_{ita} = 100{,}000\, N_{ita} \div P_{ita} \quad (1)$$

$$R_{it} = \left(\sum_a w_a r_{ita}\right) \div \left(\sum_a w_a\right) \quad (2)$$

Direct standardization was chosen to remove changes caused only by differing age structures. The weights w_a were the combined-sex 2021 population in age group a across included locations, held fixed for all countries and years. This study-standardized rate is internally consistent but is not the official IHME age-standardized rate. Secondary outcomes were DALY numbers, crude rates, incidence, prevalence, and mortality at ages ≥ 2 years because younger ages were absent from the supplied death export.

Observed country trends

We fitted a log-linear model separately in each country because a constant proportional change is more interpretable than an absolute change when baseline burden differs substantially between countries:

$$\ln\left(R_{it}\right) = a_i + b_i t + e_{it}; \quad EAPC_i = 100\left[\exp(b_i) - 1\right] \quad (3)$$

The log transformation stabilizes proportional effects and prevents fitted rates below zero. A country was classified as increasing when the entire 95% confidence interval for EAPC was above zero, decreasing when it was below zero, and stable or uncertain otherwise.

Longitudinal models of modifiable factors

The primary model related the natural logarithm of the study-standardized DALY rate to lagged exposures, country fixed effects, and year fixed effects:

$$\ln\left(R_{it}\right) = \alpha_i + \lambda_t + \Sigma_k \beta_k Z_{k,it-L_k} + \epsilon_{it} \quad (4)$$

Here, α_i controls every measured or unmeasured country characteristic that does not change over time, including geography and stable health-system features; λ_t controls shocks shared by all countries in a given year. This fixed-effects design was chosen because the scientific question concerns whether burden changes when a country's exposure changes, rather than why countries differ cross-sectionally (17, 18). The transformed coefficient $100[\exp(\beta_k) - 1]$ is the percentage difference in the DALY rate associated with the stated change in the original exposure.

Exposure scaling, lag choice, and statistical inference

To express coefficients in clinically interpretable units, the model was fitted using each exposure's typical within-country change, but results are reported in the original scale: 67.7 tuberculosis cases per 100,000 population, 4.0 percentage points of clean-cooking access, 4.1 percentage points of poverty, and 5.1 UHC-index points. Lags were specified before outcome modelling: 1, 3, and 5 years for infectious sequelae; 5 and 10 years for chronic environmental exposure; and 1 and 3 years for health-system response. These windows acknowledge latency while avoiding a biologically implausible assumption that all exposures act immediately. Exposures were grouped into etiologic/environmental, infectious, social-vulnerability, health-system/ascertainment, and contextual domains. Domain models reduced overadjustment for variables that may lie on the same causal pathway. Standard errors were clustered by country because repeated observations within a country are correlated. The cluster-robust covariance estimator was:

$$\hat{V}(\hat{\beta}) = (X'X)^{-1}\left[\Sigma_i X_i' \hat{u}_i \hat{u}_i' X_i\right](X'X)^{-1} \quad (5)$$

Tuberculosis causal-structure analyses

We did not assume that tuberculosis incidence represented an independent biological exposure. We therefore refitted the tuberculosis model on identical complete-case samples after sequential adjustment

for poverty, vulnerable employment, income inequality, secondary education, population density, urbanization, income, clean cooking, UHC, financial access, physicians, and hospital capacity. We calculated coefficient attenuation as 100×(1 − βadjusted/βunadjusted). Population density was treated only as a national settlement proxy and not as a direct measure of household crowding. No direct country-year measures of household crowding, HIV, immunosuppressive treatment, autoimmune disease, or malnutrition were available. We then used future tuberculosis incidence at 1, 3, and 5 years as negative-control exposures: future exposure cannot cause a previous outcome, so a positive association indicates residual common trends, reverse direction, or ascertainment bias rather than a specific lagged effect (19). Finally, we added country-specific linear time trends to the country and year fixed-effects model to test whether the association depended on gradual unmeasured national trajectories.

Multiplicity and interpretation

Prespecified screening estimates were controlled within their analytic family using the Benjamini-Hochberg false-discovery-rate procedure (20). This method was chosen because the study considered several correlated, policy-relevant hypotheses and control of the expected proportion of false discoveries is less conservative and more informative than controlling the probability of any false positive. The manuscript reports clinically coherent adjusted domain models and their effect sizes and confidence intervals; statistical significance alone was not used to infer causality.

Projection model

Country-specific log-linear slopes were estimated for 2000–2021 and 2010–2021. Because short or noisy country series can produce implausibly extreme extrapolations, each slope was shrunk toward the precision-weighted global mean using an empirical-Bayes estimator:

$$\tilde{b}_i = [\tau^2/(\tau^2 + v_i)]\, b_i + [v_i/(\tau^2 + v_i)]\, \mu \;; \quad \tilde{v}_i = (v_i^{-1} + \tau^{-2})^{-1} \qquad (6)$$

where b_i and v_i are the country slope and its variance, μ is the precision-weighted mean slope, and τ^2 is between-country variance. Shrinkage preserves well-estimated country signals while reducing the influence of unstable estimates.

Projection scenarios and uncertainty

For horizon h years after 2021, projected DALYs were:

$$\hat{Y}_{i,2021+h} = Y_{i,2021} \exp(\tilde{b}_i h) \qquad (7)$$

$$95\%\ \text{interval} = \exp\{\ln(Y_{i,2021}) + \tilde{b}_i h \pm 1.96 h \sqrt{\tilde{v}_i}\} \qquad (8)$$

We evaluated no change after 2021, 50% attenuation of the recent slope, continuation of 2010–2021 trends, and continuation of 2000–2021 trends. Scenario analysis was chosen instead of a single deterministic forecast to make the dependence on trend persistence explicit. Intervals quantify slope

uncertainty only; they are not official IHME forecasts and do not include future changes in population, diagnosis, treatment, climate, conflict, or policy (21).

Sensitivity and reproducibility

We repeated the integrated model using crude DALY rates, mortality, incidence, and prevalence and restricted the observation period to 2000–2019 to assess sensitivity to the COVID-19 period. Analyses used complete cases without imputation. The self-contained Python supplement embeds the analytical data, model specifications, expected outputs, and cryptographic checksums; it regenerates the complete results and figures without external files.

Post-hoc analyses of confounding, heterogeneity, and policy translation

Following the primary analyses, we added four post-hoc analyses. First, we decomposed the attenuation of the tuberculosis coefficient after adjustment for the available structural block using the method of Gelbach (22), which partitions the change in a coefficient between a short and a long regression into the exact contribution of each added covariate, computed as the product of that covariate's coefficient in the long regression and its own association with tuberculosis incidence in an auxiliary regression. Second, we tested whether the four modifiable associations differed by country income by classifying countries into two groups using their 2000–2021 average GDP per capita (PPP) and re-estimating each two-way fixed-effects model within each half. Third, we translated the clean-cooking and UHC coefficients into illustrative averted burden by applying each country's 2021 coverage gap to a universal clean-cooking access target and to a UHC service-coverage index of 80, multiplying the implied proportional reduction by each country's 2021 DALY number; these calculations are exploratory policy simulations under the estimated ecological association and do not assume a causal guarantee. Fourth, we tested spatial dependence in country-level DALY-rate trends (EAPC) using Moran's I (23) with inverse-distance weights between country centroids, with significance assessed by 9,999 permutations.

Results

The associated factors spanned infectious, household-energy, social, and health-system domains. When tuberculosis incidence increased by approximately 68 cases per 100,000 above a country's usual level, the ILD DALY rate was 1.00% higher three years later (95% CI, 0.48%–1.53%; $p < 0.001$; 2,936 country-years). Expanding clean-cooking access by 4 percentage points was associated with a 1.00% lower rate ten years later (95% CI, 0.32%–1.68% lower; $p = 0.004$; 2,081 country-years). A 4.1-percentage-point increase in societal poverty was followed by a 1.52% higher rate five years later (95% CI, 0.04%–3.04%; $p = 0.045$; 1,285 country-years), whereas a 5.1-point improvement in the UHC service-coverage index was followed by a 1.90% lower rate one year later (95% CI, 0.32%–3.46% lower; $p = 0.019$; 3,955 country-years) (Table 1; Fig. 1). The comparable effect magnitudes across different policy domains indicate that ILD burden tracked not one isolated exposure but a broader configuration of preventable pulmonary injury, social vulnerability, and access to care.

Table 1
Modifiable factors associated with national DALY rates

Factor	Policy-relevant change	Lag	Difference, %	95% CI	p
Tuberculosis incidence	≈ 68 cases/100,000	3 years	+ 1.00	+ 0.48 to + 1.53	< 0.001
Clean-cooking access	+ 4 percentage points	10 years	−1.00	−1.68 to − 0.32	0.004
Societal poverty	+ 4.1 percentage points	5 years	+ 1.52	+ 0.04 to + 3.04	0.045
UHC service coverage	+ 5.1 index points	1 year	−1.90	−3.46 to − 0.32	0.019

Differences compare each country with itself and are adjusted for country and year fixed effects and domain covariates; standard errors are clustered by country. Changes correspond to a typical observed within-country variation and are reported in original units. The four estimates identify complementary domains: higher tuberculosis and poverty accompanied higher burden, whereas expansion of clean household energy and health coverage accompanied lower burden. These are longitudinal ecological associations, not causal intervention effects.

Sensitivity analyses showed that the four indicators captured both disease-producing conditions and the way disease is detected and survives. Tuberculosis remained associated after adjustment for clean cooking, income, UHC, and out-of-pocket spending (+ 0.69%; 95% CI, 0.19%–1.19%), but poverty and the broader structural block reduced its coefficient by 23% and 34%; future-exposure and country-trend tests therefore argued against interpreting tuberculosis as an independent cause. Clean cooking retained an inverse association in the development-and-access model (− 1.12%; 95% CI, − 1.94% to − 0.29%), whereas the UHC estimate weakened after adjustment for correlated development indicators. Across alternative outcomes, tuberculosis accompanied higher DALYs and mortality, consistent with post-infectious damage and broader respiratory vulnerability. Clean cooking accompanied lower mortality but higher recorded incidence and prevalence, a pattern compatible with countries simultaneously reducing harmful household exposure and improving survival and case recognition. Higher UHC accompanied lower crude DALY and incidence rates, consistent with prevention and earlier care, although greater diagnostic capacity can also increase the number of cases initially recorded (Supplementary Figure S1; Supplementary Table S3). Taken together, the indicators may affect lung injury, diagnosis, and survival at the same time; therefore, opposite directions across recorded cases and fatal burden do not imply that clean energy or health coverage causes ILD.

The increase in ILD burden was not explained solely by population growth or population ageing. DALYs increased from 2,093,588 in 2000 to 4,039,790 in 2021 (+ 93.0%), while the crude rate increased from 35.81 to 52.94 per 100,000 (+ 47.8%) and the study-standardized rate—calculated using the same age weights in every country and year—increased from 46.24 to 51.73 (+ 11.9%) (Fig. 2). Because an age-standardized rate is not increased merely by having a larger or older population, its rise indicates that the underlying recorded burden also worsened. Consistently, standardized rates increased in 107 countries

(52.5%), decreased in 63 (30.9%), and were stable or uncertain in 34 (16.7%) (Fig. 3). Demographic change therefore amplified the absolute number of DALYs, but did not fully account for the global deterioration.

Alternative outcomes separated a probable burden signal from changes in diagnosis and survival. In the integrated sensitivity models, higher tuberculosis incidence was associated with standardized DALY rates (+ 0.85%), crude DALY rates (+ 0.98%), and mortality (+ 1.21%). Greater clean-cooking access was associated with lower mortality (− 1.21%) but with higher recorded incidence and prevalence, while greater UHC coverage was associated with lower crude burden and incidence (Supplementary Figure S1). The divergent incidence and prevalence estimates should not be interpreted as energy access causing disease; improved survival, ascertainment, and the composition of the broad GBD cause can move recorded cases differently from deaths and DALYs.

Continuation of recent trends would produce a very high burden by mid-century. Under the prespecified continuation of 2010–2021 country trends, global DALYs reached 5.25 million in 2030, 7.05 million in 2040, and 9.53 million in 2050; the slope-based 2050 interval was 8.77–10.39 million. Scenario choice mattered substantially: 2050 burden ranged from 4.04 million with no post-2021 change to 6.64 million with 50% trend attenuation and 11.14 million under continuation of the full 2000–2021 trend (Fig. 4). These estimates quantify the cost of persistence, not an unavoidable future.

Post-hoc analyses clarified confounding, heterogeneity, and policy relevance

A formal decomposition showed that poverty and clean-cooking access, not an independent effect of tuberculosis, accounted for most of the coefficient's attenuation. Of the 34% attenuation observed after adjustment for the available structural block, societal poverty explained 49.3% and clean-cooking access explained 45.5%—together accounting for essentially all of the change—while vulnerable employment explained a further 16.4% and education and population density acted as small suppressors (Supplementary Figure S2; Supplementary Table S9). This decomposition reinforces the interpretation that tuberculosis incidence is chiefly a marker of shared social and household-energy vulnerability rather than an independent modifiable target.

The four associations were more often significant in the higher-income country half. Stratifying countries by their 2000–2021 average income, tuberculosis (+ 2.06%; 95% CI, 0.05%–4.11%), clean-cooking access (− 2.22%; −3.20% to − 1.24%), and UHC coverage (− 3.93%; −6.33% to \u22121.48%) were significantly associated with burden only in the above-median-income half; none reached significance in the below-median half, including poverty (+ 1.79%, not significant) (Supplementary Figure S3; Supplementary Table S10). This pattern is more consistent with greater within-country variation and administrative data quality in transitioning middle- and higher-income economies than with a larger biological effect in wealthier settings, and should not be read as evidence that these factors matter less for low-income countries.

Applying the clean-cooking and UHC coefficients to each country's 2021 coverage gap suggested a modest but tangible illustrative burden reduction. Closing the clean-cooking gap to universal access in all countries corresponded to approximately 162,600 fewer DALYs (4.0% of the 2021 global burden), and raising all countries to a UHC service-coverage index of 80 corresponded to approximately 103,400 fewer DALYs (2.6%); India, Bangladesh, and Pakistan accounted for the largest single-country contributions under both scenarios (Supplementary Figure S4; Supplementary Table S11). These figures are illustrative extrapolations of the estimated ecological associations to current coverage gaps, not causal guarantees of achievable prevention.

Country-level trends were not spatially independent. Moran's I on the fitted EAPC of the DALY rate, using inverse-distance weights between country centroids, was 0.125 (permutation $p < 0.0001$; Supplementary Figure S5, Supplementary Table S12), indicating that neighboring countries had more similar trends than expected by chance. This finding does not overturn the country-and-year fixed-effects estimates, which absorb common annual shocks and stable country characteristics, but it indicates that regional processes—shared epidemiological, environmental, or health-policy diffusion across neighboring countries—are not fully captured by the primary model and warrants explicit spatial modelling in future work.

Discussion

The principal finding is that national ILD burden was associated with four distinct and potentially modifiable domains: tuberculosis burden, household energy, poverty, and health-service coverage. Their similar effect magnitudes but different temporal windows argue against a single explanatory pathway. Instead, the results identify complementary opportunities to prevent infectious and combustion-related lung injury, reduce exposure-producing deprivation, and improve timely respiratory diagnosis and care. This multidomain interpretation is the study's principal contribution to global respiratory epidemiology.

Our work addresses a different question from the closest global comparators. Ma et al. documented that ILD incidence, deaths, and DALYs increased by 118.6%, 166.6%, and 122.9%, respectively, from 1990 to 2019 (1). Zeng and Jiang attributed changing burden to age, period, and cohort patterns and projected continued growth to 2030 (2). Chen and Li showed both a rising standardized DALY rate and widening absolute inequality by SDI through 2021 (3). Zhao et al. combined joinpoint trends, demographic decomposition, SDI correlation, and Bayesian age-period-cohort projections and estimated annual increases of 0.61% in incidence, 1.32% in mortality, and 0.83% in DALY rates (4). We similarly found increasing underlying rates and a large future burden, but extended this literature by replacing a composite development gradient with specific, potentially modifiable domains and by estimating longitudinal within-country associations. We found no previous global study that jointly tested these four domains with country and year fixed effects; the novelty is therefore multidomain and longitudinal, not a claim that tuberculosis, pollution, or disadvantage have never been linked to lung disease.

The infectious-domain association is clinically plausible but causally complex. Post-tuberculosis lung disease includes fibrosis, volume loss, bronchiectasis, obstruction, vascular injury, and pleural disease

(6–8), supporting systematic respiratory assessment after treatment. However, poverty adjustment reduced the national tuberculosis coefficient by 23%, the broader structural block by 34%, future tuberculosis tracked current ILD, and country-specific trends removed the lagged association. Tuberculosis incidence therefore appears to combine possible post-infectious injury with unmeasured structural vulnerability, rather than identifying an isolated causal effect. This nuance does not diminish tuberculosis control; it specifies that prevention, prompt treatment, and post-TB respiratory care should operate alongside social and environmental interventions.

The inverse association with clean-cooking access connects household energy policy with fibrotic respiratory health. Biomass and coal combustion generate fine and ultrafine particles, nitrogen oxides, carbon monoxide, and organic compounds capable of epithelial injury, oxidative stress, macrophage activation, and profibrotic signalling. In ILD cohorts, indoor and ambient pollution have been associated with lower lung function and adverse outcomes (24–27). Johannson et al. reported lower mean lung function with greater particulate and nitrogen-dioxide exposure in idiopathic pulmonary fibrosis (25), and Goobie et al. linked particulate exposure with impaired lung function and mortality in fibrotic ILD (26). Sack et al. found worse symptoms, quality of life, FVC, and diffusing capacity with long-term PM2.5 exposure, although one-year progression and mortality were not significant (27). That heterogeneity is compatible with our findings: national clean-cooking access is a proxy for chronic household combustion, not a personal exposure measurement. The ten-year lag and lower DALY and mortality estimates support prevention, while higher recorded incidence and prevalence in sensitivity analyses may reflect better diagnosis, survival, urbanization, or residual confounding rather than a harmful effect of clean energy.

Poverty was positively associated with burden, and UHC coverage was inversely associated. These are complementary rather than competing findings. Poverty increases the probability of hazardous work, biomass use, crowded housing, untreated infection, poor nutrition, delayed presentation, and inability to sustain specialist care. Reviews of ILD disparities identify differences by income, race and ethnicity, geography, insurance, referral, transplantation, and access to antifibrotic treatment (12). Lamas et al. found that delayed access to a tertiary ILD centre was independently associated with poorer survival in idiopathic pulmonary fibrosis (11). A five-point improvement in the UHC index was associated with approximately 1.9% lower burden in our domain model, consistent with earlier recognition, longitudinal follow-up, oxygen and rehabilitation, and access to multidisciplinary diagnosis. Yet greater diagnostic capacity can initially increase recorded incidence and prevalence. The positive association observed for health expenditure in secondary models likely reflects this ascertainment and survival pathway, together with reverse causation, and should not be interpreted as expenditure causing ILD.

Occupational and ambient-air indicators require careful interpretation. A systematic review and meta-analysis supports associations of metal, wood, silica, agricultural dusts, fumes, and vapours with idiopathic pulmonary fibrosis (10), but national employment shares cannot recover cumulative task-specific dose, informal work, personal protective equipment, or the healthy-worker effect. Similarly, absence of an independent national PM2.5 association does not contradict cohort evidence (25–27). Country means obscure within-country hotspots, indoor exposure, ultrafine particles, wildfire smoke,

occupational mixtures, and residential mobility. Nondifferential exposure error would generally attenuate associations toward no difference. The clean-cooking indicator may have performed better because it captures a concrete household technology transition with large exposure contrast.

Our burden estimates reinforce the urgency of prevention. Global DALYs increased by 93% between 2000 and 2021, while the study-standardized rate increased by 11.9% and rose in 107 countries. The difference between numbers and standardized rates is expected: population growth and ageing amplify total need, but the rate increase shows that demography is not the complete explanation. The direction agrees with prior GBD analyses (1–4). Under recent trend continuation, DALYs reached 9.53 million in 2050. Zhao et al. likewise projected rising incidence and DALY rates through 2050 (4), and broader GBD scenario work shows how demographic and risk trajectories can generate widely separated futures (21). Our range of 4.04–11.14 million is intentionally scenario-dependent; its policy value lies in showing how much burden is contingent on persistence rather than claiming forecasting certainty.

The findings suggest an integrated clinical and public-health response. Tuberculosis programmes should incorporate chronic respiratory assessment and referral after treatment; energy and climate policy should accelerate affordable clean cooking without shifting costs to poor households; social protection should reduce exposure-producing deprivation and financial barriers; and UHC packages should include spirometry, high-resolution computed tomography when indicated, multidisciplinary review, pulmonary rehabilitation, oxygen, vaccination, palliative care, and evidence-based pharmacotherapy. Occupational surveillance and exposure control remain essential even when national proxy variables are imprecise. These actions could improve outcomes beyond the GBD ILD category, including chronic obstructive, post-infectious, and cardiovascular disease.

Four post-hoc analyses refined these interpretations. A formal decomposition confirmed that poverty and clean-cooking access, rather than an independent effect of tuberculosis, accounted for nearly all of the coefficient's attenuation, reinforcing tuberculosis incidence as a marker of shared vulnerability. Country-income stratification showed that the four associations were detectable mainly among above-median-income countries, a pattern we interpret as reflecting greater within-country variation and administrative data quality during economic and health-system transitions rather than a larger true effect in wealthier settings. Translating the clean-cooking and UHC coefficients into each country's 2021 coverage gap suggested a modest illustrative reduction—about 4.0% and 2.6% of the 2021 global burden, respectively—concentrated in a small number of high-burden countries led by India, Bangladesh, and Pakistan; these are extrapolations of the estimated associations, not causal guarantees. Finally, positive spatial autocorrelation in country trends (Moran's $I = 0.125$, $p < 0.0001$) indicates that neighboring countries share more similar trajectories than expected by chance, a regional dependency not fully captured by the country-and-year fixed-effects design and worth addressing with explicit spatial models in future work.

This study has important strengths: global annual coverage; transparent reconstruction of both-sex estimates; a fixed age standard; country and year fixed effects; biologically motivated lags; country-clustered inference; domain-specific and integrated models; alternative outcomes; a pre-COVID sensitivity analysis; empirical-Bayes stabilization; and a fully embedded reproducibility file. Limitations are equally

important. Ecological associations cannot establish individual causation or intervention effects. GBD outcomes and several covariates are modelled estimates whose uncertainty draws were unavailable, and the composite outcome combines heterogeneous ILDs with pulmonary sarcoidosis. The study-standardized rate is not the official IHME rate. Complete-case samples differed across exposures, measurement error was likely substantial, and unmeasured time-varying confounding, reverse causation, changing ascertainment, spatial spillovers, and policy co-adoption remain possible. Mortality analyses excluded ages younger than two years because of source-file coverage. Fixed effects remove stable differences but do not correct time-varying bias. Positive spatial autocorrelation in country trends indicates that regional diffusion processes are not fully captured by this design. Finally, projections assume continuation of estimated slopes and omit future population changes, migration, conflict, climate hazards, diagnostic innovation, and treatment uptake; their intervals include slope uncertainty but not the full uncertainty of GBD estimates or future shocks.

The expanded evidence base strengthens the environmental and occupational interpretation. A 2025 systematic review of 96 studies (40.8 million participants) found higher ILD risk with smoking (OR 1.69), organic exposures (OR 1.56), metals (OR 1.52), dust (OR 1.45), and asbestos (OR 1.53) (28). Patient cohorts add outcome-specific evidence: ozone and nitrogen dioxide preceded acute IPF exacerbations (29); PM2.5 and PM10 were associated with mortality in the French COFI cohort (30); PM10 accompanied an additional 46 mL/year decline in FVC (31); and ultrafine particles were associated with transplantation-free survival and functional decline when residential exposure was measured precisely (32). These quantitative findings support prevention while explaining why a coarse national annual PM2.5 mean may not reproduce individual-level associations.

Exposure assessment itself is a clinically actionable component of ILD care. In a tertiary cohort, 65% of patients reported a potentially relevant domestic or occupational exposure (33), while population-based MESA analyses linked occupational vapours, gases, dusts, and fumes to 2.64% greater high-attenuation area and higher odds of interstitial lung abnormalities in relevant subgroups (34). Reviews and meta-analyses consistently implicate organic dust, metal, wood, silica, fumes, and vapours (35, 36), and low income has been associated with worse IPF prognosis and greater occupational exposure (37). These data reinforce interpreting clean cooking and poverty as markers of preventable exposure environments, not merely national development scores.

The literature also clarifies why burden metrics diverge. In fibrotic ILD, PM2.5 exposure has been associated with baseline severity, progression, and mortality (38), but exposure histories, latency, diagnostic access, and phenotype classification vary across settings (39). Post-tuberculosis reviews document persistent structural disease, disability, recurrent respiratory events, and impaired quality of life (40, 41). Environmental remediation reviews emphasize structured exposure histories and avoidance when a plausible antigen or toxicant is found (42). Together, this evidence supports our integrated interpretation: national indicators may influence true disease, recognition, survival, and coding simultaneously, so concordance across DALYs and mortality is more persuasive than any single incidence estimate.

In conclusion, tuberculosis burden, household energy, poverty, and UHC were the principal potentially modifiable national factors associated with ILD DALY rates, while global burden trajectories indicated substantial growth through 2050 if recent trends persist. The findings support an integrated strategy combining infectious-disease control and post-TB care, clean household energy, social and occupational protection, and equitable multidisciplinary respiratory services. Because associations are ecological and mechanisms differ, intervention effects require confirmation in individual-level and quasi-experimental studies.

Conclusions

In conclusion, tuberculosis burden, household energy, poverty, and UHC were the principal potentially modifiable national factors associated with ILD DALY rates, while global burden trajectories indicated substantial growth through 2050 if recent trends persist. The findings support an integrated strategy combining infectious-disease control and post-TB care, clean household energy, social and occupational protection, and equitable multidisciplinary respiratory services. Because associations are ecological and mechanisms differ, intervention effects require confirmation in individual-level and quasi-experimental studies.

Declarations

Competing interests:

The authors declare no competing interests.

Consent for publication:

Not applicable.

Ethics approval and consent to participate:

Not applicable. This study was only based on published data. Thus, an ethical approval or consent to participate was not necessary.

Funding:

This study was supported by the funding of authors.

Author Contribution

JAB. wrote the main manuscript text

Data availability:

No datasets were generated or analysed during the current study.

References

1. Ma X, Zhu L, Kurche JS, Xiao H, Dai H, Wang C. Global and regional burden of interstitial lung disease and pulmonary sarcoidosis from 1990 to 2019. Thorax. 2022;77:596–605. 10.1136/thoraxjnl-2020-216732.
2. Zeng Q, Jiang D. Global trends of interstitial lung diseases from 1990 to 2019: an age-period-cohort study and projections until 2030. Front Med (Lausanne). 2023;10:1141372. 10.3389/fmed.2023.1141372.
3. Chen S, Li Y. Global health inequalities in the burden of interstitial lung disease and pulmonary sarcoidosis from 1990 to 2021. BMC Public Health. 2024;24:2892. 10.1186/s12889-024-20430-y.
4. Zhao L, Zhou Y, Jia Y, et al. Assessing the global burden of interstitial lung disease and pulmonary sarcoidosis using multiple statistical models. BMC Pulm Med. 2025;25:352. 10.1186/s12890-025-03813-8.
5. Spagnolo P, Ryerson CJ, Putman RK, et al. Global epidemiology and burden of interstitial lung disease. Lancet Respir Med. 2025. 10.1016/S2213-2600(25)00129-8.
6. Allwood BW, Byrne A, Meghji J, et al. Post-tuberculosis lung disease: clinical review of an under-recognised global challenge. Respiration. 2021;100:751–63. 10.1159/000512531.
7. Meghji J, Simpson H, Squire SB, Mortimer K. A systematic review of imaging-defined post-tuberculosis lung disease. PLoS ONE. 2016;11:e0161176. 10.1371/journal.pone.0161176.
8. Meghji J, Lesosky M, Joekes E, et al. Patient outcomes associated with post-tuberculosis lung damage in Malawi. Thorax. 2020;75:269–78. 10.1136/thoraxjnl-2019-213808.
9. Cui F, Sun Y, Xie J, et al. Air pollutants, genetic susceptibility and risk of incident idiopathic pulmonary fibrosis. Eur Respir J. 2023;61:2200777. 10.1183/13993003.00777-2022.
10. Gandhi SA, Li N, Zhu Z, et al. Occupational exposures and idiopathic pulmonary fibrosis: a systematic review and meta-analysis. Ann Am Thorac Soc. 2024;21:486–98. 10.1513/AnnalsATS.202305-402OC.
11. Lamas DJ, Kawut SM, Bagiella E, et al. Delayed access and survival in idiopathic pulmonary fibrosis. Am J Respir Crit Care Med. 2011;184:842–7. 10.1164/rccm.201104-0668OC.
12. Tikellis G, Hoffman TW, Oldham JM, et al. Health disparities and associated social determinants of health in interstitial lung disease. Eur Respir Rev. 2025;34:240176. 10.1183/16000617.0176-2024.
13. GBD 2021 Diseases and Injuries Collaborators. Global incidence, prevalence, years lived with disability, disability-adjusted life-years, and healthy life expectancy for 371 diseases and injuries in 204 countries and territories, 1990–2021. Lancet. 2024;403:2133–61. 10.1016/S0140-6736(24)00757-8.

14. World Development Indicators. World Bank, Washington DC. World Bank; accessed 2026. https://databank.worldbank.org/source/world-development-indicators
15. World Health Organization. Global Health Observatory. Geneva: WHO; accessed 2026. https://www.who.int/data/gho
16. Stevens GA, Alkema L, Black RE, et al. Guidelines for Accurate and Transparent Health Estimates Reporting: the GATHER statement. Lancet. 2016;388:e19–23. 10.1016/S0140-6736(16)30388-9.
17. Mundlak Y. On the pooling of time series and cross section data. Econometrica. 1978;46:69–85. 10.2307/1913646.
18. Wooldridge JM. Econometric Analysis of Cross Section and Panel Data. 2nd ed. MIT Press; 2010.
19. Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. Epidemiology. 2010;21:383–8. 10.1097/EDE.0b013e3181d61eeb.
20. Benjamini Y, Hochberg Y. Controlling the false discovery rate. J R Stat Soc Ser B. 1995;57:289–300. 10.1111/j.2517-6161.1995.tb02031.x.
21. Vollset SE, Ababneh HS, Abate YH, et al. Burden of disease scenarios for 204 countries and territories, 2022–2050. Lancet. 2024;403:2204–56. 10.1016/S0140-6736(24)00685-8.
22. Gelbach JB. When do covariates matter? And which ones, and how much? J Labor Econ. 2016;34(2):509–43. 10.1086/683668.
23. Moran PAP. Notes on continuous stochastic phenomena. Biometrika. 1950;37(1/2):17–23. 10.2307/2332142.
24. Yoon HY, Lee JH, Lee CT et al. Effects of indoor air pollution on clinical outcomes in patients with idiopathic pulmonary fibrosis. Sci Rep. 2024;14. PMID:38262669.
25. Johannson KA, Vittinghoff E, Morisset J, et al. Air pollution exposure is associated with lower lung function in idiopathic pulmonary fibrosis. Chest. 2018;154:119–25. 10.1016/j.chest.2018.01.015.
26. Goobie GC, Carlsten C, Johannson KA, et al. Association of particulate matter exposure with lung function and mortality among patients with fibrotic interstitial lung disease. JAMA Intern Med. 2022;182:1248–59. 10.1001/jamainternmed.2022.4696.
27. Sack C, Wojdyla DM, MacMurdo MG, et al. Long-term air pollution exposure and severity of idiopathic pulmonary fibrosis. Ann Am Thorac Soc. 2025;22:378–86. 10.1513/AnnalsATS.202404-382OC.
28. Lee CT, et al. Inhalational exposures associated with risk of interstitial lung disease: a systematic review and meta-analysis. Thorax. 2025;80:918–26. 10.1136/thorax-2024-222306.
29. Johannson KA, Vittinghoff E, Lee K, et al. Acute exacerbation of idiopathic pulmonary fibrosis associated with air pollution exposure. Eur Respir J. 2014;43:1124–31. 10.1183/09031936.00122213.
30. Sesé L, Nunes H, Cottin V, et al. Role of atmospheric pollution on the natural history of idiopathic pulmonary fibrosis. Thorax. 2018;73:145–50. 10.1136/thoraxjnl-2017-209967.
31. Winterbottom CJ, Shah RJ, Patterson KC, et al. Exposure to ambient particulate matter is associated with accelerated functional decline in idiopathic pulmonary fibrosis. Chest. 2018;153:1221–8. 10.1016/j.chest.2017.07.034.

32. Goobie GC, et al. Ambient ultrafine particulate matter and clinical outcomes in fibrotic interstitial lung disease. Am J Respir Crit Care Med. 2024;209:1111–22. 10.1164/rccm.202306-1027OC.
33. Lee CT, Adegunsoye A, Chung JH, et al. Characteristics and prevalence of domestic and occupational inhalational exposures across interstitial lung diseases. Chest. 2021;160:209–18. 10.1016/j.chest.2021.02.026.
34. Sack CS, Doney BC, Podolanczuk AJ, et al. Occupational exposures and subclinical interstitial lung disease: the MESA Air and Lung Studies. Am J Respir Crit Care Med. 2017;196:1031–9. 10.1164/rccm.201612-2431OC.
35. Walters GI. Occupational exposures and idiopathic pulmonary fibrosis. Curr Opin Allergy Clin Immunol. 2020;20:103–11. 10.1097/ACI.0000000000000617.
36. Huntley CC, Patel K, Mughal AZ, et al. Airborne occupational exposures associated with idiopathic pulmonary fibrosis: a systematic review and meta-analysis. Occup Environ Med. 2023;80:517–26. 10.1136/oemed-2023-108833.
37. Sesé L, Khamis W, Jeny F, et al. Low income and outcome in idiopathic pulmonary fibrosis: an association to uncover. Respir Med Res. 2021;79:100809. 10.1016/j.resmer.2021.100809.
38. Goobie GC, et al. Association of particulate matter exposure with lung function and mortality among patients with fibrotic interstitial lung disease. JAMA Intern Med. 2022;182:1243–53. 10.1001/jamainternmed.2022.4696.
39. Lee CT, Strek ME. Environmental and occupational exposures in interstitial lung disease. Clin Chest Med. 2022;43:631–46. 10.1016/j.ccm.2022.06.006.
40. Cupido G, Meghji J. Post-tuberculosis lung disease and tuberculosis sequelae. Clin Chest Med. 2024;45:355–70. 10.1016/j.ccm.2024.02.010.
41. Allwood BW, Byrne A, Meghji J, Rachow A, van der Zalm MM, Schoch OD. Post-tuberculosis lung disease: clinical review of an under-recognised global challenge. Respiration. 2021;100:751–63. 10.1159/000512531.
42. Copeland CR, et al. Identification and remediation of environmental exposures in patients with interstitial lung disease. Curr Opin Pulm Med. 2021;27:393–400. 10.1097/MCP.0000000000000799.

Figures

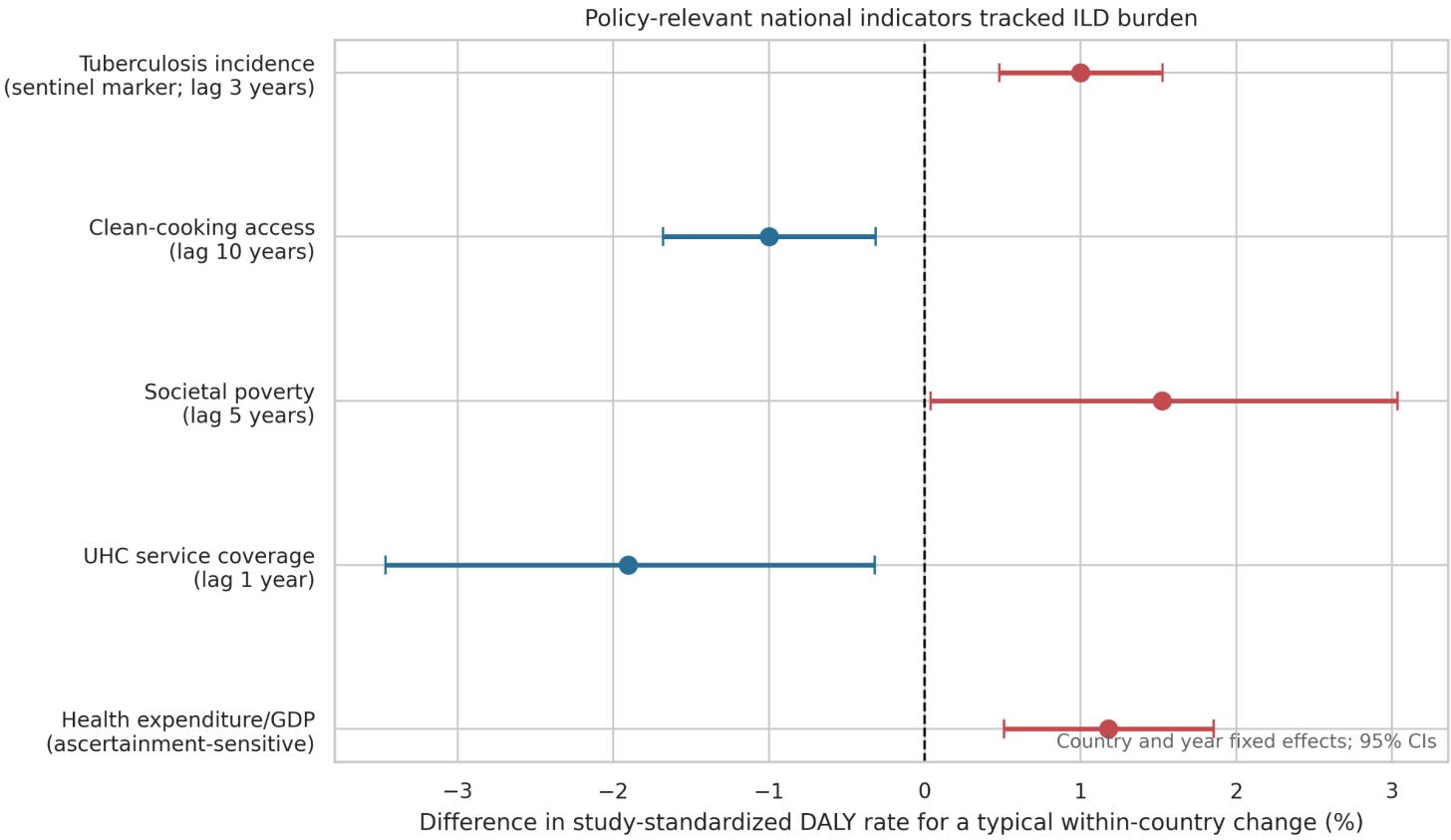


Figure 1

Adjusted associations between modifiable national factors and study-standardized DALY rates. ***(Supplementary Tables S2 and S4).*** The estimates point in a coherent public-health direction across four domains: lower infectious and social vulnerability and better household energy and health coverage accompanied lower burden.

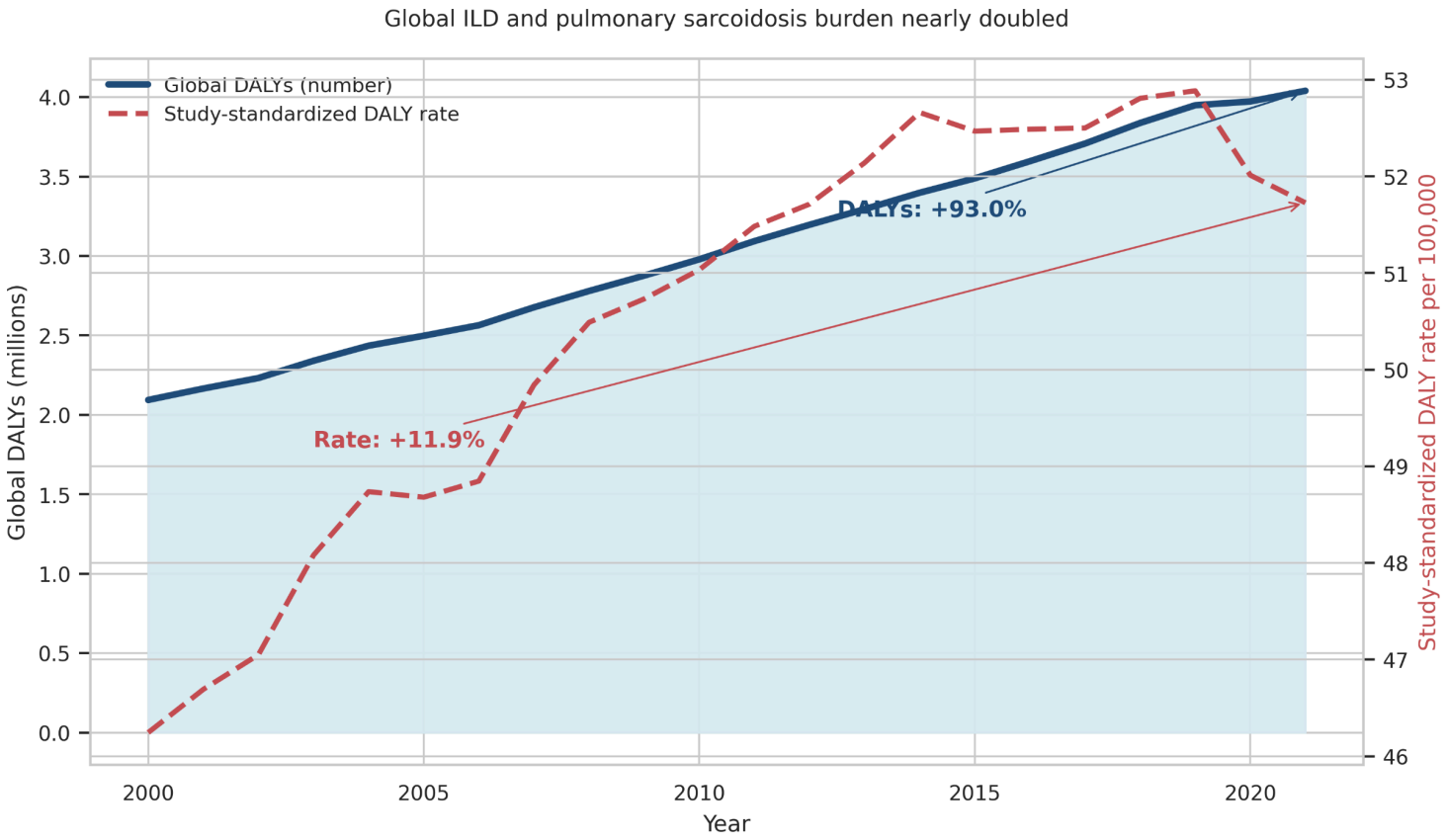


Figure 2

Global ILD and pulmonary sarcoidosis DALYs and study-standardized DALY rate, 2000–2021. ***(Supplementary Table S7).*** DALY numbers rose much faster than the standardized rate, showing that population growth and ageing amplified—without fully explaining—the deterioration in underlying burden.

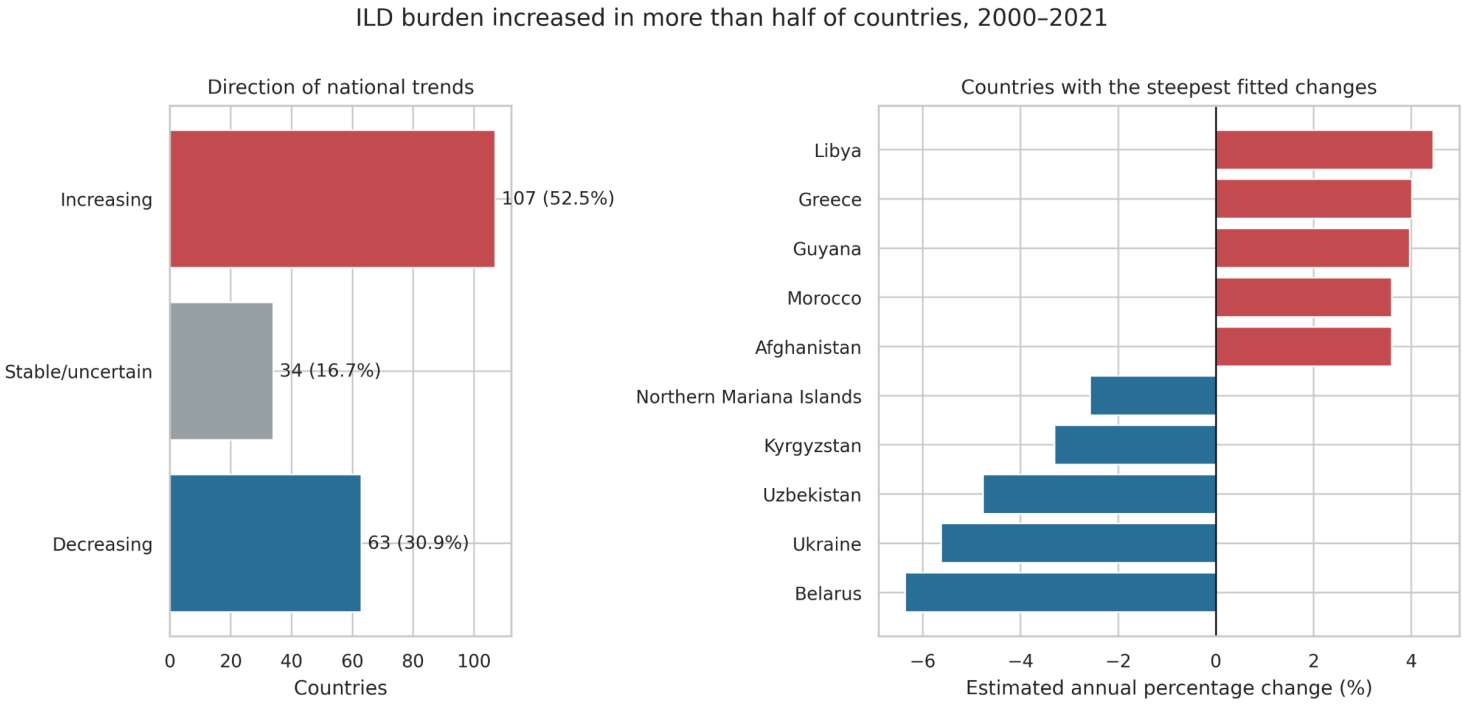


Figure 3

Direction of country trends and countries with the steepest fitted changes. ***(Supplementary Table S1).*** More than half of countries had increasing standardized rates, but the extremes in both directions show that the global trajectory is heterogeneous and potentially alterable.

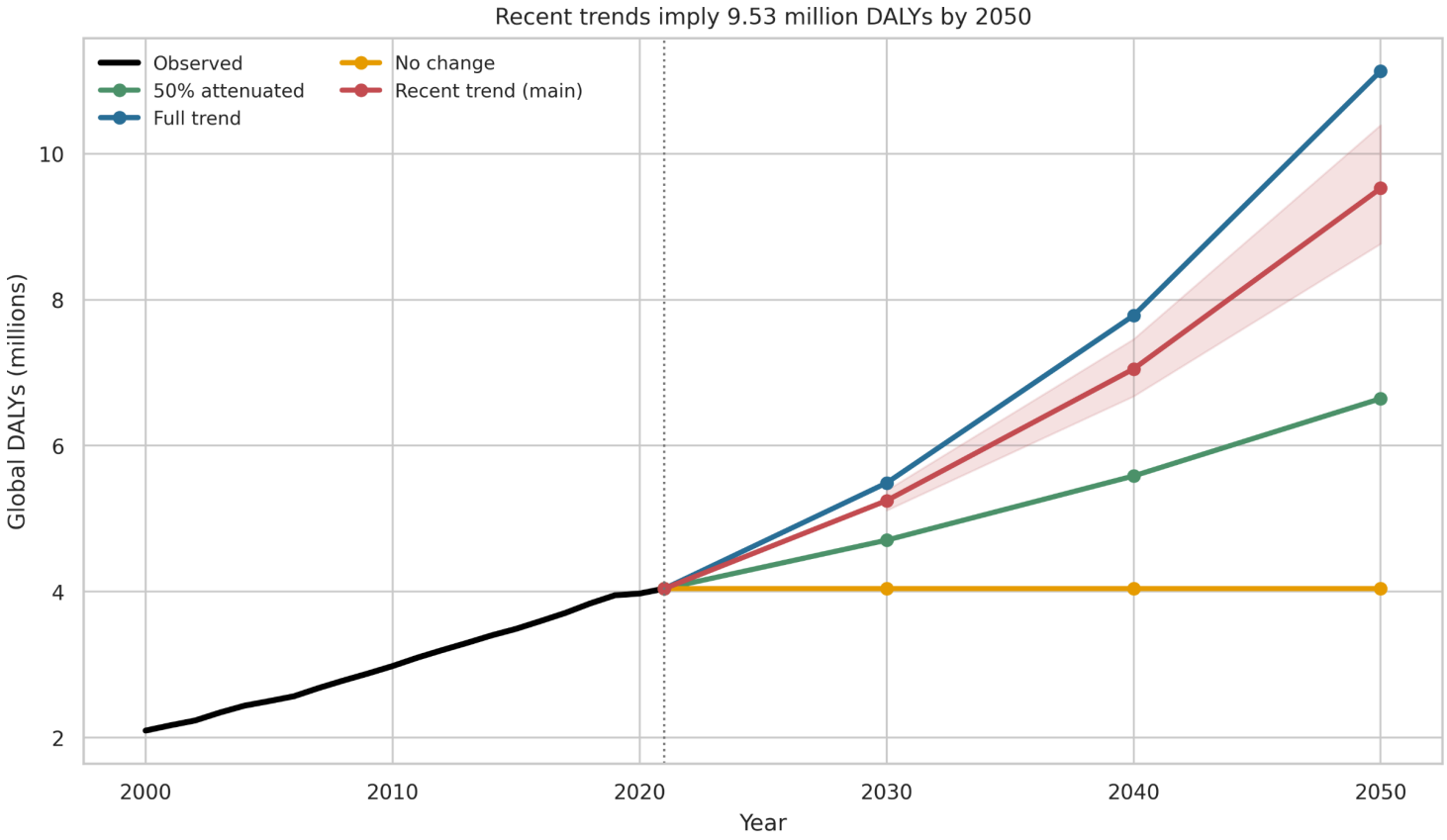


Figure 4

Observed burden and four conditional continuation scenarios to 2050. ***(Supplementary Tables S5, S6, and S8).*** The wide separation between scenarios by 2050 quantifies the consequence of trend persistence; these are conditional trajectories rather than inevitable forecasts.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ILDSupplementaryMaterialEnglishv2.docx