INPLASY

Risk Factors for Lung Cancer in Patients with Idiopathic Pulmonary Fibrosis: A Systematic Review and Meta-analysis Protocol

INPLASY202690108

doi: 10.37766/inplasy2026.9.0108

Received: 24 September 2026

Published: 24 September 2026

Corresponding author:
Mei Tian

tianmeicrx@126.com

Author Affiliation:
Shandong First Medical University.

Tian, RX; Tian, Y; Li, ZQ; Liu, X; Tian, L; Tian, M.

ADMINISTRATIVE INFORMATION

Support - No specific financial or non-financial support has been received for this review, and no sponsor will have a role in its design, conduct, analysis, interpretation, or dissemination.

Review Stage at time of this submission - The review has not yet started.

Conflicts of interest - None declared.

INPLASY registration number: INPLASY202690108

Amendments - This protocol was registered with the International Platform of Registered Systematic Review and Meta-Analysis Protocols (INPLASY) on 24 September 2026 and was last updated on 24 September 2026.

INTRODUCTION

Review question / Objective To systematically review and synthesize evidence on prespecified candidate factors associated with lung cancer among patients with idiopathic pulmonary fibrosis, and to estimate the direction and magnitude of these associations. Prespecified subgroup and sensitivity analyses will be performed where data permit. Where at least 10 clinically and methodologically comparable studies are available for a specific exposure–outcome association within the same evidence stratum and effect measure, and with sufficiently comparable exposure definitions, reference groups, and measurement time points, exploratory random-effects univariable meta-regression will be considered to investigate potential sources of heterogeneity. Candidate study-level covariates will include effect-estimate adjustment status (multivariable-adjusted versus unadjusted or univariable), total Newcastle–Ottawa Scale score, mean participant age, proportion of male participants, and prespecified exposure-definition categories. Because cohort and case-control evidence will be synthesized in separate strata, study design will not be used as a covariate unless both designs are included in one prespecified, sufficiently comparable pool. Multivariable meta-regression will be considered only when at least 10 studies are available per covariate, the covariates show adequate variation and no important collinearity, and no more than two covariates are included in a model. All meta-regression analyses will be exploratory and interpreted cautiously.

Rationale Idiopathic pulmonary fibrosis (IPF) is associated with an increased risk of lung cancer. Previous systematic reviews have examined the incidence, prevalence, clinical characteristics, and overall burden of lung cancer in patients with IPF. Some have also examined selected factors, including sex and smoking (JafariNezhad and YektaKooshali, 2018; Brown et al., 2019; Kim et al., 2026). However, evidence concerning a broader

prespecified range of candidate factors associated with lung cancer among patients with IPF remains fragmented. Studies vary in their exposure definitions, reference groups, adjustment strategies, designs, and reported effect measures. A preliminary search of PubMed, INPLASY, and PROSPERO identified related work but did not identify a clearly overlapping INPLASY protocol. A related PROSPERO record (CRD42024420691) examines antifibrotic treatment in patients with both IPF and lung cancer, rather than factors associated with lung cancer occurrence. This review will synthesize associations between prespecified candidate factors and lung cancer in patients with IPF. Distinct evidence strata, exposure definitions, and effect measures will be analyzed separately. Prespecified subgroup and sensitivity analyses, and exploratory meta-regression where sufficient comparable studies are available, will investigate potential sources of heterogeneity.

Condition being studied Idiopathic pulmonary fibrosis is a chronic, progressive fibrosing interstitial pneumonia of unknown cause, characterized by progressive pulmonary fibrosis and decline in lung function. Lung cancer is an important comorbidity in patients with idiopathic pulmonary fibrosis and may develop during follow-up or be identified in comparative studies. This review focuses on factors associated with lung cancer among patients with idiopathic pulmonary fibrosis.

METHODS

Search strategy The first author will search MEDLINE via PubMed, Embase, Web of Science Core Collection, the Cochrane Library, SinoMed/CBM, CNKI, Wanfang Data, and VIP for studies published from January 1985 to the date of the final search. In PubMed, “Idiopathic Pulmonary Fibrosis”, “Lung Neoplasms”, “Small Cell Lung Carcinoma”, and “Carcinoma, Non-Small-Cell Lung” will be used as MeSH headings. Appropriate free-text terms and subheadings will be selected accordingly. The draft PubMed search strategy is:

((("Idiopathic Pulmonary Fibrosis"[Mesh]) AND ("Lung Neoplasms/complications"[Mesh] OR "Small Cell Lung Carcinoma/complications"[Mesh] OR "Carcinoma, Non-Small-Cell Lung/complications"[Mesh])) OR (("Familial Idiopathic Pulmonary Fibrosis"[Title/Abstract] OR "Pulmonary Fibrosis, Idiopathic"[Title/Abstract] OR "IPF"[Title/Abstract]) AND ("Lung Cancer"[Title/Abstract] OR "Cancer of Lung"[Title/Abstract] OR "Neoplasms, Lung"[Title/Abstract] OR "Neoplasms, Pulmonary"[Title/Abstract] OR "Pulmonary Cancer"[Title/Abstract] OR "Pulmonary Neoplasms"[Title/Abstract] OR "Carcinoma, Small Cell Lung"[Title/Abstract] OR "Small Cell Lung Cancer"[Title/Abstract] OR "Carcinoma, Non-Small Cell Lung"[Title/Abstract] OR "Non-Small Cell Lung Cancer"[Title/Abstract] OR "Non-Small Cell Lung Carcinoma"[Title/Abstract] OR "Non-Small-Cell Lung Carcinoma"[Title/Abstract] OR "Nonsmall Cell Lung Cancer"[Title/Abstract]))) NOT (("Animals"[Mesh] NOT "Humans"[Mesh] OR "Meta-Analysis" [Publication Type] OR "Systematic Review" [Publication Type] OR "Randomized Controlled Trial" [Publication Type]))

The strategy will be adapted to each database’s indexing and syntax. The corresponding author will review the planned strategies before execution. The first author will document the executed strategies, database interfaces, search dates, limits, and numbers of records retrieved in a search log. Reference lists of included studies and relevant reviews will be screened, and forward citation searching will be performed where available. One final search is planned; no scheduled update search will be conducted.

Participant or population Adults with idiopathic pulmonary fibrosis diagnosed according to high-resolution computed tomography and/or histopathological criteria will be eligible. Studies must assess lung cancer development or occurrence or compare participant characteristics according to lung cancer status. Eligible participants will be aged 18–85 years; studies with a broader age range will be eligible only if data for participants aged 18–85 years can be extracted separately.

Intervention Not applicable. This is an observational systematic review and meta-analysis of risk-factor and etiologic associations rather than an intervention review. The exposures of interest are prespecified candidate factors associated with lung cancer among patients with idiopathic pulmonary fibrosis.

Comparator For each candidate exposure, the reference group will be the study-defined unexposed group or the lowest prespecified exposure category. For cohort studies, incident lung cancer during follow-up will be compared with no incident lung cancer among patients with idiopathic pulmonary fibrosis. For case-control studies, patients with idiopathic pulmonary fibrosis and lung cancer will be compared with patients with idiopathic pulmonary fibrosis without lung cancer.

Study designs to be included Cohort studies and case-control studies.

Eligibility criteria Inclusion criteria (all criteria must be met):
1. Studies involving adults with idiopathic pulmonary fibrosis (IPF) diagnosed according to high-resolution computed tomography and/or histopathological criteria. Participants will be aged 18–85 years; studies with broader age ranges will be eligible only if data for participants aged 18–85 years can be extracted separately. This restriction is based on the age range commonly used in previous studies of adult IPF populations. The lower age limit is intended to restrict the review to adults, whereas the upper age limit is a methodological criterion intended to reduce heterogeneity related to extreme age, comorbidity burden, competing mortality, and potentially incomplete diagnostic or follow-up assessments. This age range is not intended to represent a biological age range for IPF-associated lung cancer.
2. Studies reporting at least one prespecified candidate exposure or risk factor associated with lung cancer development or occurrence among patients with IPF. Candidate factors may include demographic characteristics; lifestyle and exogenous exposures; anthropometric, metabolic, and comorbidity-related factors; baseline IPF disease status and pulmonary function; blood gas and oxygenation; imaging and pulmonary structural phenotypes; treatment and healthcare exposures; and specific serum biomarkers. Symptoms, signs, and general inflammatory or blood-cell indices will not be considered core candidate exposures for the primary synthesis.
3. Studies reporting an effect estimate for the association between a candidate factor and lung cancer, such as a hazard ratio (HR), odds ratio (OR), or risk ratio (RR) with a 95% confidence interval (CI), or providing sufficient raw data to calculate an effect estimate and its standard error or 95% CI.
4. Cohort studies or case-control studies.
5. For cohort studies, a follow-up period of at least 12 months is required. Case-control studies must report a clearly defined comparison between patients with IPF and lung cancer and patients with IPF without lung cancer.
6. Publications in English or Chinese.
7. Studies published from January 1985 to the date of the final search.
Exclusion criteria (any one criterion will result in exclusion):
1. Studies involving non-IPF populations or studies that do not report a clear basis for the diagnosis of IPF.
2. Studies that do not report a directly extractable effect estimate and do not provide sufficient raw data to calculate an effect estimate and its precision.
3. Multiple publications reporting data from the same cohort will be linked as one study. For the primary synthesis, the report with the most complete exposure–outcome data and the most fully adjusted estimates will be prioritized. If reports are otherwise equivalent, the report with the largest sample size and longest follow-up will be retained. Other reports will be used only for non-overlapping information and will not be double-counted.
4. Publications in languages other than English or Chinese.
5. Non-original studies, including narrative reviews, systematic reviews, meta-analyses, guidelines, expert consensuses, commentaries, editorials, conference abstracts, and letters.
Uncertain eligibility will be retained for full-text assessment and resolved through discussion between the first and second authors. Unresolved disagreements will be adjudicated by the third author and, if necessary, the fourth author.

Information sources The primary information sources will be MEDLINE via PubMed, Embase, Web of Science Core Collection, the Cochrane Library, SinoMed/CBM, CNKI, Wanfang Data, and VIP. Additional targeted searches using exposure-related terms, key authors, and relevant journals will be undertaken where needed. Reference lists of included studies and relevant reviews will be screened, and forward citations will be checked through available citation databases and cited-by functions. Reviews will be used to identify potentially eligible original studies but will not themselves be included as primary studies. Full texts and supplementary materials will be obtained through journal websites, institutional access, or other publicly available sources. Grey literature and study registries will not be searched as sources of primary studies. Study authors may be contacted to clarify unclear information or obtain missing data. No scheduled update search is planned. Records from all sources will be documented, deduplicated, and screened against the prespecified eligibility criteria.

Main outcome(s) The primary outcome is lung cancer development or occurrence among patients with idiopathic pulmonary fibrosis. In cohort studies, this will be defined as incident lung cancer diagnosed during follow-up after IPF diagnosis or baseline assessment. In case-control studies, the outcome will be lung cancer status, comparing patients with IPF and lung cancer with patients with IPF without lung cancer. Reported effect

estimates will include hazard ratios (HRs), odds ratios (ORs), and risk ratios (RRs), together with their 95% confidence intervals. When a directly reported effect estimate is unavailable but sufficient raw data are available, an effect estimate will be derived; an OR will be calculated preferentially for binary comparative data when appropriate. Adjusted estimates will be prioritized when available. HRs, ORs, and RRs will be synthesized separately where appropriate. Incidence-cohort evidence and comparative case-control evidence will be analyzed as separate evidence strata.

Data management Records will be imported into Zotero 10 reference-management software for deduplication and record management. The first and second authors will independently screen titles and abstracts and assess potentially eligible full texts against the prespecified eligibility criteria. A standardized data-extraction form will be piloted before formal extraction. The first and second authors will independently extract study and participant characteristics, exposure and outcome definitions, follow-up duration, effect estimates, adjustment status, and available study-level covariates, and will cross-check the extracted data. Multiple reports from the same study will be linked and treated as one study, and reasons for full-text exclusion will be recorded. Missing or unclear information will be requested from study authors where necessary. Disagreements will be resolved through discussion, followed by adjudication by the third author and, if necessary, the fourth author. Data will be stored in a secure project repository.

Quality assessment / Risk of bias analysis The first and second authors will independently assess the methodological quality and risk of bias of each included cohort and case-control study using the appropriate cohort or case-control version of the Newcastle–Ottawa Scale (NOS). The assessment will cover the domains of selection, comparability, and outcome or exposure ascertainment. The total NOS score and domain-level judgments will be recorded. The total NOS score will be included as a prespecified study-level covariate in exploratory meta-regression when the conditions specified in Item 8 are met. Disagreements will be resolved through discussion; if consensus cannot be reached, the third author will adjudicate. If the disagreement remains unresolved, the fourth author will provide final adjudication. NOS assessments will be considered in the interpretation of findings and prespecified sensitivity analyses.

Strategy of data synthesis Where at least two clinically and methodologically comparable studies are available for a given exposure–outcome association, quantitative synthesis will be performed using a random-effects model. Studies will be pooled only when exposure definitions, reference groups, measurement time points, outcome definitions, and evidence strata are sufficiently comparable. Different definitions or thresholds of the same exposure will be synthesized separately when at least two comparable studies are available for each definition; findings based on a single study will be summarized narratively. Effect estimates will be pooled separately according to the effect measure; odds ratios (ORs), hazard ratios (HRs), and risk ratios (RRs) will not be combined in the same primary analysis. Adjusted or multivariable estimates will be prioritized when available. If a study does not report an adjusted or multivariable estimate, its univariable estimate may be included in the primary synthesis when it meets the prespecified comparability criteria. If multiple estimates are reported for the same study, exposure definition, and outcome, one estimate will be selected for each primary synthesis according to the following hierarchy: the most fully adjusted estimate, followed by the most appropriate univariable estimate, and then an eligible crude estimate when no model-based estimate is available. Other estimates will be extracted and considered in separate sensitivity analyses or strata. A study may contribute to analyses of different exposure definitions, but its correlated estimates will not be treated as independent observations within the same meta-analysis. ORs and RRs will not be converted into one another. When 2×2 data are available, a crude OR may be calculated as ad/bc from a 2×2 table, where a and b are exposed and unexposed lung cancer cases, and c and d are exposed and unexposed non-cases, respectively; this will be treated as a calculated OR. HRs, MDs, and SMDs will not be converted to other effect measures. For continuous variables, MDs will be used when the same scale and unit are reported, whereas SMDs will be used when different scales measure the same construct. When at least 10 clinically and methodologically comparable studies are available within the same evidence stratum and effect measure, exploratory random-effects univariable meta-regression will be considered using the prespecified study-level covariates. Each study will contribute one effect estimate to a given meta-regression, and ratio measures will be analyzed on the log scale. Multivariable meta-regression will be considered only when at least 10 studies are available per covariate, the covariates show

adequate variation and no important collinearity, and no more than two covariates are included in a model. Meta-regression findings will be interpreted as exploratory evidence of potential sources of heterogeneity rather than definitive or causal associations. Statistical heterogeneity will be assessed using I^2 and τ^2. When fewer than two comparable studies are available or quantitative pooling is otherwise inappropriate, findings will be synthesized narratively.

Subgroup analysis Prespecified subgroup analyses will be conducted when at least two clinically and methodologically comparable studies are available. For each exposure, studies will first be stratified according to exposure definition. Studies with identical or clinically comparable exposure definitions will be pooled. Studies with different exposure definitions will be synthesized separately when at least two studies are available within each definition-specific subgroup; an exposure definition represented by only one study will be described narratively and will not be quantitatively pooled. For the same effect measure, multivariable-adjusted and unadjusted or univariable estimates may be included in the primary synthesis when otherwise comparable; when both are reported by the same study, the multivariable estimate will be prioritized. For continuous variables, studies will be pooled only when the construct, unit, reference group, and measurement time point are sufficiently comparable. Mean differences will be used for the same scale and unit, whereas standardized mean differences will be used for different scales measuring the same construct. Studies with clearly different measurement time points will not be forced into the same subgroup. Where at least 10 clinically and methodologically comparable studies are available for a specific exposure–outcome association within the same evidence stratum and effect measure, exploratory random-effects univariable meta-regression will be considered. Candidate study-level covariates will include effect-estimate adjustment status, total Newcastle–Ottawa Scale score, mean participant age, proportion of male participants, and prespecified exposure-definition categories. Meta-regression findings will be interpreted as exploratory evidence of potential sources of heterogeneity. If meta-regression is not feasible, eligible studies may still be pooled under the prespecified criteria; otherwise, findings will be summarized narratively.

Sensitivity analysis Prespecified sensitivity analyses will be conducted to assess the robustness of the primary findings. First, univariate and crude estimates will be excluded, with adjusted or multivariable estimates retained when available. Second, studies with a Newcastle–Ottawa Scale (NOS) score of 4 or less will be excluded and the meta-analysis will be repeated. Sensitivity analyses will be performed within the relevant exposure-definition, effect-measure, and evidence strata. ORs, HRs, and RRs will remain separate throughout these analyses. If at least two studies remain after an exclusion, a new quantitative meta-analysis will be performed. If only one study remains, no quantitative sensitivity meta-analysis will be conducted, and the limited amount of higher-quality evidence will be reported narratively.

Language restriction Due to translation limitations, only publications in English or Chinese will be included.

Country(ies) involved China.

Keywords Idiopathic pulmonary fibrosis; lung cancer; risk factors; observational studies; systematic review; meta-analysis.

Dissemination plans Dissemination is planned through publication in a peer-reviewed journal.

Contributions of each author

Author 1 - Ruixi Tian - The first author will conduct the literature search, screen and select studies, extract data, assess study quality and risk of bias, perform the statistical analysis, and draft the initial manuscript.
Email: 13156101782@163.com
Author 2 - Yuan Tian - The second author will independently screen and select studies, extract data, and assess study quality and risk of bias; these tasks will be performed in parallel with the first author, and the second author will prepare the figures.
Email: tytytianyuan@aliyun.com
Author 3 - Zhuoqi Li - The third author will contribute to the conceptualization and design of the study and will adjudicate disagreements arising during study selection, data extraction, or risk-of-bias assessment.
Email: lzq940705@bjmu.edu.cn
Author 4 - Xue Liu - The fourth author will provide statistical advice, contribute to the interpretation of the findings, and critically review and revise the manuscript.
Email: liuxueszy@126.com
Author 5 - Li Tian - The fifth author will provide statistical advice, contribute to the interpretation of the findings, and critically review and revise the manuscript.

Email: tianli3432@163.com
Author 6 - Mei Tian - The corresponding author will contribute to the conceptualization and design of the study, including the development of the search strategy, supervise the statistical analysis, and critically review and revise the manuscript.
Email: tianmeicrx@126.com