

Beyond Reported Rates: Detection-Adjusted COPD Prevalence and Underdiagnosis Patterns in Colombia

Jorge Ospina ¹, Olga Milena Garcia-Morales ², Maria Clara Gaviria ³

¹Nextmove Consulting, Bogotá, Colombia; ²Department of Pulmonology, Fundación Santa Fe de Bogotá, Bogotá, Colombia; ³Department of Cardiology, Clínica Las Vegas, Medellín, Colombia

Correspondence: Jorge Ospina, Nextmove Consulting, Bogotá, Colombia, Email jorge.ospina@nextmove.com.co

Background: Chronic obstructive pulmonary disease (COPD) is widely underdiagnosed in Colombia, especially in rural departments with limited access to spirometry. We conducted a department-level ecological study using aggregated administrative data from 2020–2023 to generate diagnosis-based COPD prevalence estimates that explicitly account for regional disparities in diagnostic capacity and socioeconomic conditions.

Methods: We assembled department-year data from the Individual Registry of Health Services Delivery, national mortality statistics, and the National Quality of Life Survey. A Bayesian generalized additive model with a Gamma family and log link was fitted to a Composite Bias-Correction Multiplier that captured under-ascertainment as a function of spirometry utilization, COPD lethality, outpatient contact rates, multidimensional poverty, household fuel type, and age structure. Posterior estimates of this multiplier were applied to diagnosis-based COPD prevalence in adults aged ≥ 40 years to obtain detection-adjusted departmental and national estimates. Model performance was summarized using the Bayesian R^2 (proportion of variability in the multiplier explained by the covariates) and the leave-one-out information criterion (LOOIC) as a measure of expected predictive fit.

Results: The model estimated a population-weighted national COPD prevalence of 2.22% (95% credible interval [CrI], 2.21–2.23). Detection-adjusted departmental prevalence ranged from 0.81% in Vichada to 3.50% in Caldas, whereas diagnosis-based prevalence ranged from 0.27% to 2.22%. Spirometry utilization correlated strongly with diagnosis-based prevalence ($r = 0.85$, $p < 0.001$), and departments with higher COPD lethality and greater multidimensional poverty required larger adjustment multipliers. The model explained most of the variability in the Composite Bias-Correction Multiplier (Bayesian $R^2 = 0.99$) and showed good expected predictive performance (LOOIC = -461.1).

Conclusion: COPD prevalence in Colombia shows marked regional heterogeneity driven by demographic risk and uneven diagnostic capacity. Detection-adjusted estimates indicate that the highest burden lies in Andean departments such as Caldas, Boyacá, and Risaralda, while remote Amazon and Orinoco territories experience substantial underdiagnosis. These findings support targeted expansion of spirometry and chronic respiratory care in underserved regions and illustrate how accounting for detection bias can improve chronic disease surveillance in low- and middle-income settings.

Plain Language Summary:

Why was this study done?

Chronic Obstructive Pulmonary Disease (COPD) is a serious lung condition that makes breathing difficult. In Colombia, many people with COPD are not properly diagnosed, especially in rural areas where healthcare services are limited. Previous studies estimating how many people have COPD in Colombia may not be accurate because they did not consider these differences in healthcare access between regions.

What did the researchers do and find?

We studied COPD rates across all 33 departments (states) in Colombia using advanced statistical methods. We looked at factors such as how often spirometry tests are performed, poverty levels, and access to modern cooking stoves. We estimated that 2.22% of Colombians have COPD (95% credible interval [CrI]: 2.21–2.23%), but this varies widely between regions—from 0.81% to 3.50%.

Areas with better access to spirometry, a breathing test used to confirm COPD, showed higher diagnosis rates, which suggests that many cases are missed where testing capacity is limited.

What do these results mean?

These results indicate that diagnosis-based COPD prevalence during 2020–2023 is lower than historical spirometry-based estimates, likely reflecting under-ascertainment in places where spirometry access is limited. None of Colombia's 33 departments reaches the National Administrative Benchmark of 1105 spirometry tests per 100,000 adults aged ≥ 40 years, and only Bogotá approaches half of this target. This information can guide health authorities to expand lung function testing and chronic respiratory care in underserved regions. The same modeling strategy could be applied to other chronic diseases that are underdiagnosed in low-resource settings.

Keywords: chronic obstructive pulmonary disease, Bayesian analysis, spirometry, health equity, epidemiology, ecological, Colombia

Introduction

Chronic Obstructive Pulmonary Disease (COPD) remains a leading cause of morbidity and mortality worldwide, disproportionately affecting low- and middle-income countries where access to timely diagnosis and care is often limited.¹ Underdiagnosis persists even in high-resource systems: about one fifth of patients hospitalized with severe COPD exacerbations have no prior COPD diagnosis, underscoring systematic failures in detection.² Although cigarette smoking is the most recognized risk factor, COPD can also arise from other exposures, including ambient air pollution, biomass smoke, and early-life respiratory infections.^{3–5}

In Colombia, millions of people—particularly in rural areas—continue to rely on biomass fuels like wood for cooking and heating. Biomass smoke exposure has been associated with distinct clinical and inflammatory patterns in COPD, described in GOLD reports as COPD associated with biomass or environmental exposures⁶ and characterized in Latin American cohorts as wood-smoke-related COPD (W-COPD).⁷ Recent histopathologic work has documented specific mast cell-rich inflammation and airway remodeling in biomass-related COPD compared with tobacco-related disease.⁸ These differences matter in practice when diagnosis depends on recognizing less typical presentations in populations with limited access to diagnostic tools.

Furthermore, COPD frequently coexists with cardiovascular disease,^{9–12} creating a complex syndemic burden that complicates diagnosis and increases mortality.

Despite multiple efforts, Colombia still lacks COPD prevalence estimates that fully account for regional disparities in diagnostic capacity. The PREPOCOL study, conducted more than two decades ago in five cities, reported a spirometry-based prevalence of 8.9% but did not include rural departments where biomass exposure and health care access differ substantially.¹³ A later analysis by Gil et al reported a diagnosis-based national prevalence of 5.13% using administrative data,¹⁴ but applied uniform assumptions across departments and did not explicitly adjust for variation in spirometry use or service utilization. Global estimates, such as the 10.2% prevalence reported for Latin America in the GB-COPD study,¹⁵ are informative but cannot fully reflect Colombia's environmental exposures, health system organization, and social gradients.

These limitations have consequences. Without reliable, regionally sensitive estimates, policymakers and health administrators face serious challenges in planning services, allocating resources, or targeting interventions. Spirometry, the reference standard for confirming COPD, remains scarce in rural areas and is underutilized even in many primary care settings worldwide, where most at-risk patients are seen,¹⁶ leading to underdiagnosis and delayed care. Traditional prevalence models often overlook the structural barriers that prevent people from being diagnosed in the first place. Socioeconomic factors add another layer of complexity. Poverty influences both disease risk, through continued biomass exposure, and detection, by limiting access to healthcare services. Yet in many models, these variables are treated as noise rather than core determinants of the observed data.^{1,17,18}

The study period (2020–2023) overlaps with the coronavirus disease 2019 (COVID-19) pandemic, which disrupted outpatient care and spirometry and may have altered coding practices for respiratory diagnoses. Interpretation of diagnosis-based COPD prevalence during these years must consider both underlying disease burden and uneven diagnostic capacity.

To address these gaps, we conducted a department-level ecological study that combines national administrative records, mortality statistics, demographic projections, and socioeconomic indicators to generate diagnosis-based prevalence estimates for all 33 Colombian territories. We used a Bayesian generalized additive model to estimate a Composite Bias-Correction Multiplier that captures detection bias associated with spirometry utilization, COPD lethality, outpatient service use, multidimensional poverty, and age structure. Applying this multiplier to diagnosis-based prevalence yields estimated true prevalence at the department and national levels. Our primary objective is to quantify COPD prevalence in Colombia while explicitly adjusting for detection bias and to describe regional heterogeneity in a way that can guide the expansion of diagnostic and chronic respiratory care in underserved areas.

Materials and Methods

Study Design and Setting

We performed a population-level ecological study that analyzed diagnosis-based COPD prevalence alongside socio-demographic, economic, and health service indicators across Colombia's 33 administrative territories, including 32 departments and the capital district, from 2020 through 2023. All analyses were conducted at the department-year level. Because of this ecological design, findings relate to regional patterns and are not intended to represent individual-level causal associations.

Data Sources

COPD Diagnoses and Health Service Information

Health service information was obtained from the Individual Registry of Health Services Delivery (Registros Individuales de Prestación de Servicios de Salud, RIPS), a national administrative database that compiles outpatient visits, emergency encounters, and hospitalizations reported by providers.¹⁹ We identified COPD diagnoses in adults aged ≥ 40 years using International Classification of Diseases, Tenth Revision (ICD-10) codes J40–J44 recorded as primary or secondary diagnoses. For each department-year, diagnosis-based prevalence was calculated as the number of individuals with at least one recorded COPD diagnosis divided by the population aged ≥ 40 years. These departmental estimates are crude proportions within the ≥ 40 -year population and were not age-standardized, as the outcome is intended to reflect health-system ascertainment rather than age-adjusted disease risk. COPD diagnoses in RIPS reflect physician-entered ICD-10 codes and do not require documented spirometric confirmation, as Colombian health system reporting does not mandate spirometry for diagnostic coding. Our spirometry utilization variable captures all procedures performed regardless of their diagnostic outcome, thus reflecting diagnostic infrastructure availability rather than creating circular dependence with the prevalence outcome.

Spirometry utilization was derived from procedure records in RIPS obtained through the national SISPRO platform. For this study, we counted spirometry procedures coded with either of the following Colombian CUPS codes: 893805 (spirometry or flow–volume curve, pre- and post-bronchodilator) and 893808 (simple spirometry or flow–volume curve). For each department-year, we calculated the annual rate of completed spirometry procedures per 100,000 inhabitants aged ≥ 40 years, using the department population aged ≥ 40 years as the denominator to align with the COPD prevalence outcome. When multiple spirometry procedures were recorded for the same individual within a calendar year, they were counted once for the purpose of calculating departmental rates. At the national level, the Ministry of Health defines an a priori financial planning target of 1105.08 spirometries per 100,000 adults aged ≥ 40 years, which corresponds to the sum of the expected rates for CUPS 893805 (1085.05 per 100,000) and CUPS 893808 (20.03 per 100,000) in the capitation payment schedule (Unidad de Pago por Capitación).²⁰ This National Administrative Benchmark is a budgetary parameter used to allocate annual health system resources, not a clinical guideline or biological standard of care.

Smoking, Socioeconomic, and Demographic Indicators

Data on smoking prevalence and the use of modern cooking stoves (gas or electric) were obtained from the National Quality of Life Survey (Encuesta Nacional de Calidad de Vida, ECV) conducted by the National Administrative Department of Statistics (Departamento Administrativo Nacional de Estadística, DANE).²¹ We used the proportion of households using modern stoves as an inverse proxy for biomass fuel exposure. Population data were taken from the

2018 National Population and Housing Census with projections through 2023.²² We calculated the proportion of the population aged ≥ 40 years in each department-year to represent age structure.

Mortality data were obtained from DANE's official death certificate database, focusing on non-fetal deaths attributed to COPD (ICD-10 J40–J44) in individuals aged ≥ 40 years.²³ Socioeconomic conditions were characterized using DANE's multidimensional poverty index, which incorporates deprivation in health, education, employment, and basic services and is reported as the proportion of individuals living in multidimensional poverty in each department-year.²⁴

Reproducible Workflow and Data Management

All data processing and statistical analyses were executed within a version-controlled pipeline using R (version 4.5.2).²⁵ To ensure integrity and auditability, we strictly avoided manual spreadsheet manipulation. Raw administrative files remain immutable in a structured directory (data/raw), while the preprocessing script (analysis/00_data_preprocessing.R) performs automated cleaning, variable harmonization, and indicator calculation to generate the analysis-ready dataset. The complete analytical workflow—from raw inputs to final posterior estimates—is publicly available in our GitHub repository (<https://github.com/jeospinaa/colombia-copd-bayesian-model>). This repository hosts the raw dataset, all analysis scripts, and the generated tables and figures, alongside a comprehensive data dictionary to enable full independent replication.

Composite Bias-Correction Multiplier

We developed a Composite Bias-Correction Multiplier (Φ) to quantify the magnitude of under-ascertainment for each department-year. This multiplier is constructed from three objective quality-of-care benchmarks: (1) spirometry deficit, defined as the shortfall below the National Administrative Benchmark of 1105 spirometries per 100,000 inhabitants aged ≥ 40 years, a priori financial planning target derived from the expected rates for CUPS 893805 and 893808 in the national capitation schedule and used to allocate annual diagnostic resources; (2) excess lethality, identifying departments where COPD mortality exceeds the national 75th percentile; and (3) access barriers, flagging outpatient contact rates falling below the national 25th percentile. Deviations from these benchmarks contribute additive penalties to a bias score, which is transformed into the multiplier ($\Phi = 1 + \text{BiasScore}$). Consequently, health systems meeting all quality standards receive a baseline multiplier of 1, while structural deficits result in proportionally higher multipliers, correcting for expected underdiagnosis. This algorithm is applied uniformly across all territories via the preprocessing script.

Bayesian Modeling Strategy

We employed a Bayesian Generalized Additive Model (GAM) to estimate the relationship between the Composite Bias-Correction Multiplier and regional determinants. The outcome was modeled using a Gamma distribution with a log link function to accommodate its positive, right-skewed nature. Predictors included the spirometry rate, lethality rate, outpatient service utilization, multidimensional poverty index, biomass fuel usage, and the proportion of the population aged ≥ 40 years. Continuous predictors were modeled using thin-plate regression splines to capture non-linear effects, while the diagnosis-based prevalence was included as a linear anchor.

Model fitting was performed using the brms package, interfacing with Stan for full Bayesian inference via Markov Chain Monte Carlo (MCMC) sampling (4 chains; 4000 iterations; 2000 warm-up).^{25,26} We specified a weakly informative Normal (0, 1) prior on the intercept on the log (linear predictor) scale and regularizing priors for spline coefficients. On the response scale, this intercept prior corresponds to a median multiplier of $\exp(0) = 1.0$, reflecting an expectation of minimal adjustment in departments with optimal diagnostic capacity while remaining sufficiently diffuse to allow data-driven estimates across the observed range. Convergence was verified using trace plots and R-hat statistics ($R^2 < 1.01$). Model fit was assessed via the Bayesian R-squared and the Leave-One-Out Information Criterion (LOOIC).

Prevalence Estimation and Aggregation

We generated the full posterior predictive distribution of the Composite Bias-Correction Multiplier using the *posterior_epred* function, producing 4000 simulated draws that preserve model uncertainty. For each draw, we calculated the department-specific true prevalence by applying the simulated multiplier to the observed diagnosis rate. We derived

national estimates by computing the population-weighted average for each individual posterior draw, utilizing the population aged ≥ 40 years as weights. We report the median and 95% Credible Interval (CrI) of the resulting national distribution, a method that propagates both statistical uncertainty and regional demographic differences.

Results

Diagnosis-Based COPD Prevalence

Based on RIPS records, diagnosis-based COPD prevalence in adults aged ≥ 40 years during 2020–2023 ranged from 0.27% in Vichada to 2.22% in Risaralda. These values reflect recorded diagnoses without adjustment for regional differences in diagnostic capacity.

Estimated True Prevalence and Underdiagnosis by Department

The model estimated a population-weighted national COPD prevalence of 2.22% (95% Credible Interval [CrI]: 2.21–2.23%). Estimated true prevalence varied widely across departments, from 0.81% in Vichada to 3.50% in Caldas.

Table 1 summarizes the diagnosis-based prevalence, estimated true prevalence, adjustment ratio, and 95% CrI for each department. The largest multipliers were observed in remote territories with sparse populations, including Vichada (ratio 2.95), Casanare (2.85), Vaupés (2.79), and Guainía (2.70), indicating substantial under-ascertainment. In contrast, Andean departments such as Caldas, Boyacá, Risaralda, Quindío, and Antioquia exhibited higher estimated true prevalence with more moderate multipliers, consistent with advanced demographic aging and better access to diagnostic services. The geographical distribution of the disease burden is visualized in [Figure 1](#).

Table 1 Diagnosis-Based and Estimated True COPD Prevalence by Department, Adults ≥ 40 years, Colombia 2020–2023 (Median Over Period)

Department	Diagnosis-Based Prevalence	Estimated True Prevalence	Ratio (Estimated / Diagnosis-Based)	95% CrI for Estimated True Prevalence
Caldas	0.0204	0.0350	1.72	[0.0345, 0.0356]
Boyacá	0.0185	0.0334	1.80	[0.0328, 0.0339]
Risaralda	0.0222	0.0329	1.48	[0.0325, 0.0334]
Quindío	0.0196	0.0310	1.58	[0.0305, 0.0314]
Antioquia	0.0178	0.0290	1.63	[0.0286, 0.0294]
Chocó	0.0124	0.0258	2.08	[0.0252, 0.0265]
Bogotá, D.C.	0.0177	0.0248	1.40	[0.0243, 0.0252]
Norte de Santander	0.0131	0.0240	1.83	[0.0237, 0.0243]
Tolima	0.0126	0.0233	1.85	[0.0229, 0.0238]
Valle del Cauca	0.0121	0.0213	1.76	[0.0210, 0.0217]
Huila	0.0114	0.0212	1.86	[0.0209, 0.0215]
Cundinamarca	0.0114	0.0209	1.83	[0.0205, 0.0213]
Santander	0.0127	0.0204	1.61	[0.0201, 0.0208]
Caquetá	0.0102	0.0199	1.95	[0.0196, 0.0202]
Guaviare	0.0107	0.0187	1.75	[0.0184, 0.0190]

(Continued)

Table 1 (Continued).

Department	Diagnosis-Based Prevalence	Estimated True Prevalence	Ratio (Estimated / Diagnosis-Based)	95% CrI for Estimated True Prevalence
Cauca	0.0092	0.0168	1.82	[0.0166, 0.0170]
Nariño	0.0086	0.0165	1.92	[0.0163, 0.0166]
Cesar	0.0084	0.0163	1.94	[0.0161, 0.0164]
Meta	0.0080	0.0160	1.99	[0.0159, 0.0161]
Magdalena	0.0078	0.0154	1.98	[0.0152, 0.0155]
Sucre	0.0081	0.0147	1.81	[0.0145, 0.0149]
Guainía	0.0053	0.0143	2.70	[0.0140, 0.0146]
Arauca	0.0074	0.0142	1.91	[0.0139, 0.0145]
Amazonas	0.0062	0.0142	2.28	[0.0139, 0.0144]
Bolívar	0.0076	0.0141	1.85	[0.0140, 0.0143]
Vaupés	0.0050	0.0141	2.79	[0.0137, 0.0144]
Atlántico	0.0073	0.0134	1.84	[0.0132, 0.0135]
Putumayo	0.0060	0.0133	2.21	[0.0132, 0.0135]
Córdoba	0.0056	0.0126	2.25	[0.0125, 0.0128]
La Guajira	0.0054	0.0121	2.23	[0.0119, 0.0123]
Casanare	0.0036	0.0104	2.85	[0.0102, 0.0107]
Archipiélago de San Andrés	0.0045	0.0097	2.14	[0.0095, 0.0099]
Vichada	0.0027	0.0081	2.95	[0.0079, 0.0082]

Notes: Prevalence values are expressed as proportions (eg, 0.0222 = 2.22%). Estimated prevalence represents the median of the posterior distribution. National Estimate (Population-weighted): 0.0222 (95% CrI: 0.0221–0.0223).

Association Between Spirometry Utilization and Prevalence

Figure 2 illustrates the strong positive correlation ($r=0.85$, $p<0.001$) between spirometry rates and diagnosis-based prevalence. Departments with higher spirometry rates per 100,000 inhabitants consistently reported higher prevalence, suggesting that observed variation is driven largely by case detection rather than biological disease burden alone.

Regional Disparities in Spirometry Utilization

Spirometry utilization varied widely across departments, and no territory reached the National Administrative Benchmark of 1105 procedures per 100,000 inhabitants aged ≥ 40 years.²⁰ Only Bogotá exceeded 50% of this benchmark. Figure 3 illustrates this diagnostic gap, which aligns with the high adjustment ratios observed in most departments.

Model Performance and Covariate Effects

The Bayesian Generalized Additive Model demonstrated excellent convergence, with all R-hat values ≈ 1.00 and robust effective sample sizes (Table 2).

Model fit statistics are summarized in Table 3. The high Bayesian R^2 indicates that the model explains most of the variability in the Composite Bias-Correction Multiplier across departments. The negative LOOIC reflects good expected

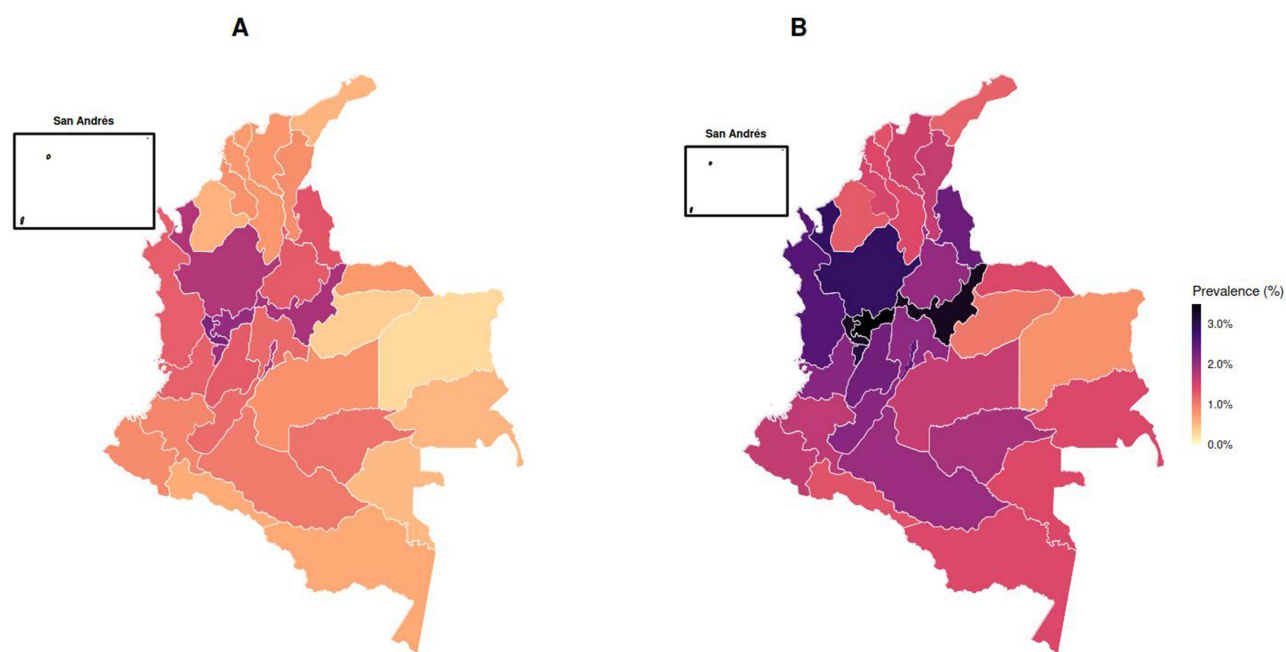


Figure 1 Choropleth maps of COPD prevalence in Colombia. **(A)** Diagnosis-based prevalence. **(B)** Model-estimated true prevalence using the Composite Bias-Correction Multiplier. Both panels use a unified color scale to visualize the magnitude of underdiagnosis.

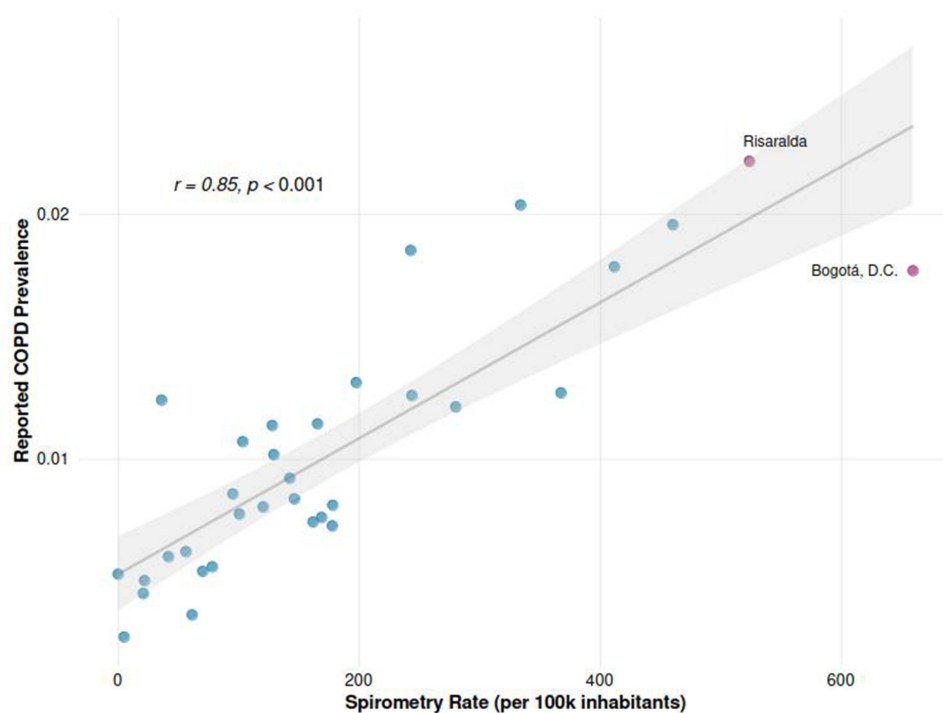


Figure 2 Association between spirometry utilization and diagnosis-based prevalence. Scatterplot with linear regression line and 95% confidence band, illustrating the dependence of case ascertainment on diagnostic infrastructure ($r=0.85$).

predictive performance, and the `elpd_loo` provides a summary of out-of-sample predictive density that can be used for comparison with alternative specifications.

Collectively, these results demonstrate that the model successfully corrects for substantial detection bias, producing estimates that are both statistically stable and epidemiologically coherent with Colombian demographic patterns.

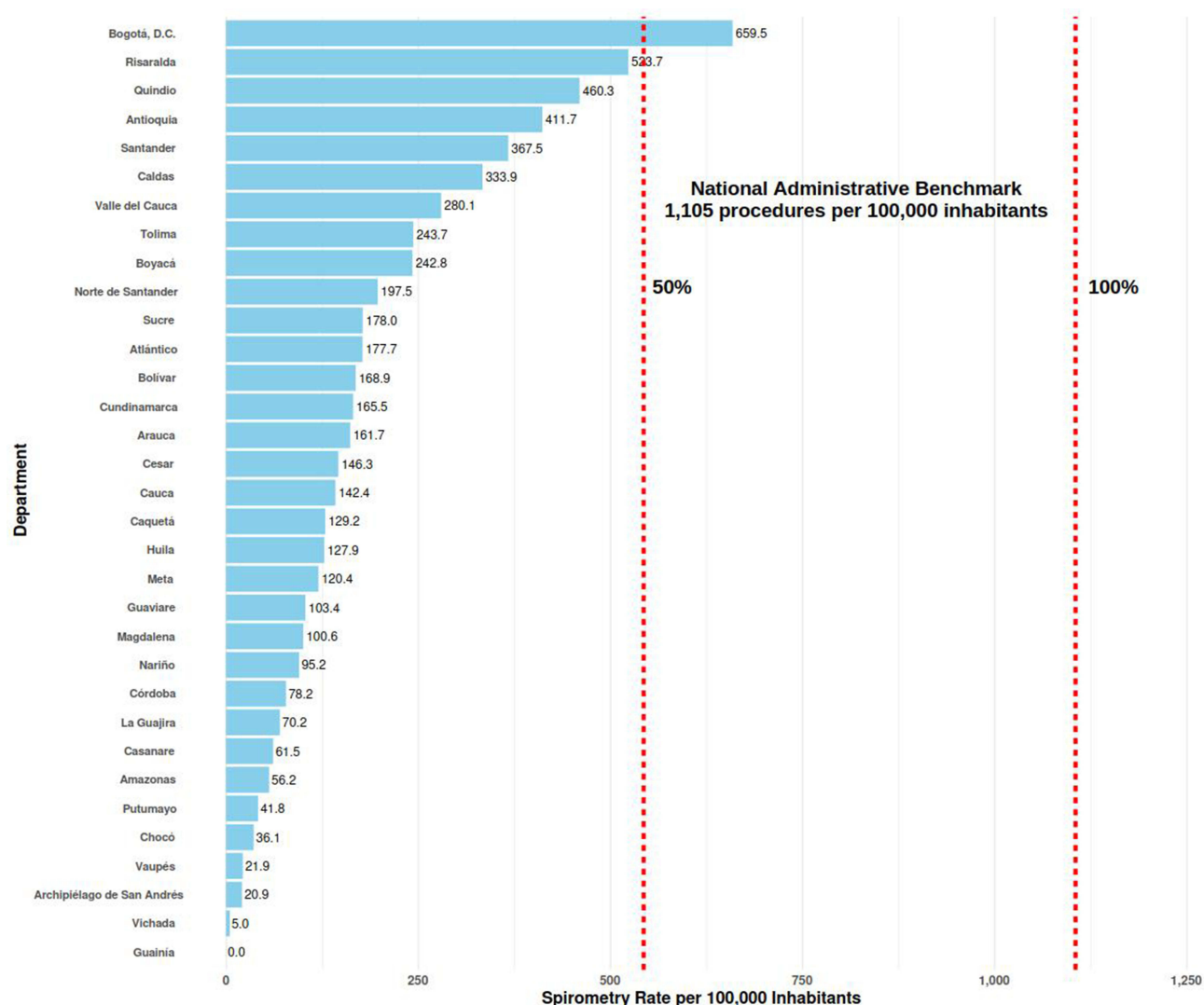


Figure 3 Reported rates of spirometry use in Colombia per 100,000 inhabitants. Bar chart displaying the rate of spirometry procedures performed per 100,000 inhabitants aged ≥ 40 years, with the dashed line marking the National Administrative Benchmark (1105 procedures/100,000), a financial planning benchmark rather than a clinical target.

Discussion

This ecological analysis across Colombia's 33 departments shows that diagnosis-based prevalence largely reflects variation in diagnostic capacity rather than a homogeneous burden of disease. The model-estimated national prevalence was 2.22% (95% CrI: 2.21–2.23), well below the 8.9% spirometry-based prevalence reported in the PREPOCOL survey of five major cities. This difference is not merely numerical; it highlights a deeper issue in how COPD is identified and measured across different settings.

PREPOCOL focused on five urban centers with strong healthcare infrastructure and easy access to spirometry testing, which likely increased case detection in those areas. Their study also reported lower prevalence using medical diagnosis (3.3%) and clinical criteria (2.7%), figures more aligned with our estimates. This suggests that our model, which accounts for access and underdiagnosis, may be capturing symptomatic and clinically relevant COPD rather than detecting asymptomatic individuals based solely on spirometry criteria.¹³

This methodological distinction is clinically important. Population-based studies that rely exclusively on spirometry often classify individuals as having COPD even if they lack symptoms or meaningful exposure histories.²⁷ Colombia's smoking prevalence has declined from 19% to under 7% over recent decades.²¹ However, given the 20–30 year lag between smoking initiation and COPD development, this decline would not yet substantially affect current prevalence

Table 2 Posterior Summaries for Fixed Effects of the Composite Bias-Correction Multiplier Model

Parameter	Estimate	Std. Error	2.5% CrI	97.5% CrI	R-hat	Bulk ESS	Tail ESS
Intercept	0.921	0.058	0.805	1.034	1.001	1596.597	2580.727
Total_prevalence	-22.882	5.669	-33.818	-11.551	1.001	1595.874	2584.444
s(spirometry_rate) (smooth term)	-1.089	0.106	-1.337	-0.920	1.003	1870.456	3020.124
s(lethality_rate) (smooth term)	0.173	0.154	-0.124	0.476	1.001	3739.412	4425.966
s(patients_rate) (smooth term)	-0.028	0.370	-0.766	0.684	1.000	1577.931	2824.674
s(biomass_stove_usage) (smooth)	-0.018	0.092	-0.229	0.169	1.001	2254.335	2840.253
s(multidimensional_poverty_index)	0.144	0.110	-0.037	0.381	1.001	1972.805	3916.279
s(pop_over_40_percent) (smooth)	-0.340	0.097	-0.533	-0.150	1.001	1564.757	3288.027

Notes: Estimates are on the log scale of the Gamma mean. Smooth terms represent the overall effect of each predictor through the spline basis; signs indicate the direction of association across the observed range.

Table 3 Model Performance Metrics

Metric	Value
Number of department-year observations	132
Bayesian R ²	0.99
Expected log pointwise predictive density (elpd_loo)	230.5
Leave-one-out information criterion (LOOIC)	-461.1

among adults aged ≥ 40 years. The lower estimates in our study more likely reflect methodological differences—diagnosis-based ascertainment versus spirometry-based screening conducted more than two decades ago—than temporal changes in exposure patterns.

A key finding is the concentration of COPD burden in the Andean region. Departments such as Caldas (3.50%), Boyacá (3.34%), and Risaralda (3.29%) show the highest estimated true prevalence. These territories combine older age structures with a long history of tobacco use and biomass exposure related to coffee-growing and rural cooking practices, which is consistent with prior clinical descriptions of wood-smoke-related COPD in Colombia.^{7,13} In contrast, remote territories in the Amazon and Orinoco regions, including Vichada (0.81%) and Guainía (1.43%), have lower estimated prevalence despite severe access barriers. Their younger demographic profile and different exposure history likely moderate the absolute burden, even though the diagnostic gap is wide.

Contrary to expectation, the proportion of households using modern cooking stoves did not significantly predict the adjustment multiplier in our model (Table 2). This likely reflects that biomass exposure primarily drives true disease prevalence rather than detection bias, or that current household fuel type incompletely captures the decades-long historical exposure patterns relevant to COPD burden among adults currently aged ≥ 40 years.

The strong correlation between spirometry utilization and diagnosis-based prevalence ($r = 0.85$, $P < 0.001$) underscores how strongly case ascertainment depends on diagnostic infrastructure. Only Bogotá, D.C. approaches at least half of the National Administrative Benchmark of 1105 spirometries per 100,000 inhabitants aged ≥ 40 years, and no department reaches the benchmark itself. Departments with the lowest spirometry rates frequently show higher COPD lethality, which supports the interpretation that late diagnosis and missed opportunities for treatment contribute to excess mortality rather than an intrinsically mild disease profile.

The gap between observed spirometry use and the National Administrative Benchmark likely reflects both system and patient barriers. Spirometers are concentrated in urban centers, many facilities lack staff trained to perform and interpret

the test, and the National Administrative Benchmark itself is a financial planning parameter rather than a clinical quality standard. Current payment schemes do not offer incentives for physicians to order spirometry. In addition, travel distance, indirect costs, and low awareness of chronic respiratory symptoms limit test uptake in rural areas. Addressing these issues will require coordinated changes in equipment distribution, workforce training, and referral organization.

Our findings align with work from other settings showing that health-system constraints shape measured COPD prevalence and widen inequities in diagnosis and outcomes.^{1,15,28,29} Recent analyses have shown that late presentation, limited access to basic diagnostic tools such as spirometry, and failure to deliver core elements of respiratory care closely track social and geographic disadvantage.^{17,18} By combining spirometry rates, outpatient contact, multidimensional poverty, and age structure into a Composite Bias-Correction Multiplier, our model approximates this diagnostic gradient and separates, as far as ecological data allow, disease occurrence from the observation process. The resulting departmental estimates align more closely with demographic and exposure patterns than raw diagnosis rates and provide a more realistic basis for planning services and targeting diagnostic expansion.

Several limitations should be considered. First, this is an ecological study based on department-level aggregates; results describe geographic patterns and cannot be interpreted as causal relationships at the individual patient level. Residual confounding is likely because we lacked individual-level measures of smoking intensity, occupational exposures, and indoor air quality. Second, the study period coincides with the COVID-19 pandemic. We did not perform sensitivity analyses excluding 2020 data; therefore, part of the modeled under-ascertainment may reflect pandemic-related service disruptions superimposed on long-standing structural deficits. Third, our model did not explicitly incorporate spatial random effects or geographic correlation structures; neighboring departments may share unobserved characteristics affecting both disease burden and diagnostic practices. Fourth, we did not age-standardize departmental prevalence estimates, which limits direct comparability with age-standardized survey-based estimates from other sources. Fifth, biomass exposure was approximated using household fuel type, which does not fully capture duration, intensity, or occupational sources of smoke exposure. Finally, administrative data such as RIPS are vulnerable to under-reporting and miscoding, particularly in remote departments, although incorporating COPD lethality and outpatient contact rates into the model partly mitigates this concern.

Despite these limitations, the clinical and policy implications are clear. The highest burden of COPD is concentrated in aging Andean populations, where integrated programs for chronic respiratory and cardiovascular disease could have substantial impact.^{4,5,9–12,28,29} In remote territories with lower estimated prevalence but large diagnostic gaps, priorities include establishing basic spirometry capacity, ensuring referral pathways, and strengthening primary care recognition of chronic respiratory symptoms in people with biomass exposure or prior tuberculosis. Expanding spirometry toward the National Administrative Benchmark, while aligning its use with guideline-based indications rather than indiscriminate screening, is a realistic goal for health-system planning.

Taken together, these results show that COPD prevalence in Colombia is shaped as much by health-system design as by underlying exposures. A modeling framework that accounts for detection bias can support more equitable allocation of diagnostic resources and provide a template for other low- and middle-income countries facing similar constraints in chronic disease surveillance.

Conclusions

This study estimates the national diagnosis-adjusted prevalence of COPD in Colombia at 2.22% (95% CrI: 2.21–2.23), with departmental values ranging from 0.81% to 3.50%. The highest burden is concentrated in Andean departments with older populations and long-standing exposure to tobacco and biomass fuels, whereas remote Amazon and Orinoco territories show lower absolute prevalence but wide diagnostic gaps. The strong correlation between spirometry use and diagnosis-based prevalence ($r = 0.85$, $p < 0.001$) underscores the central role of healthcare infrastructure in shaping surveillance data.

These findings support targeted investments in spirometry access, especially in departments with low diagnostic coverage. Given the frequent overlap with cardiovascular disease, integrated care strategies are warranted.

The modeling approach provides a transferable framework for chronic disease surveillance in low-resource settings. Adjusting for detection bias enables more accurate prevalence estimates, guides resource allocation, and helps ensure all populations are represented in public health data.

Abbreviations

CI, Confidence Interval; COPD, Chronic Obstructive Pulmonary Disease; COVID-19, Coronavirus Disease 2019; CrI, Credible Interval; DANE, Departamento Administrativo Nacional de Estadística (National Administrative Department of Statistics); ECV, Encuesta Nacional de Calidad de Vida (National Quality of Life Survey); elpd_loo, Expected Log Pointwise Predictive Density from Leave-One-Out Cross-Validation; ESS, Effective Sample Size; GAM, Generalized Additive Model; GB-COPD, Global Burden of Chronic Obstructive Pulmonary Disease Study; ICD-10, International Classification of Diseases, Tenth Revision; LOOIC, Leave-One-Out Information Criterion; MCMC, Markov Chain Monte Carlo; National Administrative Benchmark, A priori financial planning target of 1105 spirometries per 100,000 inhabitants aged ≥ 40 years, derived from the expected combined rates for CUPS 893805 (pre- and post-bronchodilator spirometry) and CUPS 893808 (simple spirometry) in the national capitation payment schedule; R, R Statistical Computing Environment; R-hat, Potential Scale Reduction Factor (Gelman–Rubin Statistic); RIPS, Registros Individuales de Prestación de Servicios de Salud (Individual Registry of Health Services Delivery); W-COPD, Wood-Smoke–Related Chronic Obstructive Pulmonary Disease.

Data Sharing Statement

The analytical pipeline and datasets supporting the findings of this study are openly available. The analysis-ready datasets derived from public sources, along with all R code used for data processing, Bayesian modeling, and figure generation, are archived in the project's GitHub repository (<https://github.com/jeospinaa/colombia-copd-bayesian-model>). Original source data can be accessed directly from the Ministry of Health and Social Protection's SISPRO platform for health services data (<https://web.sispro.gov.co/>) and from DANE for demographic, socioeconomic, and mortality data (<http://www.dane.gov.co/>).

Ethics Approval

This study used publicly available, anonymized secondary data obtained from Colombian national health databases (RIPS), demographic surveys (ECV), and mortality registries. In accordance with Resolution 8430 of 1993 from the Colombian Ministry of Health and Social Protection, research involving non-identifiable information from public sources is classified as “research without risk” and is exempt from informed consent requirements and institutional review board approval. No individual patient-level or identifiable information was accessed. All data sources are maintained by Colombian governmental institutions (Ministry of Health and Social Protection, DANE) under their respective privacy and data protection policies.

Consent for Publication

Not applicable. This study used only aggregated, publicly available data with no individual patient information, images, or other identifiable material requiring specific consent for publication.

Acknowledgments

The authors thank the Colombian Ministry of Health and Social Protection for providing access to RIPS data through the SISPRO platform and DANE for making demographic, socioeconomic, and mortality data publicly available. We also acknowledge the healthcare professionals and institutions across Colombia whose consistent data reporting made this analysis possible.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising, or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This study received financial support from AstraZeneca Colombia for data analysis and manuscript preparation. The funder had no role in the design of the study, data collection, data management, statistical analysis, interpretation of results, decision to submit the article for publication, or writing of the manuscript. The authors retained full control over the study protocol, access to the data, and the content of the final article.

Disclosure

Dr. Jorge Ospina, Dr. Olga Milena García-Morales, and Dr. Maria Clara Gaviria report receiving personal fees and honoraria for lectures and medical advisory services from AstraZeneca and other pharmaceutical companies outside the submitted work. The authors declare that these professional activities did not influence the design, analysis, or interpretation of this study. The research grant provided by AstraZeneca Colombia was restricted to operational support and did not include any influence over the scientific content. The authors report no other conflicts of interest in this work.

References

1. Pleasants RA, Riley IL, Mannino DM. Defining and targeting health disparities in chronic obstructive pulmonary disease. *Int J Chron Obstruct Pulmon Dis*. 2016;11:2475–2496. doi:10.2147/COPD.S79077
2. Guecamburu M, Klein E, Verdy G, et al. Factors associated with undiagnosed COPD in patients hospitalised for severe COPD exacerbation. *BMJ Open Respir Res*. 2024;11(1):e002620. doi:10.1136/bmjresp-2024-002620
3. Agusti A, Fabbri LM, Baraldi E, et al. Spirometry: a practical lifespan predictor of global health and chronic respiratory and non-respiratory diseases. *Eur J Internal Med*. 2021;89:3–9. doi:10.1016/j.ejim.2021.04.027
4. Sin DD, Anthonisen NR, Soriano JB, Agusti AG. Mortality in COPD: role of comorbidities. *Eur Respir J*. 2006;28(6):1245–1257. doi:10.1183/09031936.00133805
5. Divo M, Cote C, de Torres JP, et al. Comorbidities and risk of mortality in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2012;186(2):155–161. doi:10.1164/rccm.201201-0034OC
6. Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global strategy for prevention, diagnosis and management of COPD: 2024 report. 2024. Available from: <https://goldcopd.org/2024-gold-report>. Accessed December 15, 2025.
7. Torres-Duque CA, García-Rodríguez MC, González-García M. Is chronic obstructive pulmonary disease caused by Wood Smoke a different phenotype or a different entity? *Arch Bronconeumol*. 2016;52(8):425–431. doi:10.1016/j.arbres.2016.04.004
8. Robles-Hernández R, Rivera RM, Páramo-Pérez M, et al. Immunohistochemical analysis of mastocyte inflammation: a comparative study of COPD associated with tobacco smoking and wood smoke exposure. *Biomedicines*. 2025;13(7):1593. doi:10.3390/biomedicines13071593
9. Polman R, Hurst JR, Uysal OF, Mandal S, Linz D, Simons S. Cardiovascular disease and risk in COPD: a state of the art review. *Expert Rev Cardiovasc Ther*. 2024;22(4–5):177–191. doi:10.1080/14779072.2024.2333786
10. MacLagan LC, Croxford R, Chu A, et al. Quantifying COPD as a risk factor for cardiac disease in a primary prevention cohort. *Eur Respir J*. 2023;62(2):2202364. doi:10.1183/13993003.02364-2022
11. de Miguel-Díez J, Núñez Villota J, Santos Pérez S, et al. Multidisciplinary management of patients with chronic obstructive pulmonary disease and cardiovascular disease. *Arch Bronconeumol*. 2024;60(4):226–237. doi:10.1016/j.arbres.2024.01.013
12. Shrikrishna D, Taylor CJ, Stonham C, Gale CP. Exacerbating the burden of cardiovascular disease: how can we address cardiopulmonary risk in individuals with chronic obstructive pulmonary disease? *Eur Heart J*. 2023;45(4):247–249. doi:10.1093/eurheartj/ehad669
13. Caballero A, Torres-Duque CA, Jaramillo C, et al. Prevalence of COPD in five Colombian cities situated at low, medium, and high altitude (PREPOCOL study). *Chest*. 2008;133(2):343–349. doi:10.1378/chest.07.1361
14. Gil Y, Torres-Duque C, Figueredo MDC, et al. An estimate of the prevalence of COPD in Colombia based on the diagnoses reported in the Individual Registry of Health Services Delivery (RIPS). 2019.
15. Boers E, Barrett M, Su JG, et al. Global burden of chronic obstructive pulmonary disease through 2050. *JAMA Network Open*. 2023;6(12):e2346598. doi:10.1001/jamanetworkopen.2023.46598
16. Kim SH, Han MK. Challenges and the future of pulmonary function testing in Chronic Obstructive Pulmonary Disease (COPD): toward earlier diagnosis of COPD. *Tuberc Respir Dis*. 2025;88(3):413–418. doi:10.4046/trd.2025.0009
17. Fazleen A, Freeman A, Kong A, et al. Health inequality and COPD. *Inequal Respir Health*. 2023;129.
18. Holland AE, Hamidah Eri Rusli Z, Oliveira MR, Bremner J, Rawlings S, Cox NS. Equity in pulmonary rehabilitation delivery: a systematic review and meta-analysis. *Eur Respir Rev*. 2025;34(178):250105. doi:10.1183/16000617.0105-2025
19. Ministerio de Salud y Protección Social. SISPRO - Aplicativos Misionales. 2024. Available from: <https://web.sispro.gov.co/>. Accessed May 20, 2025.

20. Ministerio de Salud y Protección Social. Unidad de pago por capitación (UPC). 2021. Available from: <https://www.minsalud.gov.co>. Accessed May 20, 2025.
21. Departamento Administrativo Nacional de Estadística – DANE. Encuesta Nacional de Calidad de Vida (ECV) 2020–2023. 2023. Available from: <http://www.dane.gov.co/>. Accessed March 15, 2025.
22. Departamento Administrativo Nacional de Estadística – DANE. Proyecciones de población con la actualización post COVID-19. Serie nacional de población por área, sexo y edad para el periodo 2020–2070. 2023.
23. Departamento Administrativo Nacional de Estadística – DANE. Defunciones no fetales en Colombia 2020–2023. 2024. Available from: <http://www.dane.gov.co/>. Accessed March 15, 2025.
24. Departamento Administrativo Nacional de Estadística – DANE. Pobreza multidimensional. 2024. Available from: <https://www.dane.gov.co>. Accessed March 15, 2025.
25. R Core Team. R: a language and environment for statistical computing. 2025. Available from: <https://www.R-project.org/>. Accessed November 7, 2025.
26. Carpenter B, Gelman A, Hoffman MD, et al. Stan: a probabilistic programming language. *J Stat Softw.* 2017;76:1. doi:10.18637/jss.v076.i01
27. Mangione CM, Barry MJ, Nicholson WK, et al. Screening for chronic obstructive pulmonary disease: US preventive services task force reaffirmation recommendation statement. *JAMA.* 2022;327(18):1806–1811. doi:10.1001/jama.2022.5692
28. Hurst JR, Gale CP; Global Working Group on Cardiopulmonary Risk. MACE in COPD: addressing cardiopulmonary risk. *Lancet Respir Med.* 2024;12(5):345–348. doi:10.1016/S2213-2600(24)00038-9
29. Morgan AD, Zakeri R, Quint JK. Defining the relationship between COPD and CVD: what are the implications for clinical practice? *Ther Adv Respir Dis.* 2018;12:1753465817750524. doi:10.1177/1753465817750524

International Journal of Chronic Obstructive Pulmonary Disease

Publish your work in this journal

The International Journal of COPD is an international, peer-reviewed journal of therapeutics and pharmacology focusing on concise rapid reporting of clinical studies and reviews in COPD. Special focus is given to the pathophysiological processes underlying the disease, intervention programs, patient focused education, and self management protocols. This journal is indexed on PubMed Central, MedLine and CAS. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/international-journal-of-chronic-obstructive-pulmonary-disease-journal>

Dovepress
Taylor & Francis Group