SUPPLEMENTARY MATERIAL

LUNG CANCER OCCURRENCE AFTER AN EPISODE OF TUBERCULOSIS: A SYSTEMATIC REVIEW AND META-ANALYSIS

AUTHORS

Cabrera-Sanchez J, Cuba V, Vega V, Van der Stuyft P, Otero L

TABLE OF CONTENTS

Figure 4. Forest plot for the association between tuberculosis and subsequent lung cancer mortality among cohort studies (model 2)

Appendix 11. Effect estimates between tuberculosis and lung cancer diagnosis from the three cohort studies included reporting by latency strata

Appendix 12. Pooled adjusted estimates from cohort studies excluding lung cancer cases detected within one or two years of tuberculosis diagnosis

Appendix 13. Funnel plots

Figure 1. Adjusted estimates from cohort studies that report the association between tuberculosis and lung cancer diagnosis

Figure 2. Adjusted estimates from case-control studies that report the association between tuberculosis and lung cancer diagnosis

Figure 3. Adjusted estimates from cohort studies that report the association between tuberculosis and lung cancer mortality

Appendix 14. GRADE assessment of the evidence

Appendix 15. List of variables adjusted for in multivariate analyses by included studies

Appendix 16. Amendments to the protocol

Appendix 17. List of excluded studies with reasons

Appendix 1. PRISMA checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	Tittle
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	Summary
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	Introduction
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	Introduction
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	Methods (Search strategy and inclusion criteria)
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	Methods (Search strategy and inclusion criteria)
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	Appendix 2
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	Methods (Search strategy and inclusion criteria)
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	Methods (Data extraction and risk of bias assessment)
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	Methods (Data extraction and risk of bias assessment)

Section and Topic	Item #	Checklist item	Location where item is reported
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	Methods (Data extraction and risk of bias assessment)
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	Methods (Data extraction and risk of bias assessment), Appendix 3 and 4
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	Methods (Statistic analysis)
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	Methods (Statistic analysis)
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	Methods (Statistic analysis)
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	Methods (Statistic analysis)
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	Methods (Statistic analysis)
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	Methods (Statistic analysis)
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	Methods (Statistic analysis)
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	Methods (Statistic analysis)
Certainty	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	Methods

Section and Topic	Item #	Checklist item	Location where item is reported
assessment			(Statistic analysis)
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	Results, figure 1; Appendix 6
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	Appendix 16
Study characteristics	17	Cite each included study and present its characteristics.	Results; Appendix 7
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	Appendix 8
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	Appendix 9
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	Results; Discussion
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	Results; Tables 1 and 2; Appendix 10
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	Table 2
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	Tables 1 and 2
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	Appendix 12
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	Appendix 13
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	Discussion
	23b	Discuss any limitations of the evidence included in the review.	Discussion
	23c	Discuss any limitations of the review processes used.	Discussion
	23d	Discuss implications of the results for practice, policy, and future research.	Discussion
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	Methods
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	Methods
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	Appendix 15

Section and Topic	Item #	Checklist item	Location where item is reported
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	Role of the funding source
Competing interests	26	Declare any competing interests of review authors.	Declaration of interests
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	Data sharing

Appendix 2. Search strategy

Search strategy: PubMed

1. Search **tuberculosis[MeSH Terms]**
2. Search **tuberculosis[Title/Abstract]**
3. Search **mycobacterium[Title/Abstract]**
4. Search **"tb"[Title/Abstract]**
5. Search **"tbc"[Title/Abstract]**
6. 1 OR 2 OR 3 OR 4 OR 5
7. Search **lung neoplasm[MeSH Terms]**
8. Search **lung cancer[MeSH Terms]**
9. Search **lung cancer*[Title/Abstract]**
10. Search **lung neoplasm*[Title/Abstract]**
11. Search **lung carcinoma*[Title/Abstract]**
12. Search **lung tumor*[Title/Abstract]**
13. Search **pulmonary cancer*[Title/Abstract]**
14. Search **pulmonary neoplasm*[Title/Abstract]**
15. Search **pulmonary carcinoma*[Title/Abstract]**
16. Search **pulmonary tumor*[Title/Abstract**
17. Search **"cancer of the lung"[Title/Abstract]**
18. Search **"neoplasm of the lung"[Title/Abstract]**
19. Search **"tumor of the lung"[Title/Abstract]**
20. 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 16 OR 17 OR 18 OR 19
21. 6 AND 20
22. Search **("case reports"[Publication Type] OR "comment*"[Publication Type] OR "Autobiography"[Publication Type] OR "Biography"[Publication Type] OR "legal case"[Publication Type])**
23. 21 NOT 22

Filters: **Publication date from 1980/01/01 to 2020/06/24; English; French; Spanish**

Search strategy: Scopus

1. TITLE-ABS-KEY ("tuberculosis")
2. TITLE-ABS-KEY ("mycobacterium infection")
3. 1 OR 2
4. TITLE-ABS-KEY ("lung cancer")
5. TITLE-ABS-KEY ("lung neoplasm")
6. TITLE-ABS-KEY ("lung tumor")
7. TITLE-ABS-KEY ("lung carcinoma")
8. TITLE-ABS-KEY ("lung adenocarcinoma")
9. TITLE-ABS-KEY ("pulmonary cancer")
10. TITLE-ABS-KEY ("pulmonary neoplasm")
11. TITLE-ABS-KEY ("pulmonary tumor")
12. TITLE-ABS-KEY ("pulmonary carcinoma")
13. TITLE-ABS-KEY ("pulmonary adenocarcinoma")
14. TITLE-ABS-KEY ("cancer of the lung")
15. TITLE-ABS-KEY ("cancer of lung")
16. TITLE-ABS-KEY ("neoplasm of the lung")
17. TITLE-ABS-KEY ("neoplasm of lung")
18. TITLE-ABS-KEY ("tumor of lung")
19. TITLE-ABS-KEY ("tumor of the lung")))
20. 4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 16 OR 17 OR 18 OR 19
21. 3 AND 20

Filters:
22. LIMIT-TO (DOCTYPE , "ar")
23. LIMIT-TO (DOCTYPE , "re")
24. LIMIT-TO (DOCTYPE , "le")
25. 22 OR 23 OR 24
26. LIMIT-TO (SUBJAREA , "MEDI")
27. LIMIT-TO (SUBJAREA , "BIOC")
28. LIMIT-TO (SUBJAREA , "IMMU")
29. 26 OR 27 OR 28
30. LIMIT-TO (LANGUAGE , "English")
31. LIMIT-TO (LANGUAGE , "French")
32. LIMIT-TO (LANGUAGE , "Spanish")
33. 30 OR 31 OR 32
34. LIMIT-TO (PUBYEAR , 1980-2021)
35. 25 AND 29 AND 33 AND 34

Search strategy: Lilacs

1. tw:(tuberculosis) OR
2. tw:(mycobacterium tuberculosis) OR
3. tw:("TB") OR
4. tw:(mycobacterium infection)
5. 1 OR 2 OR 3 OR 4
6. tw:(lung cancer) OR
7. tw:(lung neoplasm) OR
8. tw:(small cell carcinoma) OR
9. tw:(lung tumor) OR
10. tw:(lung malignancy) OR
11. tw:("cancer of the lung") OR
12. tw:(non-small cell carcinoma)
13. 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12
14. 5 AND 13

Search strategy: Scielo

1. ti:(tuberculosis) OR
2. ab:(tuberculosis) OR
3. ti:(TB) OR
4. ab:(TB) OR
5. ti:(mycobacterium infection)
6. ab:(mycobacterium infection)
7. 1 OR 2 OR 3 OR 4 OR 5 OR 6
8. ti:(lung cancer)
9. ab:(lung cancer)
10. ti:(lung neoplasm)
11. ab:(lung neoplasm)
12. ti:(lung tumor)
13. ab:(lung tumor)
14. ti:(lung malignancy)
15. ab:(lung malignancy)
16. 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14 OR 15

Search strategy: Cochrane

#1: (tuberculosis):ti,ab,kw OR (TB):ti,ab,kw OR ("Mycobacterium"):ti,ab,kw
#2: ("lung cancer"):ti,ab,kw OR ("lung neoplasm"):ti,ab,kw OR ("lung adenocarcinoma cell"):ti,ab,kw OR ("small cell lung cancer"):ti,ab,kw OR ("non small cell lung cancer"):ti,ab,kw
#3: #1 AND #2

The search strategy was developed by LO (MD, PhD), JACS (medical student) and VC (medical student).

Appendix 3. Modified Newcastle-Ottawa Quality Assessment Scale (NOS)

Modified Newcastle - Ottawa Quality Assessment Scale

Cohort Studies

Selection
1) Representativeness of the exposed cohort
a) truly representative **(one star)**
b) somewhat representative **(one star)**
c) selected group of users
d) no description of the derivation of the cohort

2) Selection of the non exposed cohort
a) drawn from the same community as the exposed cohort **(one star)**
b) drawn from a different source
c) no description of the derivation of the non exposed cohort

3) Ascertainment of exposure
a) Bacteriologically confirmed TB episode, from NTP/medical records **(two stars)**
b) Bacteriologically confirmed episode, from a structured interview **(one star)**
c) Clinically diagnosed TB episode from NTP/medical records **(one star)**
d) Structured interview with no information on bacteriological diagnosis
e) Self-report
f) No description
g) Other

4) Attempt to Demonstrate that outcome of interest was not present at start of study
a) yes (excluded cases occurring in the first year after the tuberculosis diagnosis or performed latency analysis) **(one star)**
b) no

Comparability
1) Comparability of cohorts on the basis of the design or analysis
a) study controls age AND smoking **(two stars)**
b) study controls for age OR smoking **(one star)**

Outcome
1) Assessment of outcome
a) Pathological (histological or cytological) diagnosis (for at least 80% of all lung cancer cases) **(one star)**
b) No pathological diagnosis in more than 80% cases.
c) No description
d) Other

2) Was follow-up long enough for outcomes to occur
a) yes (>5 years) **(one star)**
b) no

3) Adequacy of follow up of cohorts
a) complete follow up - all subjects accounted for **(one star)**
b) subjects lost to follow up unlikely to introduce bias – number lost less than or equal to 20%, or description of those lost suggested no difference from those followed-up **(one star)**
c) Evidence of selective losses
d) Follow-up rate less than 80% and no description of those lost
e) No statement

Overall risk of bias for cohort studies:

Low risk of bias	4 or 5 stars in selection domain AND 1 or 2 stars in comparability domain AND 2 or 3 stars in outcome domain
Moderate risk of bias	2 or 3 stars in selection domain AND 1 or 2 stars in comparability domain AND 2 or 3 stars in outcome domain
High risk of bias	0 or 1 star in selection domain OR 0 stars in comparability domain OR 0 or 1 stars in outcome domain

Newcastle - Ottawa Quality Assessment Scale
Case Control Studies

Selection
1) <u>Is the case definition adequate?</u>
a) Yes, with pathological evidence **(one star)**
b) No pathological evidence
c) No description

2) <u>Representativeness of the cases</u>
a) consecutive or obviously representative series of cases **(one star)**
b) potential for selection biases or not stated

3) <u>Selection of Controls</u>
a) community controls **(one star)**
b) hospital controls
c) no description

4) <u>Definition of Controls</u>
a) no history of disease (lung cancer) **(one star)**
b) no description of source

Comparability
1) <u>Comparability of cases and controls on the basis of the design or analysis</u>
a) study controls age AND smoking **(two stars)**
b) study controls for age OR smoking **(one star)**

Exposure
1) <u>Ascertainment of exposure</u>
a) Linked record with NTP database with bacteriological confirmation (>80%) **(two stars)**
b) Linked record with NTP database without bacteriological confirmation **(one star)**
c) Structured interview where blind to case-control status **(one star)**
d) Interview not blinded or written self-report
e) No description

2) <u>Same method of ascertainment for cases and controls</u>
a) yes **(one star)**
b) no

3) <u>Non-Response rate</u>
a) Similar for both groups and total response rate >80% **(one star)**
b) Non-response selective to one group
c) Total response rate <80%
d) No description

Overall risk of bias for case-control studies:

Low risk of bias	3 or 4 stars in selection domain AND 1 or 2 stars in comparability domain AND 3 or 4 stars in exposure domain
Moderate risk of bias	2 stars in selection domain AND 1 or 2 stars in comparability domain AND 2 stars in exposure domain
High risk of bias	0 or 1 star in selection domain OR 0 stars in comparability domain OR 0 or 1 stars in exposure domain

Appendix 4. Rationale for changes to the Newcastle-Ottawa Scale		
Cohort studies		
Original scale	Adapted scale	Rationale for changes

	Selection		
1	Representativeness of the exposed cohort A. Truly representative **(one star)** B. Somewhat representative **(one star)** C. Selected group D. No description of the derivation of the cohort	No changes made	
2	Selection of the non-exposed cohort A. Drawn from the same community as the exposed cohort (**one star**) B. Drawn from a different source C. No description of the derivation of the non-exposed cohort	No changes made	
3	Ascertainment of exposure A. Secure record (e.g., surgical record) (**one star**) B. Structured interview (**one star**) C. Written self-report D. No description E. Other	Ascertainment of exposure A. Bacteriologically confirmed TB episode, from NTP/medical records (**two star**) B. Bacteriologically confirmed TB episode, from structured interview (**one star**) C. Clinically diagnosed TB episode from NTP/medical records (**one star**) D. Structured interview with no information on bacteriological diagnosis E. Self-report with no further information on the TB symptoms or	When the episode of TB is bacteriologically confirmed, we can be almost certain that it was active TB and not an early manifestation of lung cancer. This is more reliable if it has been ascertained from a NTP or medical record. An interview is less reliable to ascertain if a diagnosis was made bacteriologically Clinical or radiological TB diagnosis is less accurate since TB and lung cancer may share symptoms and radiological findings.

		diagnosis F. No description G. Other	
4	Demonstration that outcome of interest was not present at start of study A. Yes, no history of endpoint (**one star**) B. No	4) <u>Attempt to Demonstrate that outcome of interest was not present at start of study</u> A. Yes (excluded cases occurring in the first year after the tuberculosis diagnosis or performed latency analysis) **(one star)** B) No	Ascertain that a TB patient did not have lung cancer is very difficult to even using imaging. Therefore, we allow for a period one years between the TB diagnosis and the cancer diagnosis. If less, the cancer could have been present.
Comparability			
1	Comparability of cohorts on the basis of the design or analysis controlled for confounders A. The study controls for the most important factor (**one star**) B. Study controls for any additional important factor (list) (**one star**) C. Cohorts are not comparable on the basis of the design or analysis controlled for confounders	Comparability of cohorts on the basis of the design or analysis controlled for confounders A. The study controls for age AND smoking (**two star**) B. The study controls for age OR smoking (**one star**) C. The study controls for other factors only) D. Cohorts are not comparable on the basis of the design or analysis controlled for confounders	We considered a study should control for age and smoking for it to be pooled in the adjusted effects meta-analysis. These variables were chosen from a larger list of potential cofounders after considering epidemiological evidence (see "DAG and references") Smoking and age are the main (ref) risk factors for lung cancer. Studies controlling for both, have more comparable cohorts, than those controlling for age or for other factors only.
Outcome			
1	Assessment of outcome A. Independent blind assessment (**one star**) B. Record linkage (**one star**) C. Self-report D. No description	Assessment of outcome A. Pathological diagnosis (for at least 80% of all lung cancer diagnoses) (**one star**) B. No pathological diagnosis F. No description G. Other	Since TB and lung cancer may share clinical features, we consider it necessary that the diagnosis of lung cancer is made based on pathological evidence. Otherwise, a recurrence or sequel of TB may be misdiagnosed as lung cancer.

	E. Other		
2	Was follow-up long enough for outcomes to occur A. Yes (**one star**) B. No Indicate the median duration of follow-up and a brief rationale for the assessment above:_____	Was follow-up long enough for outcomes to occur A. Yes (>= 5 years on average) (**one star**) B. No	
3	Adequacy of follow-up of cohorts A. Complete follow-up- all subject accounted for (**one star**) B. Subjects lost to follow-up unlikely to introduce bias – number lost less than or equal to 20% or description of those lost suggested no different from those followed. (**one star**) C. Follow-up rate less than 80% and no description of those lost D. No statement	Adequacy of follow-up of cohorts A. Complete follow-up- all subject accounted for (**one star**) B. Subjects lost to follow-up unlikely to introduce bias - number lost less than or equal to 20%. (**one star**) C. losses are clearly selective to one group D. Follow-up rate less than 80% and no description of those lost E. No statement	A study where losses are relatively small but selective to one group may also introduce bias.
	Overall risk of bias		

Adapted Newcastle-Ottawa Scale for case-control studies

Original scale	Adapted scale	Rationale for changes
Good quality: 3 or 4 stars in selection domain AND 1 or 2 stars in comparability domain AND 2 or 3 stars in outcome domain	Low risk of bias: 4 or 5 stars in selection domain AND 1 or 2 stars in comparability domain AND 2 or 3 stars in outcome domain	Adapted to the changes in stars assigned to ascertainment of exposure (because this item can receive up to two stars instead of only one in the original scale).
Fair quality: 2 stars in selection domain AND 1 or 2 stars in comparability domain AND 2 or 3 stars in outcome domain	Moderate risk of bias: 2 or 3 stars in selection domain AND 1 or 2 stars in comparability domain AND 2 or 3 stars in outcome domain	We substituted the terms related to "quality" for "risk of bias" as suggested by current systematic review guidelines.
Poor quality: 0 or 1 star in selection domain OR 0 stars in comparability domain OR 0 or 1 stars in outcome domain	High risk of bias: 0 or 1 star in selection domain OR 0 stars in comparability domain OR 0 or 1 stars in outcome domain	

Selection

	Original scale	Adapted scale	Rationale for changes
1	Is the case definition adequate? A. Yes, with independent validation. **(one star)** B. Based on record linkage or based on self-reports C. No description	Is the case definition adequate? A. Yes, with pathological evidence **(one star)** B. Attempt to independently validate but not enough pathological evidence C. Based on record linkage D. Based on self-reports E. No description	Since TB and lung cancer may share clinical and radiological features, we consider it necessary that the diagnosis of lung cancer is made based on pathological evidence. Otherwise, a recurrence or sequel of TB may be misdiagnosed.

2	Representativeness of the cases A. Consecutive or obviously representative series of cases (**one star**) B. Potential for selection biases or not stated	No changes made	
3	Selection of Controls This item assesses whether the control series used in the study is derived from the same population as the cases and essentially would have been cases had the outcome been present. A. Community controls (**one star**) B. Hospital controls or other health service controls C. No description	No changes made	
4	Definition of Controls A. No history of disease (endpoint) (**one star**) B. No description of source	Definition of Controls A. No history of lung cancer (**one star**) B. No description of source	

Comparability			
1	Comparability of cases and controls on the basis of the design or analysis controlled for confounders A. The study controls for the most important factor (**one star**) B. Study controls for any additional important factor (list) (**one star**) C. Cases and controls are not comparable on the basis of the design or analysis controlled for confounders	Comparability of cases and controls on the basis of the design or analysis controlled for confounders A. The study controls* for age AND smoking (**two star**) B. The study controls* for age OR smoking (**one star**) C. Study controls* for other predefined factors (socioeconomic status, passive smoking, chronic bronchitis or emphysema) D. Cases and controls are not comparable on the basis of the design or analysis controlled for confounders	We considered a study should control for age, and smoking for it to be pooled in the adjusted effects meta-analysis. These variables were chosen from a larger list of potential cofounders after considering epidemiological evidence

		*if controls were matched to cases, matched analysis needs to be conducted, in order for the factors to be controlled.(not for frequency matching)	
Exposure			
1	Assessment of exposure A. Secure record (**one star**) B. Structured interview where blind to case/control status (**one star**) C. Interview not blinded D. Written self-report or medical record only E. No description	Assessment of exposure A. Linked record with NTP database with bacteriological confirmation (>80%) (**two star**) B. Linked record with NTP database without bacteriological confirmation (**one star**). **C.** Structured interview where blind to case/control status (**one star)** D. Interview not blinded or written self-report E. No description (bacteriological confirmation of exposure would be ideal, but unlikely to be complete for all)	When the episode of TB is bacteriologically confirmed, we can be almost certain that it was active TB and not an early manifestation of lung cancer misdiagnosed as TB. Diagnosis of TB based on clinical or radiological criteria is less accurate since TB and lung cancer may share symptoms and radiological findings. An interview is less reliable to ascertain if a diagnosis was made bacteriologically and it is also prone to recall bias
2	Same method of ascertainment for cases and controls A. Yes (**one star**) B. No	No changes made	

3	Non-response rate	Non-response rate (or not **possible** to link? Which is different to "not linked")	A study where overall non-response rate is relatively small but selective to either cases or controls may introduce bias.
	A. Same for both groups (**one star**)	A. Similar for both groups and total response rate >80% and description of non-respondents suggests no difference from respondents. (**one star**)	
	B. Non respondents described	B. Non-response selective to one group	
	C. Rate different and no designation	C. Total response rate <80%	
		C. No description	

Overall risk of bias

	Good quality: 3 or 4 stars in selection domain AND 1 or 2 stars in comparability domain AND 2 or 3 stars in exposure domain	Low risk of bias: 3 or 4 stars in selection domain AND 1 or 2 stars in comparability domain AND 3 or 4 stars in exposure domain	
	Fair quality: 2 stars in selection domain AND 1 or 2 stars in comparability domain AND 2 or 3 stars in exposure domain	Moderate risk of bias: 2 stars in selection domain AND 1 or 2 stars in comparability domain AND 2 stars in exposure domain	
	Poor quality: 0 or 1 star in selection domain OR 0 stars in comparability domain OR 0 or 1 stars in exposure domain	High risk of bias: 0 or 1 star in selection domain OR 0 stars in comparability domain OR 0 or 1 stars in exposure domain	

Appendix 5. Flow diagram of study selection into the meta-analysis models

Total number of studies included in the review (n=73)

Cohort studies (n=29)

Case-control studies (n=44)

Lung cancer diagnosis (n=19)[*]

Lung cancer mortality (n=12)[*]

Lung cancer diagnosis (n=43)

Lung cancer mortality (n=1)

not included in the meta-analysis (n=6)[†]

not included in the meta-analysis (n=5)[†]

not included in the meta-analysis (n=1)[†]

Studies included in at least one model of the meta-analysis (n=13)

Studies included in at least one model of the meta-analysis (n=7)

Studies included in at least one model of the meta-analysis (n=42)

Not enough studies for a meta-analysis

included only in model 1 (n=7)

included in models 1 and 2 (n=2)

included in models 2 and 3 (n=1)

included in all 3 models (n=3)

included only in model 1 (n=5)

included in models 1 and 2 (n=1)

included in all 3 models (n=1)

included only in model 1 (n=19)

included in models 1 and 2 (n=4)

included in models 2 and 3 (n=1)

included in all 3 models (n=18)

[*]Two studies reported both lung cancer diagnosis and mortality as their study outcomes. [†]Studies that reported an estimate adjusted for variables other than smoking and age (and did not report any unadjusted estimate eligible for model 1).

Appendix 6. Summary of the characteristic of the studies included in the systematic review

	Cohorts studies (n)	Case-control studies (n)
Lung cancer diagnosis	19 studies	43 studies
Study setting	Taiwan (8), South Korea (5), USA (2), China (1), Denmark (1), Finland (1), Lithuania (1)	China (16), USA (9), Taiwan (6), Canada (3), Singapore (2), Germany (2), South Korea (1), Nepal (1), Czech Republic (1), Italy (1), United Kingdom (1)
Publication date	1980-1999 (0), 2000-2009 (3), 2010-2021 (16)	1980-1999 (14), 2000-2009 (15), 2010-2021 (14)
Risk of bias	Low (7), moderate (10), high (2)	Low (5), moderate (15), high (23)
Main variables adjusted for[*]	Smoking (7), age (18), sex (18), any socioeconomic status indicator[†] (6), any comorbidity (10), passive smoking (0)	Smoking (32), age (25), sex (32), any socioeconomic status indicator* (2), any comorbidity (6), passive smoking (6)
Lung cancer mortality	12 studies	1 study
Study setting	China (5), USA (2), Denmark (2), Japan (1), South Korea (1), Italy (1)	China (1)
Publication date	1980-1999 (7), 2000-2009 (1), 2010-2021 (4)	1980-1999 (1), 2000-2009 (0), 2010-2021 (0)
Risk of bias	Low (4), moderate (1), high (7)	Low (0), moderate (0), high (1)
Main variables adjusted for[*]	Smoking (3), age (2), sex (8), any socioeconomic status indicator[†] (2), any comorbidity (6), passive smoking (1)	None

[*]The variables controlled for in each individual study as well as the number of times each variable was adjusted for by the included studies can be found in appendix 9 and 15. [†]Either income, education or occupation

Appendix 7. Characteristics of included studies

Table 1. Characteristic of included cohort studies that report lung cancer diagnosis as the outcome

Study	Study setting (location, country)	Study population	Number of participants	Ascertainment of TB / source	Comparator group	Ascertainment of lung cancer / source	Recruitment period	Follow-up duration	Factors adjusted for
An et al (2020)	South Korea	General population, A representative sample established by the National Health Insurance Service (NHIS)	22 656	Only record linkage / NHIS database	Five matched people without TB according to the same database	Only record linkage / NHIS database	2003-2013	Follow-up until 2013	Adjustment for smoking status (ever smoker, ex-smoker or current smoker), age, sex and household income
Bae et al (2013)	Seoul, South Korea	Representative sample of current male smokers	7 009	Interview / Seoul Male Cancer Cohort (SMCC)	Males without history of TB from the same cohort	Only record linkage / Seoul Regional Cancer Registry (SRCR), the Korea Central Cancer Registry (KCCR) and death certificates at Statistics Korea	1992-1993	99 965 person-years; follow-up until 2008	Adjustment for age, intake of tomatoes and coffee
Engels et al (2009)	Xuanwei, China	Farmers	42 422	Interview	Farmers without history of TB from the same community	Death records from hospitals, public security bureaus and public health bureaus	1976-1992	Follow-up until 1996	None
Everatt et al (2016)	Lithuania	General population	21 986	Record linkage / Lithuanian Tuberculosis Registry	Estimates from the general population	Record linkage (66.9% were microscopically confirmed) / Lithuanian Cancer Registry (LCR)	1998-2012	138 811.1 person years; 6.3 years (mean)	Standardization for age and sex

Hong et al (2016)	South Korea	General population, participants of the Korean Cancer Prevention Study (KCPS)	1 607 710	Chest x-ray or past hospitalization for TB / National Health Insurance Service (NHIS)	Participants without TB that participated in the same study	Two or more hospitalizations for lung cancer / NHIS	1997-2000	23 379 734 person-years; 14.5 years (mean)	Adjustment for smoking status (current smokers, exsmokers and never-smokers), amount of cigarettes per day (1-9, 10-19 and >=20), age, sex, socioeconomic status, alcohol consumption, hospitalizations for respiratory diseases
Huang et al (2015)	Taiwan	General population	15 219 024	Record linkage and more than two outpatient visits or one admission for TB / National Health Insurance Research Database (NHIRD)	People without history of TB from the same database	Record linkage with histological confirmation / NHIRD, Taiwan Cancer Registry Database (TCRD)	2001-2003	Follow-up until 2008	Adjustment for age, sex, low income, urbanization, geographical area, asthma, COPD, diabetes, hyperlipidaemia, chronic kidney disease, smoking-related cancers
Jian et al (2016)	Taiwan	Asthma patients	54 520	Record linkage and more than two outpatient visits or one admission for TB / NHIRD	Asthma patients without history of TB	Record linkage with histological confirmation / NHIRD, Taiwan Cancer Registry Database (TCRD)	2001-2005	Follow-up until 2010	Adjustment for age, sex, urbanization, COPD, pneumonia, diabetes, hyperlipidaemia, liver cirrhosis, chronic kidney disease, autoimmune disease, atopy dermatitis, rhinosinusitis, inhaled corticosteroids use, smoking-related cancers
Kuo et al (2013)	Taiwan	General population	6 699	Record linkage including prescription of at least two ant tuberculosis drugs for 2 months / NHIRD	Estimates from the general population	Record linkage – with histological confirmation / NHIRD, Catastrophic Illness Taiwan Database	2000-2010	28 866 person-years; 3.8 years (median)	Standardization for age and sex

Lai et al (2012)	Taiwan	Diabetes Mellitus patients and matched controls (aim of the study was to study diabetes as a risk factor for lung cancer)	98 120	Only record linkage / NHIRD	People without history of TB from the same sample	Only record linkage / NHIRD	1995-2005	442 237 and 108 214 person-years for the DM and non-DM cohort respectively; follow-up until 2008	Adjustment for age, sex, COPD, diabetes mellitus
Littman et al (2004)	USA	Heavy smokers and asbestos-exposed workers that participated in a cancer prevention trial (CARET trial).	17 698	Interview	Subjects without history of TB from the same trial	Review of clinical records and pathology reports from the diagnosing physician or hospital to confirm the tumor origin, location, and histology	1985-1993	9.1 years (median); follow-up until 2002	Adjustment for years smoked and years smoked squared, average number of cigarettes smoked per day and average number of cigarettes smoked per day squared, smoking status (former or current), age, sex, body mass index, trial intervention, asbestosis, asthma, chronic bronchitis or emphysema, pneumonia
Liu et al (2017)	Taiwan	Female COPD patients	13 686	Only record linkage / NHIRD	Female COPD patients without history of TB from the same database	Only record linkage / NHIRD	1997-2011	9.78 years (median); follow-up until 2011	Adjustment for age, income, pneumonia, bronchiectasis, pulmonary fibrosis, hypertension, diabetes mellitus, inhaled corticosteroids use

Oh et al (2020)	South Korea	General population older than 40 years that participated in a nationwide survey (KNHANES study)	20 252	Interview / conducted as part of the survey	People without history of TB from the same survey	Record linkage with pathological confirmation / Korea Central Cancer Registry	2008-2013	3.85 years (mean) for the TB group, 4 years (mean) for the control group; follow-up until 2014	Adjustment for smoking status (current smokers, ex-smokers or never-smokers), age, sex, income level, education, body mass index, physical activity
Shebl et al (2010)	USA	AIDS patients	322 675	History of TB / HIV/AIDS registries	AIDS patients without TB from the same registry	Only record linkage / cancer registries in 11 US regions	1977-2002	1 032 256 person-years; 10 years (not clear if mean or median)	Adjustment for age, sex, race, mode of HIV acquisition, CD4 count at AIDS onset, calendar year of AIDS onset
Shiels et al (2011)	Southwestern regions of Finland	Male smokers, aged 50 to 69 years old, that participated in a cancer prevention trial (ATBC trial)	29 133	Only record linkage / available from the National Hospital Discharge Register	Participants without history of TB from the same trial	Record linkage with histology known for 62% cases / Finnish Cancer Registry	1985-1988	Follow-up until 2005	Adjustment for smoking measured with log cig-years (log [cigarettes smoked per day + 1] x number of years smoked) and age
Simonsen et al (2014)	Denmark	General population	15 024	Record linkage (58% cultured-confirmed) / Danish National Registry of Patients (DNRP)	Estimates from the general population	Record linkage with 89% cases verified morphologically / Danish Cancer Registry, Danish Pathology Register	1978-2011	150 400 person-years; 8.5 years (median)	Standardization for age and sex
Wu et al (2011)	Taiwan	General population	29 641	Record linkage and prescriptions of at least 2 anti-tuberculosis medications for >28 day	Four matched control subjects with no TB record matched to each TB patient from the same database	Record linkage – with histological confirmation / NHIRD, Catastrophic Illness Taiwan Database	1997-2008	5.86 years (mean) for TB patients, 6.22 years (mean) for controls	Adjustment for age, sex, COPD, diabetes mellitus, chronic renal failure, autoimmune disease

Wu et al (2016)	Taiwan	COPD patients	44 065	Record linkage and either 2 outpatients visits or one admission for TB / NHIRD	COPD patients without history of TB from the same database	Record linkage – with histological confirmation / NHIRD, Taiwan Cancer Registry	2001-2005	$(4.2 + 17.4) \times 10^5$ person-months; follow-up until 2010	Adjustment for age, sex, urbanization, number of visits for respiratory diseases within 2 years after index date, pneumonia, chronic kidney disease, diabetes mellitus, hyperlipidaemia, liver cirrhosis, autoimmune disease, atopy dermatitis, rhinosinusitis, inhaled corticosteroids use, oral corticosteroids use, bronchodilators use, statins and aspirin use
Yeo et al (2021)	South Korea	Random sample from the general population that participated in health examinations	1 875 846	Record linkage / Korean National Health Insurance (KNHI) database	People without history of TB from the same sample	Record linkage – with histological confirmation /KNHI database	2009	15 341 796 person years; 8.2 years (mean)	Adjustment for smoking (pack-years), age, sex, BMI, COPD, diabetes mellitus, alcohol consumption, insurance coverage
Yu et al (2011)	Taiwan	General population	716 872	Only record linkage / NHIRD	People without history of TB from the same database	Record linkage / NHIRD	1998-2000	37 951 person-years for the TB group and 6 571 088 person-years for the control group; follow-up until 2007	Adjustment for age, sex, occupation, COPD, diabetes mellitus, hypertension, dyslipidaemia, smoking-related cancers

Table 2. Characteristic of included case-control studies that report lung cancer diagnosis as the outcome

Study	Study setting (location, country)	Number of participants	Case description	Control description	Type of controls	Ascertainment of lung cancer	Ascertainment of TB /source	Recruitment period	Factors adjusted for
Alavanja et al (1992)	Missouri, USA	2 015	Nonsmoking women from Missouri Cancer Registry	Nonsmoking women randomly sampled from driver's license files and the HCFA	Community controls	76% histologically confirmed, others were cytologically confirmed (percentage no available)	Structured interview	1986-1991	Smoking history (lifetime nonsmoker or former smoker), age
Bodmer et al (2012)	United Kingdom	91 301	Subjects from the General Practice Research Database	Randomly sampled subjects without lung cancer from the General Practice Research Database	Community controls	Record linkage	Record linkage	1995-2009	Smoking status (non-smoker, current, past or unknown), age, sex, COPD, BMI, congestive heart failure, ischemic heart disease, stroke/transient ischemic attack, hypertension, dyslipidaemia, diabetes mellitus
Brenner et al (2001)	Pingliang and Qingyang, China	2 650	Subjects from 2 prefecture hospitals, a company hospital,15 county hospitals and local clinics	Randomly sampled subjects from a population census list	Community controls	60% clinical-radiological diagnosis, 40% pathologically confirmed	Structured interview	1994-1998	Smoking category (heavy, moderate, light o never-smokers), age, sex, prefecture
Brenner et al (2010)	Toronto, Canada	1 393	Subjects from 4 major tertiary care hospitals in metropolitan Toronto	Subjects without any cancer and randomly sampled from property tax assessment files (45%) and the Mount Sinai Hospital Family Medicine Clinic (55%)	Community and hospital controls	100% histology confirmed	Interview	1997-2002	Smoking (pack-years), age, sex, education, ethnicity
Brownson et al (2000)	Missouri, USA	1 376	Women from the Missouri Cancer Registry	Randomly sampled subjects from state driver's license files and from the HCFA	Community controls	74% histologically confirmed, others were cytologically confirmed (percentage no available)	Interview	1993-1994	Smoking (pack-years)
Chan-yeung et al (2003)	Hong Kong, China	661	Subjects from the Queen Mary Hospital	Subjects without lung cancer from the Queen Mary Hospital	Hospital or other health service controls	100% pathologically confirmed	Interview	1999-2001	Smoking duration and amount of cigarettes smoked (<20, 20-39, >40 pack-years), sex
Cheng et al (2012)	Taiwan	1 485	Women from the NHRI	Women from the NHRI hospitalized for orthopedic conditions, trauma, and other health conditions	Hospital or other health service controls	Record linkage	Record linkage	2005-2008	None

Chen et al (2021)	Xinjian, China	16 884	Subjects from a Cancer hospital	Subjects treated for benign tumors	Hospital controls	100% histologically confirmed	Medical records	2016-2018	Age and sex
Galeone et al (2008)	Harbin, China	651	Hospitalized subjects from the department of cardiothoracic surgery of the hospitals	Subjects without lung cancer from the cardiothoracic, urological and general surgery departments of the same hospitals as cases	Hospital or other health service controls	100% histologically confirmed	Structured interview	1987-1990	Smoking status (never, current or exsmokers), smoking duration (for current smokers: <25, 25-35 and >35 years; for exsmokers: <5 and > 5 years from the last cigarette) and amount of smoking (for current cmokers: <10, 10-15 and >=15 cigarettes per day; for exsmokers: <15 and >=15 cigarettes per day), income, family history of lung and other cancers, occupational exposure to lung carcinogens
Hinds et al (1982)	Hawaii, USA	629	Women from Hawaii Tumor Registry	Women from a representative sample of 38 000 adults in Hawaii	Community controls	No information (tumor registry)	Medical records	1968-1978	None
Hosgood lll et al (2013)	Xuanwei, China	996	Subjects from 4 hospitals in Xuanwei	Randomly sampled subjects from the general population	Community controls	61% clinical-radiological diagnosis, 39% pathologically confirmed	Interview	1985-1990	Smoking (never users; sole users of other types of tobacco or cigarettes smoked with a water pipe, ≤20 pack-years of cigarettes smoked without a water pipe; >20 pack-years of cigarettes smoked without a water pipe), sex, educational status, passive smoking, fuel type, family history of lung cancer
Ko et al (1997)	Kaohsiung, Taiwan	210	Women from Kaohsiung Medical College Hospital	Women with non-smoking related disease from a health check or ophthalmic department in the Kaohsiung Medical College Hospital	Hospital or other health service controls	100% pathologically confirmed	Structured interview	1992-1993	Socioeconomic status, education residential area, industrial district, cooking fuels, fume extractor, vegetable consumption
Koshiol et al (2010)	Lombardia, Italy	3 883	Subjects from 13 hospitales in Lombardia	Randomly sampled subjects from the Regional Health Service database	Community controls	95% pathologically confirmed, 5% clinical-radiological diagnosis	Structured interview	2002-2005	Smoking (pack-years and average packs/day), age, sex, region
Kreuzer et al (2001)	Germany	857	Men who were never-smokers from 15 study clinics in defined regions of East and West Germany	Men who were never-smokers and randomly sampled from mandatory registries or by a modified random-digit dialing from the same regions as cases	Community controls	100% pathologically confirmed	Structured interview	1990-1996	Age, region

Kreuzer et al (2002)	Germany	762	Women who were never-smokers from 15 study clinics in defined regions of East and West Germany	Women who were never-smokers and randomly sampled from mandatory registries or by a modified random-digit dialing from the same regions as cases	Community controls	100% pathologically confirmed	Structured interview	1991-1996	Age, region
Lai et al (2013)	Taiwan	14 110	Subjects from NHI	Subjects without lung cancer from NIH	Community controls	Record linkage	Record linkage	2000-2009	Smoking (ICD-9 codes, NIH is not reliable for this variable), age, sex, parkinson's disease, COPD, pneumoconiosis, asbestosis, alcoholism
Lai et al (2013)	Taiwan	11 450	Men from NHI	Randomly sampled men from NIH	Community controls	Record linkage	Record linkage	2000-2010	Smoking (ICD-9 codes, NIH is not reliable for this variable, COPD, asbestosis
Lee et al (2001)	Kaohsiung, Taiwan	473	Hospitalized subjects in the chest or oncology from Kaohsiung Medical University Hospital	Hospitalized subjects with conditions unrelated to tobacco use from Kaohsiung Medical University Hospital	Hospital or other health service controls	100% histologically confirmed	Structured interview	1993-1999	Smoking (cumulative pack-years), sex, socioeconomic status, education, residential area
Liang et al (2009)	Shenyang, China	505	Women who were never-smokers from 18 hospitals in Shenyang.	Women who were never-smokers and randomly sampled from the general population using the Residential Registry in Shenyang	Community controls	68% histologically confirmed, others were cytologically confirmed (percentage no available)	Structured interview	2004-2007	Age, passive smoking, years of schooling, marital status- - ethnicity, 5 years ago BMI, coal use, exposure to coal smoke and cooking fumes
Lim et al (2011)	Singapore	1 808	Women from the five major public sector hospitals in Singapore	Hospitalized women for conditions other than cancer at the same hospital as cases	Hospital or other health service controls	96% pathologically confirmed	Structured interview	1996-1998 and 2005-2008	Age, passive smoking, number of years in school, family history of cancer, fruit and vegetable consumption, country of origin, dialect group, housing type
Liu et al (1993)	Guangzhou, China	632	Subjects from 8 major hospitals in Guangzhou	Inpatients of the surgery departments at 6 of the same hospitals as cases	Hospital or other health service controls	32% pathologically confirmed	Structured interview	1983-1984	Smoking (not clear how, but they measured cigarettes smoked per day), sex, education, occupation, living area
Lo et al (2013)	Taiwan	3 055	Never-smokers from 6 tertiary medical centres in Taiwan	Never-smokers without lung cancer and randomly selected from the health-examination departments of the same six hospitals as cases	Hospital or other health service controls	100% histologically confirmed	Structured interview	2002-2009	Age, sex, years of education
Luo et al (1996)	Fuzhou, China	408	Subjects from a special reporting system designed to cover all lung cancers in hospitals in urban	Subjects randomly sampled of the general population of urban Fuzhou	Community controls	100% histologically confirmed	Structured interview	1990-1991	None

				Fuzhou					
Mayne et al (1999)	New York, USA	868	Nonsmoking subjects from a special system established to rapid ascertainment of lung cancer in New York	Nonsmoking subjects randomly sampled from the New York State Department of Motor Vehicles' file of licensed drivers	Community controls	99% histologically confirmed	Structured interview	1982-1984	None
Osann et al (2000)	California, USA	302	Women with small cell carcinoma from 28 hospitals in Orange County and neighbouring areas	Women identified through a random digit dialling in the same region as cases	Community controls	100% pathologically confirmed	Interview	1990-1993	Smoking (pack-years), years since quitting smoking, age, education
Park et al (2010)	South Korea	2 615	Subjects from 50 Korean general hospitals recruited in the nationwide KATRD study	Subjects from Chungju in the KMCC, a prospective cohort, who were voluntary participants in cancer screening surveys	Hospital or other health service controls	100% histologically confirmed	Interview	1996-2004	Smoking status (ever or never smoker), age, sex
Ramanaku mar et al (2006) – study I	Montreal, Canada	2 746	Subjects from 18 large hospitals in Metropolitan Montreal	Randomly sampled subjects from population based electoral lists in Metropolitan Montreal	Community controls *	100% histologically confirmed	Structured interview	1995-2001	Smoking status (ever or never smoker), log of cigarettes-year, number of years since quitting smoking (0-2, 2-5, 5-10, 10-15 or >15 years), age, family income, year of schooling, ethnicity, type of respondent
Ramanaku mar et al (2006) – study II	Montreal, Canada	1 287	Men from 18 large hospitals in Metropolitan Montreal	Randomly sampled men from population based electoral lists in Metropolitan Montreal	Community controls *	100% histologically confirmed	Structured interview	1979-1986	Smoking status (ever or never smoker), log of cigarettes- year, number of years since quitting smoking (0-2, 2-5, 5-10, 10-15 or >15 years), age, sex, family income, years of school attendance, ethnicity, type of respondent
Raspanti et al (2016)	Chitwan, Nepal	1 212	Subjects from Koirala Memorial Cancer Hospital	Visitors of non-lung cancer patients from Koirala Memorial Cancer Hospital	Hospital or other health service controls	92% histologically confirmed for a group of 209 cases, no data for the other 397 cases	Structured interview	2009-2012	Smoking status (ever or never smoker; they also calculated pack-years of smoking but it is not clear if this was included in the model), age, sex, socioeconomic status, passive smoking, household air pollution exposure
Samet et al (1986)	New Mexico, USA	1 287	Subjects from New Mexico Tumor Registry	Subjects randomly sampled from a list of residential telephone numbers and the New Mexico Medicare Financing Administration	Community controls	No information (tumor registry)	Interview	1980-1982	Smoking (duration of smoking in years, number of cigarettes smoked per day on average, duration of cessation in years, and a product term of smoking duration with an indication

									variable for age above and age below 65), age, sex, ethnicity
Schwartz et al (1996)	Detroit, USA	534	Subjects who were never-smokers from the OCISS	Subjects without any cancer who were never-smokers and were sampled by random-digit dialling in the OCISS	Community controls	86% histologically confirmed	Interview	1984-1987	Age, gender, race
Seow et al (2002)	Singapore	1 066	Women from 3 major hospitals in Singapore	Women without lung cancer from the same hospitals as cases	Hospital or other health service controls	100% pathologically confirmed	Structured interview	1996-1998	None
Wang et al (1996) a	Guangzhou, China	780	Inpatients from 5 hospitals in Guangzhou	Inpatients without any cancer from the same hospitals as cases	Hospital or other health service controls	100% pathologically confirmed	Interview	1990-1993	Smoking status (no more details), passive smoking, chronic bronchitis/emphysema, family history of tumors, consumption of pickled and cured foods
Wang et al (1996) b	Shenyang, China	270	Women who were never-smokers from 18 hospitals in Shenyang	Women who were never-smokers and randomly sampled from the general population in urban areas of Shenyang	Community controls	57% pathologically confirmed, 43% clinical-radiological diagnosis	Structured interview	1992-1994	None
Wang et al (2009)	Hong Kong, China	504	Women who were never-smokers from the largest oncology centre in Hong Kong	Women who were never-smokers and randomly sampled using the residential telephone in Hong Kong	Community controls	100% pathologically confirmed	Structured interview	2002-2004	Age, employment, total dish year, intake of yellow/orange vegetables, dark green vegetables, multivitamins
Wang et al (2014)	Changchun, China	1 000	Subjects from a hospital (not specified) in Changchun	Randomly selected subjects with routine physical examinations in the same hospital as cases	Hospital or other health service controls	100% histologically confirmed	Structured interview	2010-2012	Smoking (pack-years), COPD, family history of cancer
Wu et al (1988)	California, USA	672	Women with adenocarcinoma from the Cancer Surveillance Program, a population-based tumour registry, in Los Angeles County	Women selected from each case's neighbourhood in Los Angeles County	Community controls	100% histologically confirmed	Structured interview	1983-1986	Smoking (pack-years), years since smoking stopped, depth of inhalation, social class according to father occupation (blues or white collar worker)
Wu et al (1995)	USA	1 633	Women who were never-smokers from 5 major metropolitan areas in USA	Women who were never-smokers and randomly selected through digit dialing and from the HCFA from the same areas as cases	Community controls	Microscopically confirmed diagnosis	Interview	1985-1990	Age-area-ethnicity-education-passive smoking
Wu-	Harbin	1 924	Women from cancer	Randomly sample	Community	42% histologically	Structured	1985-1987	Smoking (non-smoker, smoked

Study	Setting	N	Cases source	Controls source	Control type	Confirmation	Data collection	Period	Adjustment variables
Williams et al (1990)	and Shenyang, China		registries of Harbin and Shenyang	women of the general population in the same location as cases	controls	confirmed, 32% cytology confirmed, 26% clinical-radiological diagnosis	interview		1-9 cigarettes per day and 1-29 years, 2-19 and 30-39 years, 1-19 and >40 years, >20 cigarettes per day and 1-29 years >20 and 30-39 years, >20 and >40 years), age, education, study area
Yang et al (2015)	Guangzhou and Suzhou, China	3 238	Subjects from urban hospitals and one suburb hospital in Guangzhou and Suzhou	Subjects without any cancer from healthy check-up programs in the community health stations in the same city as cases	Hospital or other health service controls	100% histologically confirmed	Structured interview	2007-2010	Smoking status (ever or never smoker), pack-years of smoking (low, 0-5; moderate, 6-20; or high, >20), age, sex, passive smoking, emphysema, education, BMI, educational experience, centre, packs-years occupational exposure to metallic toxicant, housing ventilation, biomass burning, cured meat, vegetables/fruits
Zatloukal et al (2003)	Czech Republic	1 990	Women admitted to Prague University Hospital Na Bulovce	Women who were spouses, relatives, or friends of other patients hospitalized at Prague University Hospital Na Bulovce	Hospital or other health service controls	100% histologically confirmed	Structured interview	1998-2002	Smoking (pack-years), age, education, residence, residence
Zheng et al (1987)	Shanghai, China	2 863	Men from Shanghai Cancer Registry and a specially established lung cancer rapid-reporting system operated by the Shanghai Cancer Institute	Randomly sampled subjects from the general population in Shanghai urban area	Community controls	62% histologically confirmed, 30% cytology confirmed, 7% clinical-radiological diagnosis	Structured interview	1984-1986	Smoking category (non-smoker, light, moderate or heavy smoker), age, sex, education
Zhou et al (2000)	Shenyang, China	144	Women with adenocarcinoma from 18 major hospitals in Shenyang	Randomly selected women from the general population in various areas of Shenyang	Community controls	100% histologically confirmed	Structured interview	1991-1995	None

When the setting is not specified below the country level, the participants were selected nationwide or from a large area of the corresponding country. Non-smokers include never and former smokers. Some studies report a nonsmoking definition that matches corresponding to a never-smokers definition, those studies were considered to only as never-smokers due to uniformity criteria. Pathologically confirmed lung cancer includes histological AND / OR cytological confirmation. HCFA: Health Care Financing Administration. NHRI: National Health Research Institutes from Taiwan. NHI: National Health Insurance from Taiwan. KARTD: Korean academy of tuberculosis and respiratory disease. KMCC: Korean Multi-Center Cancer Cohort. OCISS: Occupational Cancer Incidence Surveillance Study. COPD: chronic obstructive pulmonary disease. BMI: body mass index

Table 3. Characteristic of included cohort studies that report lung cancer mortality as the outcome

Study	Study setting (location, country)	Study population	Number of participants	Ascertainment of TB / source	Comparator group	Ascertainment of lung cancer death / source	Recruitment period	Follow-up duration	Factors adjusted for
Boice et al (1980)	Massachusetts, USA	General population, only females, treated for TB before availability of isoniazid	1090	Medical records / 2 Massachusetts hospitals	Estimates from females in the general population	Death certificates / hospitals	1930-1954	23 094 person-years; 21.2 years (mean); follow-up until 1975	Standardizes for age
Chen et al (1990)	Hebein province, China	Males mine workers with silicosis	5406	Interview	Workers without history of TB from the same mine	Death records / pension department	1970-1982	6102.7 person-years for the TB group and 61633.7 person-years for the non-TB group	None
Christensen et al (2014)	Denmark	General population	25608	Record linkage with 82.6 % microbiological confirmation / Danish Tuberculosis Registry, Danish National Patient Registry (DNPR)	Matched people randomly sampled from the general population	Death certificates / Danish Registry of Causes of Death	1977-2008	TB group: 64 212 person-years; 8.1 years (median) Control group: 234 484 person-years; 10.5 years (median) Total population: 298 696 person-years	Matched on age and sex
Davis et al (1989)	Massachusetts, USA	General population	13 385	Medical records / 12 hospitals	Estimates from the general population	Death certificates / hospitals	1925-1954	25 years (mean); follow-up until 1986	Standardizes for age and sex

Engels et al (2009)	Xuanwei, China	Farmers	42 422	Interview	Farmers without history of TB from the same community	Death records / hospitals, public security bureaus and public health bureaus	1976-1992	Follow-up until 1996	They adjust one variable at a time: smoking per day (cigarettes per day), age, sex, education level, use of smoky coal, asthma, chronic bronchitis, emphysema, family history of tuberculosis, walking hours spent indoors, number of rooms in house)
Floe et al (2018)	Denmark	General population	42 140	Only record linkage / National Patient Registry	Matched controls randomly sampled from the general population	No description	1998-2010	Follow-up until 2010	None
Gao et al (1992)	Shangai, China	General population	30 373	There is only and exposed group	Estimates from the general population	No description	1972-1986	Follow-up until 1986	Standardizes for age and sex
Hong et al (2016)	South Korea	General population, participants of the Korean Cancer Prevention Study (KCPS)	1 607 710	Chest x-ray or past hospitalization for TB / National Health Insurance Service (NHIS)	Participants without TB that participated in the same study	Death certificates / National Statistical Office in Korea	1997-2000	23 379 734 person-years; 14.5 years (mean)	Adjusts for smoking status (current smokers, exsmokers and never-smokers), amount of cigarettes per day (1-9, 10-19 and >=20), age, sex, alcohol consumption-socioeconomic status-diabetes mellitus and respiratory diseases hospitalizations
Leung et al (2013)	Hong Kong, China	Clients enrolled at the 18 health centers for the elderly	61 239	Medical records (49.2% bacteriologically confirmed – 50.8% clinical, radiological and/or histological plus appropriate response to anti-TB	Elderly without history of TB from the same centers	Death registry / statistical section of the Department of Health	2000-2011	490 258 person-years	Adjusts for smoking status, age, sex, passive smoking, language, education level, marital status, housing situation, public means, tested financial assistance status, alcohol use, body mass index, COPD, asthma and

				treatment) / territory-wide TB notification registry					family history of malignancy
Merlo et al (1995)	Genoa, Italy	People with silicosis	515	No description	People with silicosis without history of TB	No description	1961-1980	11.56 years (mean); follow-up until 1981	None
Ng et al (1990)	Hong Kong, China	Men with silicosis	1 419	Interview	Those without history of TB from the same sample	Death records / Registry of Persons, Registry of Deaths	1980	7 429.8 person-years; follow-up until 1986	None
Sasaki et al (1992)	Nagoya, Japan	General population	3 580	Medical records / Nagoya TB Registry	Estimates from the general population	Death certificates and/or medical records	1979-1981	12 702 person-years; follow-up until 1983	Standardizes for age and sex

Table 4. Characteristic of included case-control studies that report lung cancer mortality as the outcome

Study	Study setting (location, country)	Number of participants	Case description	Control description	Type of controls	Ascertainment of lung cancer mortality	Ascertainment of TB /source	Recruitment period	Factors adjusted for
Fu et al (1984)	Harbin, China	1 046	Lung cancer deaths from medical certificates, held by the police substation in each of 3 districts in Harbin, and data from death reports held by each district's Sanitation and Antiepidemic Station	Non-respiratory deaths in the same district of residence held by the same source as cases	Community controls	Death certificates	Interview with relatives of the dead	1977-1979	None

Appendix 8. Assessment of Risk of Bias in included studies

	Selection				Comparability	Outcome			Overall risk of bias
	Representativeness of the exposed cohort (maximum 1 star)	Selection of the non-exposed cohort (maximum 1 star)	Ascertainment of exposure (maximum 2 stars)	Dealing with reverse causation bias (maximum 1 star)	Comparability of the cohorts on the basis of the design or analysis (maximum 2 stars)	Assessment of outcome (maximum 1 star)	Was follow-up long enough for outcomes to occur (maximum 1 star)	Adequacy of follow-up (maximum 1 star)	
An et al (2020)	1 star	1 star	1 star	1 star	2 stars	none	1 star	1 star	LOW
Bae et al (2013)	1 star	1 star	none	none	1 star	none	1 star	1 star	MODERATE
Engels et al (2009)	1 star	1 star	none	1 star	none	none	1 star	none	HIGH
Everatt et al (2016)	1 star	1 star	2 stars	1 star	2 stars	none	1 star	1 star	LOW
Hong et al (2016)	1 star	1 star	1 star	none	2 stars	none	1 star	1 star	MODERATE
Huang et al (2015)	1 star	1 star	1 star	none	1 star	1 star	1 star	1 star	MODERATE
Jian et al (2016)	none	1 star	1 star	none	1 star	1 star	1 star	1 star	MODERATE
Kuo et al (2013)	1 star	1 star	1 star	1 star	1 star	1 star	1 star	1 star	LOW
Lai et al (2012)	none	1 star	1 star	none	1 star	none	1 star	1 star	MODERATE
Littman et al (2004)	none	1 star	none	none	2 stars	1 star	1 star	1 star	HIGH
Liu et al (2017)	none	1 star	1 star	none	1 star	none	1 star	1 star	MODERATE
Oh et al (2020)	1 star	1 star	none	none	2 stars	1 star	none	1 star	MODERATE
Shebl et al (2010)	none	1 star	1 star	1 star	1 star	none	1 star	1 star	MODERATE
Shiels et al (2011)	1 star	1 star	1 star	1 star	2 stars	none	1 star	1 star	LOW
Simonsen et al (2014)	1 star	1 star	1 star	1 star	1 star	1 star	1 star	1 star	LOW
Wu et al (2011)	1 star	1 star	1 star	1 star	1 star	1 star	1 star	1 star	LOW
Wu et al (2016)	none	1 star	1 star	none	1 star	1 star	1 star	1 star	MODERATE
Yeo et al (2021)	1 star	1 star	1 star	1 star	2 stars	1 star	1 star	1 star	LOW
Yu et al (2011)	1 star	1 star	1 star	none	1 star	none	1 star	1 star	MODERAT

Table 1. Risk of bias in included cohort studies for the lung cancer diagnosis outcome

Table 2. Risk of bias in included case-control studies for the lung cancer diagnosis outcome

	Selection				Comparability	Exposure			Overall risk of bias
	Is the case definition adequate? (maximum 1 star)	Representativeness of the cases (maximum 1 star)	Selection of controls (maximum 1 star)	Definition of controls (maximum 1 star)	Comparability of cases and controls on the basis of the design or analysis (maximum 2 stars)	Assessment of exposure (maximum 2 stars)	Same method of ascertainment for cases and controls (maximum 1 star)	Non-response rate (maximum 1 star)	Overall risk of bias
Alavanja et al (1992)	1 star	1 star	1 star	1 star	2 stars	none	1 star	none	HIGH
Bodmer et al (2012)	none	1 star	1 star	1 star	2 stars	1 star	1 star	1 star	LOW
Brenner et al (2001)	none	1 star	1 star	none	2 stars	none	1 star	1 star	MODERATE
Brenner et al (2010)	1 star	none	none	1 star	2 stars	none	1 star	none	HIGH
Brownson et al (2000)	1 star	1 star	1 star	1 star	1 star	none	1 star	1 star	MODERATE
Chan-yeung et al (2003)	1 star	none	none	1 star	1 star	none	1 star	none	HIGH
Cheng et al (2012)	none	1 star	none	1 star	none	1 star	1 star	1 star	HIGH
Galeone et al (2008)	1 star	none	none	none	1 star	none	1 star	none	HIGH
Chen et al (2021)	1 star	1 star	none	1 star	1 star	1 star	1 star	1 star	LOW
Hinds et al (1982)	1 star	1 star	1 star	1 star	1 star	1 star	1 star	1 star	LOW
Hosgood lll et al (2013)	none	1 star	1 star	none	1 star	none	1 star	1 star	MODERATE
Ko et al (1997)	1 star	1 star	none	1 star	1 star	1 star	1 star	1 star	LOW
Koshiol et al (2010)	1 star	1 star	1 star	none	2 stars	1 star	1 star	1 star	LOW
Kreuzer et al (2001)	1 star	none	1 star	none	2 stars	none	1 star	1 star	MODERATE
Kreuzer et al (2002)	1 star	none	1 star	none	2 stars	none	1 star	1 star	MODERATE
Lai et al (2013)	none	none	1 star	1 star	1 star	1 star	1 star	1 star	MODERATE
Lai et al (2013)	none	none	1 star	1 star	none	1 star	1 star	1 star	HIGH
Lee et al (2001)	1 star	1 star	none	1 star	1 star	none	1 star	1 star	MODERATE
Liang et al (2009)	none	1 star	1 star	none	2 stars	none	1 star	1 star	MODERATE

Study									
Lim et al (2011)	1 star	1 star	none	1 star	2 stars	none	1 star	1 star	MODERATE
Liu et al (1993)	none	1 star	none	none	1 star	none	1 star	1 star	HIGH
Lo et al (2013)	1 star	1 star	none	1 star	2 stars	none	1 star	1 star	MODERATE
Luo et al (1996)	1 star	1 star	1 star	1 star	none	none	1 star	none	HIGH
Mayne et al (1999)	1 star	none	1 star	none	none	none	1 star	none	HIGH
Osann et al (200)	1 star	none	1 star	none	2 stars	none	1 star	none	HIGH
Park et al (2010)	1 star	none	none	1 star	2 stars	none	none	1 star	HIGH
Ramanakumar et al (2006) – study l	1 star	1 star	1 star	none	2 stars	none	1 star	none	HIGH
Ramanakumar et al (2006) – study ll	1 star	1 star	1 star	none	2 stars	none	1 star	none	HIGH
Raspanti et al (2016)	1 star	1 star	none	none	2 stars	none	1 star	none	HIGH
Samet et al (1986)	none	1 star	1 star	1 star	2 stars	none	1 star	1 star	MODERATE
Schwartz et al (1996)	1 star	1 star	1 star	1 star	2 stars	none	1 star	none	HIGH
Seow et al (2002)	1 star	1 star	none	1 star	1 star	none	1 star	1 star	MODERATE
Wang et al (1996) a	1 star	none	none	1 star	1 star	none	1 star	none	HIGH
Wang et al (1996) b	none	1 star	1 star	none	1 star	none	1 star	none	HIGH
Wang et al (2009)	1 star	1 star	1 star	1 star	2 stars	none	1 star	none	HIGH
Wang et al (2014)	1 star	none	none	1 star	1 star	none	1 star	none	HIGH
Wu et al (1988)	1 star	none	1 star	none	1 star	none	1 star	none	HIGH
Wu et al (1995)	1 star	1 star	1 star	none	2 stars	none	1 star	none	HIGH
Wu-Williams et al (1990)	none	1 star	1 star	1 star	2 stars	none	1 star	none	HIGH
Yang et al (2015)	1 star	1 star	none	1 star	2 stars	none	1 star	1 star	MODERATE
Zatloukal et al (2003)	1 star	1 star	none	none	2 stars	none	1 star	1 star	MODERATE
Zheng et al (1987)	1 star	1 star	1 star	1 star	2 stars	none	1 star	1 star	MODERATE
Zhou et al (2000)	1 star	none	1 star	none	none	none	1 star	none	HIGH

Table 3. Risk of bias in included cohort studies for the lung cancer mortality outcome

	Selection				Comparability	Outcome			Overall risk of bias
	Representativeness of the exposed cohort (maximum 1 star)	Selection of the non-exposed cohort (maximum 1 star)	Ascertainment of exposure (maximum 2 stars)	Dealing with reverse causation bias (maximum 1 star)	Comparability of the cohorts on the basis of the design or analysis (maximum 2 stars)	Assessment of outcome (maximum 1 star)	Was follow-up long enough for outcomes to occur (maximum 1 star)	Adequacy of follow-up (maximum 1 star)	
Boice et al (1980)	1 star	1 star	1 star	none	1 star	1 star	1 star	1 star	MODERATE
Chen et al (1990)	none	1 star	none	none	none	1 star	1 star	none	HIGH
Christensen et al (2014)	1 star	1 star	2 stars	none	none	1 star	1 star	1 star	HIGH
Davis et al (1989)	1 star	1 star	1 star	1 star	1 star	1 star	1 star	1 star	LOW
Engels et al (2009)	1 star	1 star	none	1 star	none	1 star	1 star	none	HIGH
Floe et al (2018)	1 star	1 star	1 star	none	none	none	1 star	1 star	HIGH
Gao et al (1992)	1 star	1 star	1 star	1 star	1 star	none	1 star	none	HIGH
Hong et al (2016)	1 star	1 star	1 star	1 star	2 stars	1 star	1 star	none	LOW
Leung et al (2013)	1 star	1 star	2 stars	1 star	2 stars	1 star	1 star	1 star	LOW
Merlo et al (1995)	none	1 star	none	none	none	none	1 star	none	HIGH
Ng et al (1990)	none	1 star	none	none	none	1 star	none	1 star	HIGH
Sasaki et al (1992)	1 star	1 star	1 star	1 star	1 star	1 star	none	1 star	LOW

Table 4. Risk of bias in included case-control studies for the lung cancer mortality outcome

| Study | Selection | | | | Comparability | Outcome | | | Overall risk of bias |
	Case definition (maximum 1 star)	Representativeness of the cases (maximum 1 star)	Selection of controls (maximum 1 stars)	Definition of controls (maximum 1 star)	Comparability of cases and controls on the basis of the design or analysis (maximum 2 stars)	Ascertainment of exposure (maximum 2 stars)	Same method of ascertainment for cases and controls (maximum 1 star)	Non-response rate (maximum 1 star)	
Fu et al (1984)	1 star	none	1 star	none	none	none	1 star	none	HIGH

Appendix 9. Results of individual studies

Table 1. Effect size estimates of lung cancer diagnosis risk among persons with a previous episode of TB in cohort studies

Study	Unadjusted effect measure	Unadjusted effect estimate	Lower CI	Upper CI	2 x 2 table*				Adjusted effect measure	Adjusted effect estimate	Lower CI	Upper CI	Factors adjusted for
					A	B	C	D					
An et al (2020)	HR	4.1	3.09	5.45	86	3690	108	18772	HR	4.18	3.15	5.56	Smoking status (ever smoker or current smoker), age, sex, household income
Bae et al (2013)	RR	2.01	1.09	3.47	16	77	642	6274	RR	1.85	1.08	3.19	Age, intake of tomatoes, coffee
Engels et al (2009)	HR†	5.86	3.03	11.37	-	-	-	-	-	-	-	-	-
Everatt et al (2016)	-	not applicable §	-	-	477	21509	-	-	SIR	3.83	3.49	4.19	Standardization based on age and sex
Hong et al (2016)	RR‡	2.70	2.56	2.84	1573	77725	11246	1517166	HR¶	1.38	1.31	1.46	Smoking status (current smokers, ex-smokers and never-smokers), amount of cigarettes per day (1-9, 10-19 and >=20), age, sex, socioeconomic status, alcohol consumption, diabetes mellitus, respiratory diseases hospitalizations
Huang et al (2015)	RR‡	4.49	4.23	4.78	1052	110469	31707	15075796	HR¶	1.62	1.12	2.35	Age, sex, asthma, chronic obstructive pulmonary disease, low income, urbanization, geographical area, diabetes, hyperlipidaemia, chronic kidney disease, smoking related cancers
Jian et al (2016)	-	-	-	-	-	-	-	-	HR	1.08	0.57	2.03	Age, sex, urbanization, inhaled corticosteroids use, medication, chronic obstructive pulmonary disease, smoking related cancers, diabetes, hyperlipidaemia, liver cirrhosis, chronic kidney disease, autoimmune disease, atopy dermatitis, rhinosinusitis, pneumonia
Kuo et al (2013)	-	not applicable §	-	-	159	6540	-	-	SIR	4.09	3.48	4.78	Standardization based on age and sex
Lai et al (2012)	HR	2.96	2.17	4.03	-	-	-	-	HR	1.60	1.16	2.20	Age, sex, diabetes mellitus, chronic obstructive pulmonary disease
Littman et al (2004)	-	-	-	-	-	-	-	-	HR	1.00	0.65	1.54	Years smoked and years smoked squared, average number of cigarettes smoked per day and average number of cigarettes smoked per day squared, smoking status (former or current), age, sex, body mass index, study arm, asbestosis, asthma, chronic bronchitis or emphysema, pneumonia
Liu et al (2017)	HR	2.87	2.15	3.83	57	926	248	12455	HR	2.65	1.95	3.59	Age, income, pneumonia, bronchiectasis, pulmonary fibrosis, hypertension, diabetes, inhaled corticosteroids use

Study				A	B	C	D					Adjustment variables
Oh et al (2020)	HR	2.84	1.41	5.71	-	-	-	-	HR	1.71	0.86	3.39	Smoking status (current smokers, ex- smokers or never-smokers), age, sex, education, income level, body mass index, moderate or vigorous physical activity
Shebl et al (2010)	-	-	-	-	-	-	-	-	HR	1.11	0.81	1.51	Age, sex, race, mode of HIV acquisition, CD4 count at AIDS onset, calendar year of AIDS onset
Shiels et al (2011)	RR	1.52‡	1.16	2.00	44	229	3058	25802	HR	1.97	1.5	2.65	Smoking measured with log cig-years (log [cigarettes smoked per day + 1] x number of years smoked), age
Simonsen et al (2014)	-	not applicable §	-	-	429	14595	-	-	SIR	3.40	3.09	3.74	Standardization based on age and sex
Wu et al (2011)	RR	1.64‡	1.26	2.14	74	5583	191	23793	HR	1.64	1.24	2.15	Age, gender, chronic obstructive pulmonary disease, diabetes, chronic renal failure, autoimmune disease
Wu et al (2016)	-	-	-	-	-	-	-	-	HR	1.42	0.89	2.26	Age, sex, pneumonia, urbanization, inhaled corticosteroids, oral corticosteroids, long-acting agonists, short-acting beta agonists, theophylline, statins, aspirin, chronic kidney disease, diabetes, hyperlipidaemia, liver cirrhosis, smoking-related cancers, autoimmune disease, atopy dermatitis, rhinosinusitis, number of visits for respiratory diseases within 2 years after index data
Yeo et al (2021)	HR	2.57	2.35	2.81	485	22083	16262	1837016	HR	1.34	1.22	1.47	Smoking (pack-years), age, sex, BMI, COPD, diabetes mellitus, alcohol consumption, insurance coverage
Yu et al (2011)	HR	11.9	9.73	14.6	100	4380	1584	710808	HR	3.32	2.7	4.09	Age, sex, occupation, diabetes, hypertension, dyslipidaemia, chronic obstructive pulmonary disease, smoking related cancers

*A: number of exposed with outcome, B: exposed without outcome, C: unexposed with outcome, D: unexposed without outcome reported in the paper. † Unadjusted effect estimate and confidence intervals calculated from unadjusted effect estimates reported separately for lung cancer diagnosis 0-4.9 and >5 years after tuberculosis diagnosis. ‡ Unadjusted effect estimates and confidence intervals calculated from the 2 x 2 table. § The study reports SIR using lung cancer incidence in the general population. ¶Adjusted effect estimate and confidence intervals calculated from adjusted effect estimates reported separately for males and females.

Table 2. Effect size estimates of lung cancer diagnosis risk among persons with a previous episode of TB in case-control studies

Study	Unadjusted effect measure	Unadjusted effect estimate	Lower CI	Upper CI	2 x 2 table*				Adjusted effect measure	Adjusted effect estimate	Lower CI	Upper CI	Set of factors adjusted for
					A	B	C	D					
Alavanja et al (1992)	OR	1.82†	0.85	1.12	15	19	600	1381	OR	2.0	1.0	4.1	Smoking history (lifetime nonsmoker or former smoker), age
Bodmer et al (2012)	OR	0.97	0.84	1.12	226	1395	12817	76863	OR	0.9	0.8	1.05	Smoking status (non-smoker, current, past or unknown), age, sex, body mass index, congestive heart failure, ischemic heart disease-stroke/transient ischemic attack, hypertension, dyslipidaemia, diabetes mellitus, chronic obstructive pulmonary disease
Brenner et al (2001)	OR	3.8†	2.7	5.38	103	59	783	1705	OR	3.8	2.7	5.4	Smoking category (heavy, moderate, light o never-smokers), age, sex, prefecture
Brenner et al (2010)	OR	2.58†	0.65	10.73	6	5	439	943	OR	2.6	0.7	9.2	Pack-years of smoking, age, sex, education and ethnicity
Brownson et al (2000)	OR	0.86†	0.33	2.19	10	12	666	688	OR	0.9	0.4	2.2	Pack-years of smoking
Chan-yeung et al (2003)	OR	1.78†	1.05	3.06	45	27	285	304	OR	1.83‡	1.1	3.19	Smoking duration and amount of cigarettes smoked (<20, 20-39, >40 pack-years), sex
Cheng et al (2012)	OR	3.03	1.79	5.13	26	37	271	1151	-	-	-	-	-
Chen et al (2021)	OR	4.34†	3.53	5.33	233	175	3745	12741	OR	1.44	1.06	1.95	Age and sex
Galeone et al (2008)	OR	3.97†	2.05	7.85	30	17	186	418	OR	3.82	1.97	7.41	Smoking status (never, current or ex-smokers), duration of smoking (for current smokers <25, 25-35 and >35 years; for ex-smokers <5 and > 5 years from the last cigarette), amount of smoking (for current smokers: <10, 10-15 and ≥ 15 cigarettes per day; for ex-smokers: <15 and ≥ 15 cigarettes per day), income, family history of lung cancer and other cancers, occupational exposure to lung carcinogens
Hinds et al (1982)	OR	1.6	0.6	4.3	7	9	203	410	-	-	-	-	-
HosgoodIII et al (2013)	OR	12.56†	3.08	110	24	2	474	496	OR	83.70	11.00	634.70	Smoking (never users; sole users of other types of tobacco or cigarettes smoked with a water pipe, ≤20 pack-years of cigarettes smoked without a water pipe; >20 pack-years of cigarettes smoked without a water pipe), passive smoking, sex, fuel type, educational status, family history of lung cancer
Ko et al (1997)	OR	4.54†	1.39	19.2	16	4	89	101	OR	5.9	1.3	25.9	Socioeconomic status, residential area, education, industrial district, cooking fuels, fume extractor, vegetable consumption

Study													
Koshiol et al (2010)	OR	1.1†	0.75	1.6	60	61	1777	1985	OR	0.96	0.62	1.48	Pack-years and smoking intensity (average packs/day), age, gender, region
Kreuzer et al (2001)	OR	1.19†	0.23	3.97	3	35	55	764	OR	1.2	0.04	1.41	Age, area
Kreuzer et al (2002)	OR	1.7†	0.75	3.74	13	18	218	513	OR	1.61	0.77	3.37	Age, region
Lai et al (2013)	OR	3.66	3.23	4.14	516	655	2306	10633	OR	2.96	2.60	3.37	Smoking (ICD-9 codes, the NIH database is not reliable for this variable), age, sex, Parkinson's disease, chronic obstructive pulmonary disease, pneumoconiosis, asbestosis, alcoholism
Lai et al (2013)	OR	3.31	2.73	4.02	193	248			OR	2.42	1.98	2.95	Smoking (ICD-9 codes, the NIH database is not reliable for this variable), chronic obstructive pulmonary disease, asbestosis
Lee et al (2001)	OR	1.4†	0.87	2.25	45	51	146	231	OR	6.88‡	3.03	15.63	Smoking (cumulative pack-years), residential area, education, socioeconomic status, sex
Liang et al (2009)	OR	4.18†	1.43	14.77	16	5	210	274	OR	4.7	1.6	13.2	Age, passive smoking, marital status, years of schooling, ethnicity, 5 year ago body mass index, coal use, exposure to coal smoke and cooking fumes
Lim et al (2011)	OR	1.66†	0.99	2.73	27	53	406	1322	OR	1.58	0.95	2.62	Age, passive smoking, family history of cancer, fruit and vegetable consumption, country of origin, dialect group, housing type, number of years in school
Liu et al (1993)	OR	2.23	1.51	3.31	101	55	215	261	OR	2.10‡	1.2	3.67	Smoking (not clear how, but they measured cigarettes smoked per day), sex, education, occupation, living area
Lo et al (2013)	OR	2.35†	1.58	3.55	88	39	1433	1495	OR	2.48‡	1.45	4.25	Age, sex, years of education
Luo et al (1996)	OR	2.4	1.2	4.7	16	23	86	283	-	-	-	-	-
Mayne et al (1999)	OR	1.20	0.52	2.79	12	10	421	425	-	-	-	-	-
Osann et al (2000)	OR	1.26†	0.19	6.61	3	5	95	199	OR	1.8	0.2	14.4	Pack-years of smoking, years since quitting smoking, age, education
Park et al (2010)	OR	2.96†	2.25	3.93	276	74	1262	1003	OR	2.56‡	1.85	3.56	Smoking status (ever or never smoker), age, sex
Ramanakumar et al (2006) study I	OR	2.93†	1.17	8.75	26	6	749	506	OR	2.7	1.0	7.4	Smoking status (ever or never smoker), log of cigarettes-year, number of years since quitting smoking (0-2, 2-5, 5-10, 10-15 or >15 years), age, ethnicity, type of respondent (self or surrogate), year of schooling, family income
Ramanakumar et al (2006) study II	OR	1.2†	0.68	2.1	27	29	1178	1512	OR‡	0.90	0.48	1.67	Smoking status (ever or never smoker), log of cigarettes-year, number of years since quitting smoking (0-2, 2-5, 5-10, 10-15 or >15 years), age, sex, ethnicity, type of respondent (self or surrogate), year of schooling, family income
Raspanti et al (2016)	OR	2.17†	1.46	3.25	88	44	518	562	OR	2.30	1.50	3.51	Smoking status (ever or never smoker; they also calculated pack-years of smoking but it is not clear if this

Study	Type	Estimate	Lower	Upper	A	B	C	D	Type	Estimate	Lower	Upper	Adjustments
													was included in the model), age, sex, household air pollution exposure, socioeconomic status, passive smoking
Samet et al (1986)	OR	1.24	0.66	2.29	27	22	491	747	OR	1.40	0.69	2.87	Smoking (duration of smoking in years, number of cigarettes smoked per day on average, duration of cessation in years, and a product term of smoking duration with an indication variable for age above and age below 65), age, sex, ethnicity
Schwartz et al (1996)	OR	2.19	0.58	10.06	8	4	249	273	OR	2.1	0.6	7.1	Age, sex, race
Seow et al (2002)	OR	1.92†	1.10	3.31	27	37	276	726	-	-	-	-	-
Wang et al (1996)	-	-	-	-	-	-	-	-	OR	2.57	1.37	4.80	Smoking status (no more details), chronic bronchitis/emphysema, family history of tumours, passive smoking, consumption of pickled and cured foods
Wang et al (1996_2)	OR	1.39	0.94	3.04	-	-	-	-	-	-	-	-	-
Wang et al (2009)	OR	1.76†	0.61	5.21	10	8	202	284	OR	2.43	0.82	7.20	Age, employment, total dish year, intake of yellow/orange vegetables, dark green vegetables, multivitamins
Wang et al (2014)	OR	2.22†	1.12	4.58	30	14	470	486	-	-	-	-	-
Wu et al (1988)	OR	7.1†	0.91	322.12	7	1	329	335	RR	10.0	1.1	90.1	Pack-years of smoking, years since smoking stopped, depth of inhalation, social class according to father occupation (blues or white-collar worker)
Wu et al (1995)	OR	1.63†	0.87	2.95	19	37	378	1199	OR	1.96	0.9	3.1	Age, passive smoking, area, ethnicity, education
Wu-Williams (1990)	-	-	-	-	103	83	-	-	RR	1.3	0.9	1.7	Age, smoking (non-smoker, smoked 1-9 cigarettes per day and 1-29 years, 2-19 and 30-39 years, 1-19 and >40 years, >20 cigarettes per day and 1-29 years, >20 and 30-39 years, >20 and >40 years), education, study area
Yang et al (2015)	OR	1.74†	1.3	2.34	131	84	1428	1595	OR	1.52	1.13	2.04	Smoking status (ever or never smoker), pack-years of smoking (low, 0-5; moderate, 6-20; or high, >20), age, passive smoking, sex, emphysema, education, body mass index, educational experience, centre, packs-years occupational exposure to metallic toxicant, housing ventilation, biomass burning, cured meat, vegetables/fruits consumption
Zatloukal et al (2003)	OR	0.81	0.47	1.34	20	108	346	1516	OR	1.75‡	1.01	3.05	Smoking (pack-years), age, residence, education, residence
Zheng et al (1987)	OR	1.45†	1.18	1.77	266	213	1105	1279	OR	1.5	1.2	1.8	Smoking category (non-smoker, light, moderate or heavy smoker), age, sex, education
Zhou et al (2000)	OR	1.63	0.63	4.29	15	10	57	62	-	-	-	-	-

*A: number of exposed cases, B: exposed controls, C: unexposed cases, D: unexposed controls reported in the paper. †Unadjusted effect estimate and confidence intervals calculated from the 2 x 2 table. ‡ Adjusted effect estimate and confidence intervals calculated from adjusted effect estimates reported by subgroups (e.g.: males and females or lung cancer subtypes).

Table 3. Effect size estimates of lung cancer mortality risk among persons with a previous episode of TB in cohort studies

Study	Unadjusted effect measure	Unadjusted effect estimate	Lower CI	Upper CI	2 x 2 table*				Adjusted effect measure	Adjusted effect estimate	Lower CI	Upper CI	Set of factors adjusted for
					A	B	C	D					
Boice et al (1980)	-	not applicable ‡	-	-	3	-		-	SMR	1.54§	0.29	3.14	Standardization based on age and sex
Chen et al (1990)	RR	2.72†	1.17	6.33	7	560	22	4817	-	-	-	-	-
Christensen et al (2014)	RR	not applicable			171	6231	278	18928	RR	2.25	1.86	2.72	Age and sex
Davis et al (1989)	-	not applicable ‡	-	-	-	-	-	-	SMR	1.05§	0.71	1.55	Standardization based on age and sex
Engels et al (2009)	HR	6.1	4.3	8.7	31	215	2428	39748	HR	9.70 4.30	4.8 1.8	19 10	Smoking per day (cigarettes per day). Other variables separately adjusted for (results not shown): age, sex, education level, use of smoky coal, asthma, chronic bronchitis, emphysema, family history of tuberculosis, walking hours spent indoors, number of rooms in house
Floe et al (2018)	RR	4.76†	3.81	5.93	165	6536	146	28055	-	-	-	-	-
Gao et al (1992)	-	not applicable ‡	-	-	-	-	-	-	SMR	1.91§	0.98	3.73	Standardization based on age and sex
Hong et al (2016)	RR	3.07†	2.90	3.26	1315	77983	8247	1520165	HR	1.44¶	1.36	1.53	Smoking (current smokers, ex-smokers and never-smokers), amount of cigarettes per day (1-9, 10-19 and >=20), age, sex, alcohol consumption, socioeconomic status, diabetes mellitus, respiratory diseases hospitalizations
Leung et al (2013)	RR	2.61	1.82	3.74	30	486	1314	59409	HR	2.01	1.40	2.90	Smoking status (never-smokers and ever-smokers), age, sex, passive smoking, language, education level, marital status, housing situation, public means, tested financial assistance status, alcohol use, body mass index, chronic obstructive pulmonary disease, asthma, family history of malignancy
Merlo et al (1995)	RR	0.71†	0.32	1.59	7	110	28	305	-	-	-	-	-
Ng et al (1990)	RR	1.64†	0.75	3.59	19	758	9	593	-	-	-	-	-
Sasaki et al (1992)	-	not applicable ‡	-	-	-	-	-	-	SMR	4.57§	2.81	7.42	standardization based on age and sex

*A: number of exposed with outcome, B: exposed without outcome, C: unexposed with outcome, D: unexposed without outcome reported in the paper. † Unadjusted effect estimate and confidence intervals calculated from the 2 x 2 table. ‡ The study reports SMR using lung cancer death rates in the general population. § Adjusted SMR and confidence intervals calculated from data reported in the paper. ¶ Adjusted effect estimate and confidence intervals calculated from adjusted effect estimates reported separately for males and females.

Table 4. Effect size estimates of lung cancer mortality risk among persons with a previous episode of TB in case-control studies

Study	Unadjusted effect measure	Unadjusted effect estimate	Lower CI	Upper CI	numerators and denominators *				Adjusted effect measure	Adjusted effect estimate	Lower CI	Upper CI	Set of factors adjusted for
					A	B	C	D					
Fu et al (1984)	OR	2.86†	1.87	4.45	89	35	434	488	-	-	-	-	-

*A: number of exposed cases, B: exposed controls, C: unexposed cases, D: unexposed controls reported in the paper. † Unadjusted OR and confidence intervals calculated from the 2 x 2 table (the study reports an unadjusted RR).

Appendix 10. Additional forest plots

Figure 1. Forest plot for the association between tuberculosis and subsequent lung cancer diagnosis among cohort studies (model 1)

Study	Hazard Ratio	HR	95%-CI	Weight
An et al (2020)		4.10	[3.09; 5.45]	13.6%
Bae et al (2013)		2.01	[1.13; 3.59]	9.0%
Engels et al (2009)		5.86	[3.03; 11.35]	7.9%
Hong et al (2016)		2.70	[2.56; 2.84]	16.2%
Huang et al (2015)		4.49	[4.22; 4.77]	16.1%
Oh et al (2020)		2.84	[1.41; 5.72]	7.4%
Shiels et al (2011)		1.52	[1.16; 2.00]	13.8%
Yeo et al (2021)		2.57	[2.35; 2.81]	16.0%
Random effects model		**2.96**	**[2.28; 3.83]**	**100.0%**

Heterogeneity: $I^2 = 97\%$
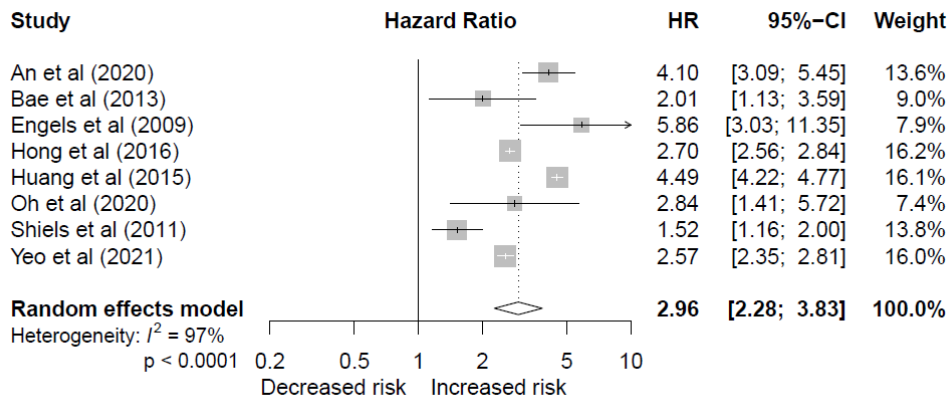p < 0.0001   0.2   0.5   1   2   5   10

Decreased risk   Increased risk

Figure 2. Forest plot for the association between tuberculosis and subsequent lung cancer diagnosis among case-control studies (model 1)

Study	Hazard Ratio	HR	95%-CI	Weight
Alavanja et al (1992)		1.82	[1.59; 2.09]	3.5%
Bodmer et al (2012)		0.97	[0.84; 1.12]	3.5%
Brenner et al (2001)		3.80	[2.69; 5.36]	3.2%
Brenner et al (2010)		2.58	[0.64; 10.48]	1.2%
Brownson et al (2000)		0.86	[0.33; 2.22]	1.9%
Chan−yeung et al (2003)		1.78	[1.04; 3.04]	2.8%
Chen et al (2021)		4.34	[3.53; 5.33]	3.5%
Cheng et al (2012)		3.03	[1.79; 5.13]	2.8%
Galeone et al (2008)		3.97	[2.03; 7.77]	2.5%
Hinds et al (1982)		1.60	[0.60; 4.28]	1.8%
HosgoodIII et al (2013)		12.56	[2.10; 75.06]	0.8%
Ko et al (1997)		4.54	[1.22; 16.87]	1.3%
Koshiol et al (2010)		1.10	[0.75; 1.61]	3.2%
Kreuzer et al (2001)		1.19	[0.29; 4.94]	1.2%
Kreuzer et al (2002)		1.70	[0.76; 3.80]	2.2%
Lai et al (2013) a		3.66	[3.23; 4.14]	3.6%
Lai et al (2013) b		3.31	[2.73; 4.02]	3.5%
Lee et al (2001)		1.40	[0.87; 2.25]	2.9%
Liang et al (2009)		4.18	[1.30; 13.43]	1.5%
Lim et al (2011)		1.66	[1.00; 2.76]	2.9%
Liu et al (1993)		2.23	[1.51; 3.30]	3.1%
Lo et al (2013)		2.35	[1.57; 3.52]	3.1%
Luo et al (1996)		2.40	[1.21; 4.75]	2.4%
Mayne et al (1999)		1.20	[0.52; 2.78]	2.1%
Osann et al (2000)		1.26	[0.21; 7.43]	0.9%
Park et al (2010)		2.96	[2.24; 3.91]	3.3%
Ramanakumar et al (2006) study II		1.20	[0.68; 2.11]	2.7%
Ramanakumar et al (2006) study I		2.93	[1.07; 8.01]	1.8%
Raspanti et al (2016)		2.17	[1.45; 3.24]	3.1%
Samet et al (1986)		1.24	[0.67; 2.31]	2.6%
Schwartz et al (1996)		2.19	[0.53; 9.12]	1.2%
Seow et al (2002)		1.92	[1.11; 3.33]	2.8%
Wang et al (1996_2)		1.39	[0.77; 2.50]	2.7%
Wang et al (2009)		1.76	[0.60; 5.14]	1.7%
Wang et al (2014)		2.22	[1.10; 4.49]	2.4%
Wu et al (1988)		7.13	[0.38; 134.15]	0.4%
Wu et al (1995)		1.63	[0.89; 3.00]	2.6%
Yang et al (2015)		1.74	[1.30; 2.33]	3.3%
Zatloukal et al (2003)		0.81	[0.48; 1.37]	2.8%
Zheng et al (1987)		1.45	[1.18; 1.78]	3.5%
Zhou et al (2000)		1.63	[0.62; 4.25]	1.9%
Random effects model		**2.00**	**[1.65; 2.41]**	**100.0%**

Heterogeneity: I^2 = 89%
p < 0.0001

0.2   0.5   1   2   5   10
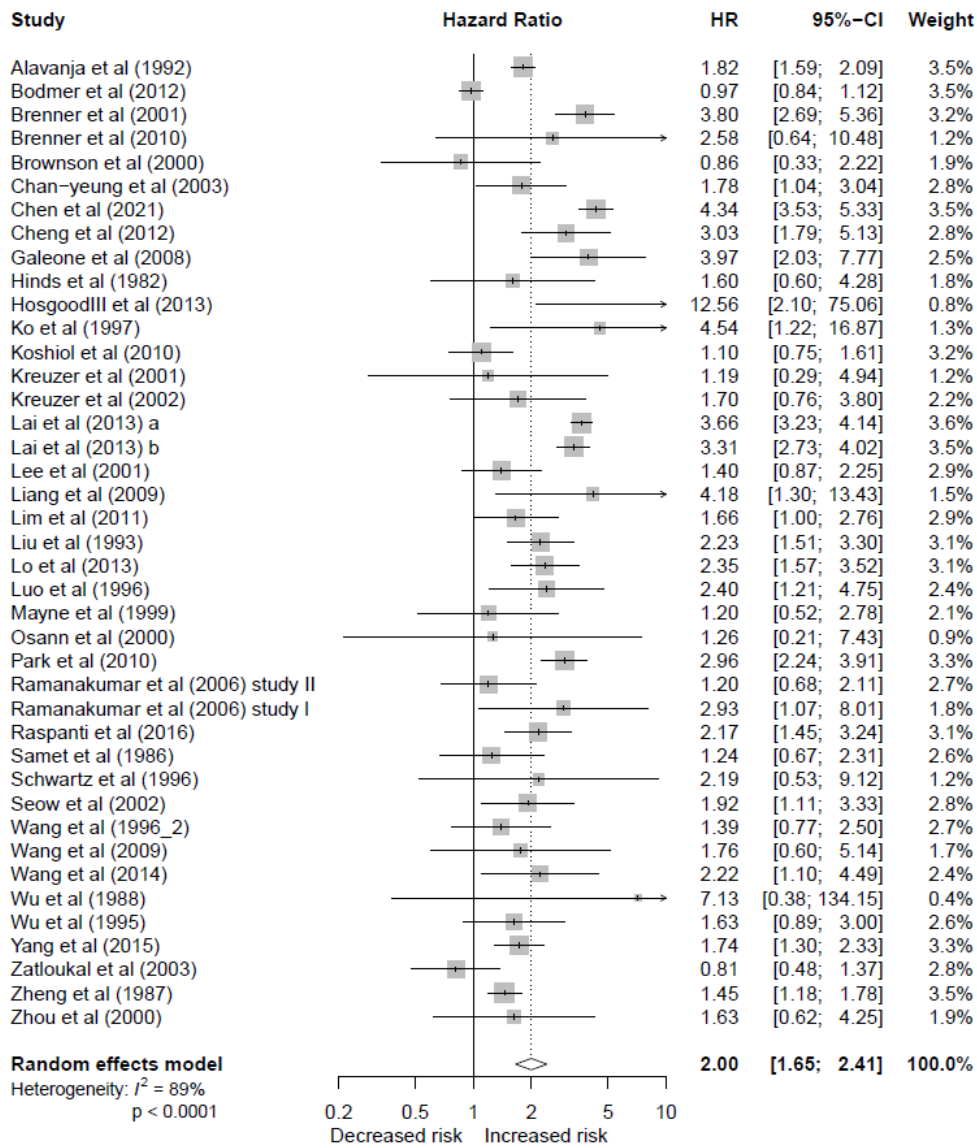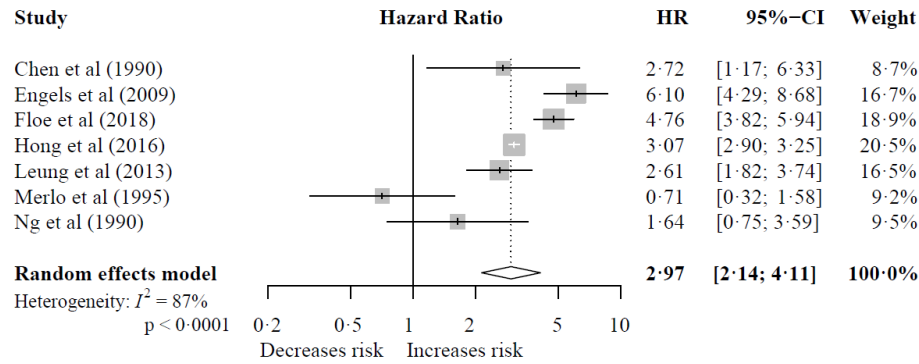Decreased risk   Increased risk

Figure 3. Forest plot for the association between tuberculosis and subsequent lung cancer mortality among cohort studies (model 1)
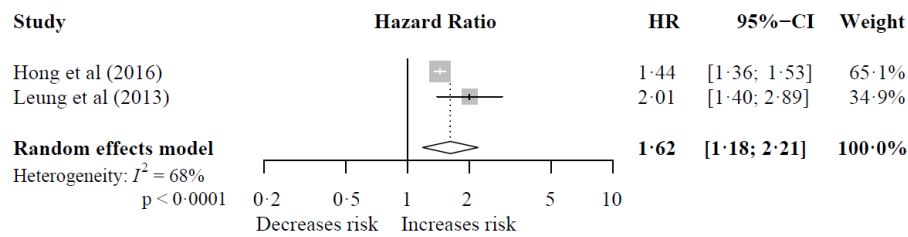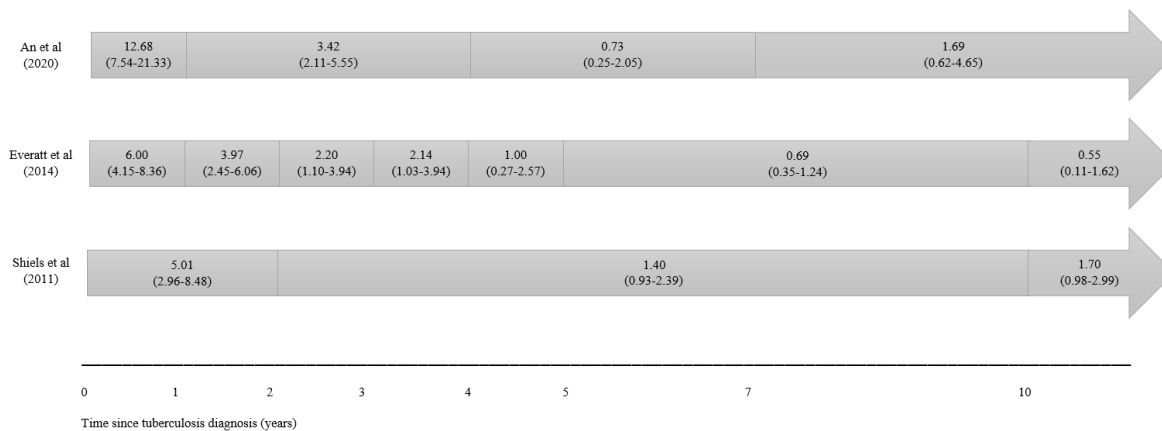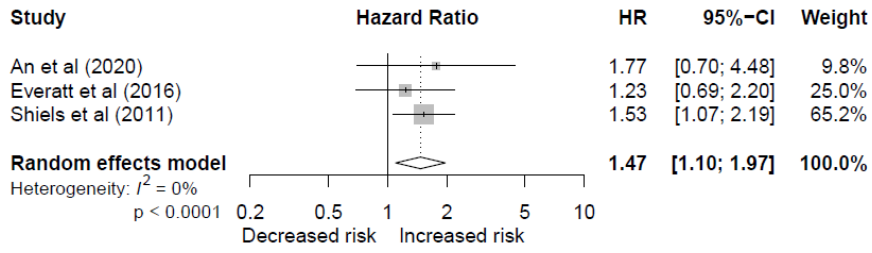
Study	Hazard Ratio	HR	95%−CI	Weight
Chen et al (1990)		2·72	[1·17; 6·33]	8·7%
Engels et al (2009)		6·10	[4·29; 8·68]	16·7%
Floe et al (2018)		4·76	[3·82; 5·94]	18·9%
Hong et al (2016)		3·07	[2·90; 3·25]	20·5%
Leung et al (2013)		2·61	[1·82; 3·74]	16·5%
Merlo et al (1995)		0·71	[0·32; 1·58]	9·2%
Ng et al (1990)		1·64	[0·75; 3·59]	9·5%
Random effects model		**2·97**	**[2·14; 4·11]**	**100·0%**

Heterogeneity: $I^2 = 87\%$
$p < 0.0001$

0·2 0·5 1 2 5 10
Decreases risk Increases risk

Figure 4. Forest plot for the association between tuberculosis and subsequent lung cancer mortality among cohort studies (model 2)

Study	Hazard Ratio	HR	95%−CI	Weight
Hong et al (2016)		1·44	[1·36; 1·53]	65·1%
Leung et al (2013)		2·01	[1·40; 2·89]	34·9%
Random effects model		**1·62**	**[1·18; 2·21]**	**100·0%**

Heterogeneity: $I^2 = 68\%$
$p < 0.0001$

0·2 0·5 1 2 5 10
Decreases risk Increases risk

Appendix 11. Effect estimates by time between tuberculosis diagnosis and lung cancer diagnosis from the three cohort studies reporting by latency strata.

An et al (2020)	12.68 (7.54-21.33)	3.42 (2.11-5.55)	0.73 (0.25-2.05)	1.69 (0.62-4.65)

Everatt et al (2014): 6.00 (4.15-8.36) | 3.97 (2.45-6.06) | 2.20 (1.10-3.94) | 2.14 (1.03-3.94) | 1.00 (0.27-2.57) | 0.69 (0.35-1.24) | 0.55 (0.11-1.62)

Shiels et al (2011): 5.01 (2.96-8.48) | 1.40 (0.93-2.39) | 1.70 (0.98-2.99)

0   1   2   3   4   5   7   10

Time since tuberculosis diagnosis (years)

Appendix 12. Pooled adjusted estimates from cohort studies excluding lung cancer cases detected within one or two years of tuberculosis diagnosis

Study	Hazard Ratio	HR	95%−CI	Weight
An et al (2020)		1.77	[0.70; 4.48]	9.8%
Everatt et al (2016)		1.23	[0.69; 2.20]	25.0%
Shiels et al (2011)		1.53	[1.07; 2.19]	65.2%
Random effects model		1.47	[1.10; 1.97]	100.0%

Heterogeneity: $I^2 = 0\%$

p < 0.0001   0.2   0.5   1   2   5   10

Decreased risk   Increased risk

All the studies included in this analysis controlled for age and any assessment of smoking (model 2). The HR from Shiels et al (2011) and Everatt et al (2016) was calculated after excluding lung cancer cases detected within the first two years of tuberculosis diagnosis. In the study by An et al (2020), cancer cases within the first year of tuberculosis diagnosis were excluded.

Appendix 13. Funnel plots

Figure 1. Adjusted estimates from cohort studies that report the association between tuberculosis and lung cancer diagnosis
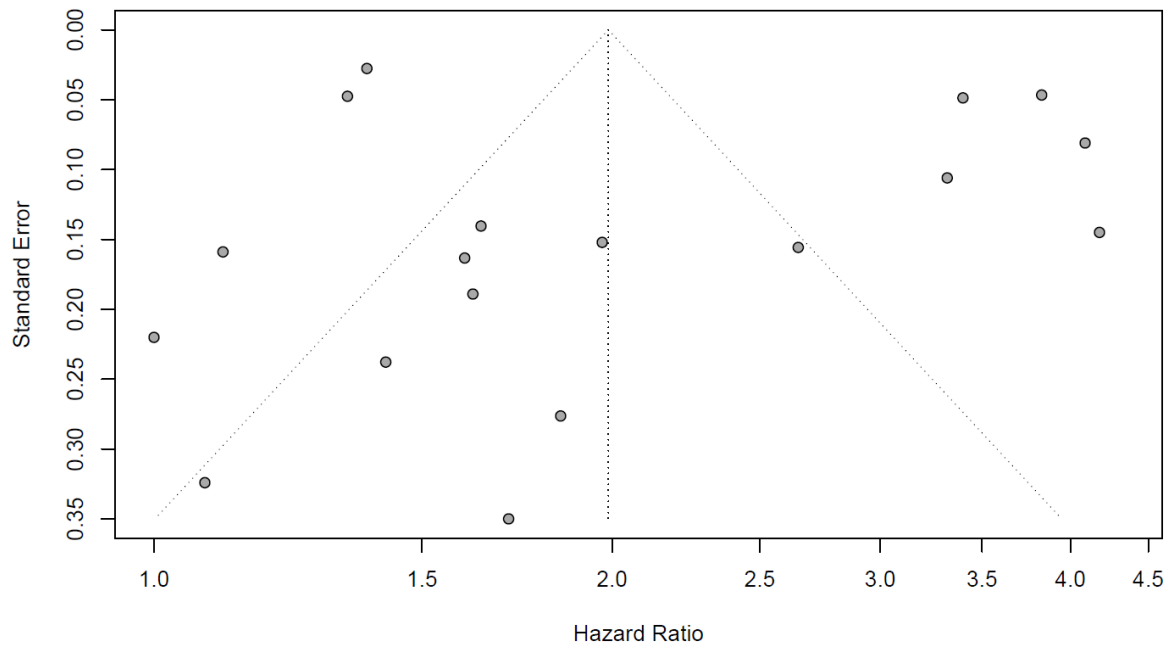
Figure 2. Adjusted estimates from case-control studies that report the association between tuberculosis and lung cancer diagnosis
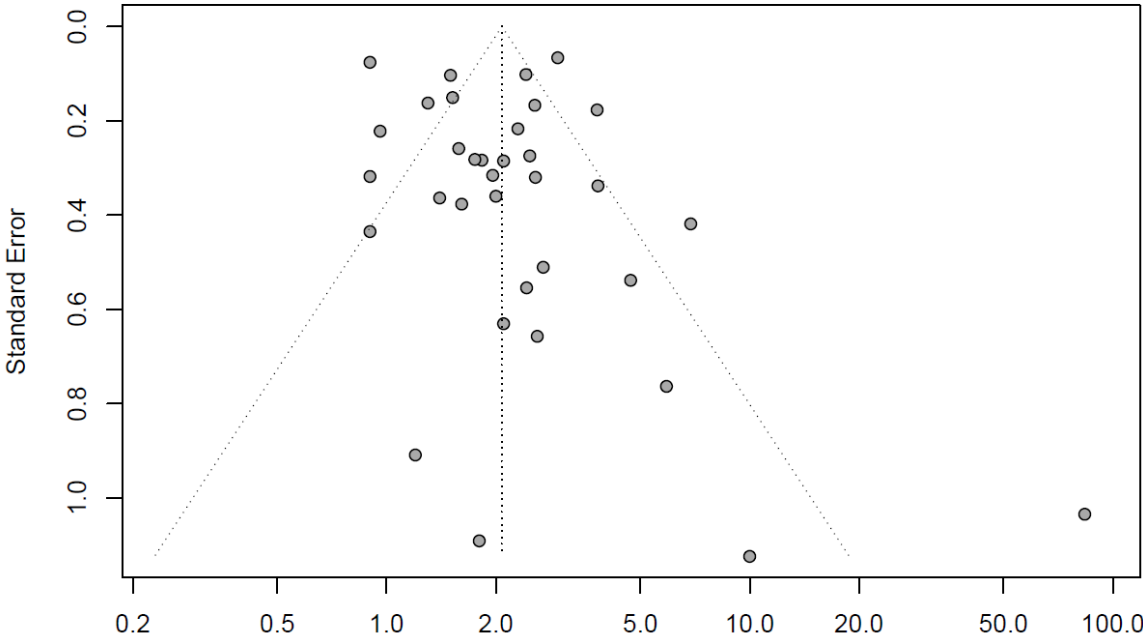
Appendix 14. GRADE assessment of the evidence

Cohort studies									
	Certainty assessment						Effect		
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Pooled HR[*] (95% CI)	Certainty	Importance
Lung cancer diagnosis									
5	Observational	Not serious[†]	Not serious	Not serious	Not serious	None	1.51 (1.30-1.76)	⊕⊕◯◯ Low	Important
Lung cancer mortality									
2	Observational	Serious[‡]	Not serious	Not serious	Not serious	None	1.62 (1.18-2.21)	⊕◯◯◯ Very low	Important

HR: Hazard ratio. [*]Here we considered the pooled estimates form the available model with the most rigorous adjustment for smoking. For lung cancer diagnosis, model 3 provides the most accurate pooled estimate (HR adjusted for age and quantitatively assessed smoking=1.51). For lung cancer mortality, model 3 was not performed so the most accurate estimate is provided by model 2 (HR adjusted for age and any assessment of smoking=1.62). [†]Three of the five studies included in this meta-analysis had low risk of bias. Furthermore, stratified analysis to the three cohorts with low risk of bias yielded consistent results (HR=1.72, 95% CI 1.25-2.38). [‡]Only two studies included in this meta-analysis and therefore not possible to perform an analysis restricted to studies with low risk of bias.

OR=Odds ratio. [*]Here we considered the pooled estimates form the available model with the most rigorous adjustment for smoking (model 3, table 2 in the main manuscript). [†]Only one out of the 19 case-control studies in model 3 had low risk of bias. Ten and eight had moderate and high risk of bias, respectively.

Case-control studies									
	Certainty assessment						Effect		
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Pooled OR[*] (95% CI)	Certainty	Importance
Lung cancer diagnosis									
19	Observational	Serious[†]	Not serious	Not serious	Not serious	None	1.74 (1.42-2.13)	⊕◯◯◯ Very low	Important

Appendix 15. List of variables adjusted for in the included studies

Adjustment for	Lung cancer diagnosis		Lung cancer mortality
	Cohort studies (n=19)	Case-control studies (n=43)	Cohort studies (n=12)
Smoking	7*	32¶	3
Age	18	25	7
Sex	18†	32#	9**
Any socioeconomic status indicator (income, education or occupation)	6	18	2
Ethnicity	1	4	0
Location (urbanization, area, residence or prefecture)	3	10	0
Passive smoking	0	6	1
Comorbidities			
Any comorbidity	10	6	4
COPD	8‡	4	1
Diabetes	7	1	1
Pneumonia	4	0	0
Dyslipidaemia	4	1	0
Chronic kidney disease/failure	4	4	0
Asthma	3§	0	1
Smoking-related cancers	3	0	0
Autoimmune disease	3	0	0
Hypertension	2	1	0
Liver cirrhosis	2	0	0
Atopy dermatitis	2	0	0
Rhinosinusitis	2	0	0
Chronic bronchitis/emphysema	1	3	0
Bronchiectasis	1	0	0
Pulmonary fibrosis	1	0	0
Asbestosis	1	2	0
Silicosis	0	0	2††
Pneumoconiosis	0	1	0
AIDS	1	0	0
Congestive heart failure	0	1	0
Ischemic heart disease	0	1	0
Stroke/transient ischemic attack	0	1	0
Inhaled corticosteroids use	3	0	0
Oral corticosteroids use	1	0	0
Body mass index	3	3	1
Physical activity	1	0	0
Alcohol	2	1	1
Hospitalizations for respiratory diseases	1	0	1
Any adjustment for diet	1	5	0
Family history of lung cancer	0	4	1

*Everatt et al (2016) reports an adjusted estimate for "non-smokers" equivalent to "never-smokers" definition. † Bae et al (2013), Liu et al (2017), and Shiels et al (2011) restricted intake to one sex (females or males). Huang et al (2015) reported adjusted estimates stratified by sex. ‡ Liu et al (2017) and Wu et al (2016) restricted intake to COPD patients. § Jian et al (2016) restricted intake to asthma patients. ¶ Kreuzer et al (2001), Kreuzer et al (2002), Liang et al (2009), Lo et al (2013), Schwartz et al (1996), Wang et al (1996) b, Wang et al (2009), and Wu et al (1995) restricted intake to never-smokers. # Alavanja et al (1992), Brownson et al (2000), Cheng et al (2012), Hinds et al (1982), Ko et al (1997), Kreuzer et al (2001), Kreuzer et al (2002), Lai et al (2013), Liang et al (2009), Lim et al (2011), Osann et al (2000), Ramanakumar et al (2006) – study l, Seow et al (2002), Wang et al (1996) b, Wang et al (2009), Wu et al (1988), Wu et al (1995), Wu-Williams et al (1990), Zatloukal et al (2003), and Zhou et al (2000) restricted intake to one sex (females or males). Chan-Yeung et al (2003), Lee et al (2001), Liu et al (1993), Lo et al (2013), Park et al (2010), Ramanakumar et al (2006) – study ll reported results stratified by sex. ** Boice et al (1980), Chen et al (1990), and Ng et al (1990) restricted intake to one sex (females or males). ††Merlo et al (1995) and Ng et al (1990) restricted to patients with silicosis.

Appendix 16. Amendments to the protocol

The study was prospectively registered in PROSPERO. A first version of the protocol was published in PROSPERO on 05/07/2020. The start of the review was delayed due to the COVID-19 outbreak in Peru, and the authors uploaded a new version of the protocol before starting the data extraction. We added a secondary outcome (lung cancer mortality), further databases to search (Scopus, conference abstracts) to make the review more comprehensive, and pre-specified the subgroup analysis (available online on 26/02/2021). No more versions of the protocol were published.

Some details of the analysis could not possibly be pre-specified in the protocols. We had to define the core set of factors for the meta-analysis of adjusted effect estimates after ascertaining, during data extraction, adjustment approaches used in the studies. We then also established the different methods studies used to control for smoking and decided to develop two models to pool adjusted estimates. Importantly, we defined these details before starting the statistical analysis. Apart from that, all further not pre-specified analyses are labelled as such in the manuscript.

Appendix 17. List of excluded studies with reasons

Study ID	Reason for exclusion and study reference
Abou et al (2017)	Wrong study design: Systematic Review Abou Chakra C, Cheng M, Cnossen S, et al. Risk of Active Tuberculosis in Patients with Cancer: A Systematic Review and Meta-Analysis. Clin Infect Dis 2017; 64: 635–44.
Aerts et al (2012)	Wrong study design Aerts J, Bakker M, Hegmans, et al. History of tuberculosis as an independent prognostic factor for lung cancersurvival. Lung Cancer 2012; 76: 452–6.
Ahmed et al (2014)	Exposure of active TB cannot be determined Ahmed F, Al Emran A, Bin Imran I, et al. Score based risk assessment of lung cancer and its evaluation for Bangladeshipeople. Asian Pac J Cancer Prev 2014; 15: 7021–7.
Ahrens et al (2014)	This article was a pooled analysis of case-control studies Ahrens W, Behrens T, Bencko V, Boffetta P, DR B, Bruning T, et al. Is previous respiratory disease a risk factor for lung cancer? Am J Respir Crit Care Med [Internet]. 2014;190(5 PG-549–59):549–59.
An et al (2020)	Wrong study design: Genetic study An S, Ashikawa K, Bassig B, et al. Tuberculosis infection and lung adenocarcinoma: Mendelian randomization andpathway analysis of genome-wide association study data from never-smoking Asianwomen. Genomics 2020; 112: 1223–32.
Andrew et al (2012)	This article was a pooled analysis of observational studies Andrew A, Bencko V, Bickeboller H, et al. Previous lung diseases and lung cancer risk: a pooled analysis from theInternational Lung Cancer Consortium. Am J Epidemiol 2012; 176: 573–85.
Ariannia et al (2020)	Wrong study design Ariannia A, Ashaari M, Fazel A, Ferlay J, Ghasemi-Kebria F, Ghayoriardahaei H, et al. Increasing trends of lung cancer in Golestan province, Northern Iran (2004-2016). Cancer Epidemiol [Internet]. 2020;65(PG-101687):101687.
Au et al (2012)	Exposure of interest was asbestosis and previous history of TB Au R, Chen M, Tse L, Wang F, XR W, IT Y. Pulmonary tuberculosis and lung cancer mortality in a historical cohort ofworkers with asbestosis. Public Health 2012; 126: 1013–6.
Ba et al (2019)	No control group Ba O, Baddredine H, Cisse M, et al. [Epidemiology of primary lung cancer among non-smokers in Senegal]. Rev Mal Respir 2019; 36: 15–21.
Boffetta et al (2012)	Wrong study design: Review Boffetta P, Sisti J. What proportion of lung cancer in never-smokers can be attributed to known riskfactors? Int J Cancer [Internet]. 2012;131(2 PG-265–75):265–75.
Boice et al (2019)	Exposure of active TB cannot be determined Boice Jr J, Cohen S, Ellis E, Golden A, Mumma M, Zablotska L. Sex-specific lung cancer risk among radiation workers in the million-person studyand patients TB-Fluoroscopy. Int J Radiat Biol 2019; : 1–12.
Brenner et al (2011)	Wrong study design: Systematic Review Brenner D, Hung R, McLaughlin J. Previous lung diseases and lung cancer risk: a systematic review andmeta-analysis. PLoS One 2011; 6: e17479.
Caporaso et al (2006)	Exposure of active TB cannot be determined Caporaso N, Consonni D, Gao Y, et al. Family history of cancer and nonmalignant lung diseases as risk factors for lung cancer. Int J Cancer 2009; 125: 146–52.
Cha et al (2009)	Wrong study design Cha S, Jung T, Kim C, et al. The clinical course of respiratory tuberculosis in lung cancer patients. Int J Tuberc Lung Dis 2009; 13: 1002–7.
Chang et al (2009)	Exposure of active TB cannot be determined Chang YL, Chul MA, Jeong HJ, Hyung JK, Se KK, Chang J, et al. Association of insulin receptor substrate-1 G972R variant with non-small cell lung cancer risk. Tuberc Respir Dis [Internet]. 2009;67(1 PG-8–13):8–13.
Chen et al (2007)	Exposure of active TB cannot be determined

	Chen W, Yang J, Chen J, Bruch J. Exposures to silica mixed dust and cohort mortality study in tin mines: Exposure-response analysis and risk assessment of lung cancer. Am J Ind Med [Internet]. 2006;49(2 PG-67–76):67–76.
Chia et al (2014)	Wrong study design Chia K, Lim W, Loy E, Omkar Prasad R, Seow A, Tan C. Lung cancer incidence in Singapore: ethnic and gender differences. Lung Cancer 2014; 84: 23–30.
Chou et al (2011)	Duplicate article Chou Y, Hu H, Huang N, et al. Pulmonary tuberculosis increases the risk of lung cancer: a population-basedcohort study. Cancer 2011; 117: 618–24.
Christopoulos et al (2014)	Wrong study design: Systematic Review Christopoulos A, Saif M, Sarris E, Syrigos K. Epidemiology of active tuberculosis in lung cancer patients: a systematic review. Clin Respir J 2014; 8: 375–81.
Colletti et al (2015)	No control group Colletti P, Fellner F, Gabriel M, et al. Malignant disease as an incidental finding at (1)(8)F-FDG-PET/CT scanning inpatients with granulomatous lung disease. Nucl Med Commun 2015; 36: 430–7.
Cukic et al (2017)	No control group Cukic V. The Association Between Lung Carcinoma and Tuberculosis. Med Arch [Internet]. 2017;71(3 PG-212–214):212–4.
Dong et al (2009)	Exposure of active TB cannot be determined Dong D, Hu P, Sun Y, Xu G. Lung cancer among workers exposed to silica dust in Chinese refractory plants. Scand J Work Env Heal [Internet]. 1995;21 Suppl 2(PG-69-72):69–72.
Esfahani et al (2018)	Wrong type of publication: Letter to the editor Esfahani B, Keikha M. The Relationship between Tuberculosis and Lung Cancer. Adv Biomed Res 2018; 7: 58.
Figueroa et al (2018)	Exposure of active TB cannot be determined Figueroa CGS, Plata RF, Briseño DM, de la Garza SR, Pizano AM, Marina FF, et al. Analysis of a routine database to identify risk factors of the host and the environment associated with respiratory diseases . Rev Inst Nac Enferm Respir [Internet]. 2012;71(1 PG-11–20):11–20.
Fol et al (2021)	Wrong study design: Review Fol M, Koziński P, Kulesza J, Białecki P, Druszczyńska M. Dual Nature of Relationship between Mycobacteria and Cancer. Int J Mol Sci [Internet]. 2021;22(15 PG-).
Fukami et al (2020)	Wrong study design Fukami T, Hebisawa A, Takahashi F, Tamura A. Recent trends in the incidence of latent tuberculosis infection in Japanesepatients with lung cancer: A small retrospective study. J Infect Chemother [Internet]. 2020;26(3 PG-315–317):315–7.
Furlow et al (2018)	Wrong type of publication: Brief discussion Furlow B. Tobacco control, lung cancer, and tuberculosis in Singapore. Lancet Respir Med [Internet]. 2018;6(10 PG-741–742):741–2.
Golsha et al (2009)	No control group Golsha R, Rezaei SR, Shafiee A, Najafi L, Dashti M, Roshandel G. Pulmonary tuberculosis and some underlying conditions in Golestan Province of Iran, during 2001-2005. J Clin Diagn Res [Internet]. 2009;3(1 PG-1302–1306):1302–6.
Goo et al (2006)	Exposure of active TB cannot be determined Goo J, Kim H, Kim H, Kim Y. Lung Cancer CT Screening and Lung-RADS in a Tuberculosis-endemic Country: TheKorean Lung Cancer Screening Project (K-LUCAS). Radiology 2020; : 192283.
Guan et al (2009)	Wrong study design: Systematic Review Guan P, He Q, Li X, et al. Facts and fiction of the relationship between preexisting tuberculosis and lungcancer risk: a systematic review. Int J Cancer 2009; 125: 2936–44.
Han et al (2018)	Wrong study design Han K, Hong S, Kim S, et al. Effect of pre existing respiratory conditions on survival of lung cancerpatients: A nationwide population-based cohort study. Asia Pac J Clin Oncol 2018; 14: e71–80.

Ho et al (2014)	Duplicate article Ho C, Huang J, Jan S, et al. The coexistence of common pulmonary diseases on the histologic type of lungcancer in both genders in Taiwan: a STROBE-compliant article. Med 2014; 93: e127.
Ho et al (2015)	Wrong study design Ho C, Huang J, Jan S, et al. Impact of coexisting pulmonary diseases on survival of patients with lungadenocarcinoma: a STROBE-compliant article. Med 2015; 94: e443.
Ho et al (2020)	Lung cancer as outcome cannot be determined Ho L, Yang H, Chung C, et al. Increased risk of secondary lung cancer in patients with tuberculosis: A nationwide, population-based cohort study. PLoS One 2021; 16: e0250531.
Hoshuyama et al (1995)	Exposure of active TB cannot be determined Hoshuyama T, Pan G, Tanaka C, Feng Y, Yu L, Liu T, et al. Mortality of iron-steel workers in Anshan, China: A retrospective cohort study. int J Occup Environ Heal [Internet]. 2006;12(3 PG-193–202):193–202.
Hughes et al (2001)	Exposure of active TB cannot be determined Hughes J, McDonald A, McDonald J, RJ R, Shi R, Weill H. Cohort mortality study of North American industrial sand workers. II.Case-referent analysis of lung cancer and silicosis deaths. Ann Occup Hyg 2001; 45: 201–7.
Jo et al (2015)	Wrong study design Jo YS, Choi SM, Lee J, Park YS, Lee S-M, Yim J-J, et al. The relationship between chronic obstructive pulmonary disease and comorbidities: A cross-sectional study using data from KNHANES 2010-2012. Respir Med [Internet]. 2015;109(1 PG-96–104):96–104.
Kikuchi et al (2007)	Exposure of active TB cannot be determined Kikuchi S. Family history and mortality in the Japan Collaborative Cohort Study forEvaluation of Cancer (JACC). Asian Pac J Cancer Prev [Internet]. 2007;8 Suppl(PG-21-4):21–4.
Kinlen et al (1988)	Exposure of active TB cannot be determined Kinlen L, Willows A. Decline in the lung cancer hazard: a prospective study of the mortality of ironore miners in Cumbria. Br J Ind Med 1988; 45: 219–24.
Marsh et al (2017)	Exposure of active TB cannot be determined Marsh GM, Buchanich JM, Zimmerman S, Liu Y, Balmert LC, Esmen NA, et al. Mortality among Hardmetal Production Workers: US Cohort and Nested Case-Control Studies. J Occup Environ Med [Internet]. 2017;59(12 PG-306–326):e306–26.
Moshammer et al (2004)	Exposure of active TB cannot be determined Moshammer H, Neuberger M. Lung cancer and dust exposure: Results of a prospective cohort study following 3260 workers for 50 years. Occup Environ Med [Internet]. 2004;61(2 PG-157–162):157–62.
Ording et al (2018)	Exposure of active TB cannot be determined Ording AG, Veres K, Farkas DK, Adelborg K, Sørensen HT. Risk of cancer in patients with epistaxis and haemoptysis. Br J Cancer [Internet]. 2018;118(6 PG-913–919):913–9.
Pira et al (2017)	Exposure of active TB cannot be determined Pira E, Romano C, Donato F, Pelucchi C, Vecchia CL, Boffetta P. Mortality from cancer and other causes among Italian chrysotile asbestos miners. Occup Environ Med [Internet]. 2017;74(8 PG-558–563):558–63.
Rodescu et al (1981)	Wrong type of publication: Case reports or case series; Rodescu D, Abeles H, Zelefsky M, Henry Williams Jr. M. Accelerated growth of lung cancer in asociation with rifampicin administration for tuberculosis. Lancet [Internet]. 1981;318(8253 PG-983):983.
Saleh et al (2019)	Wrong study design Saleh P, Hosseini M-S, Piri R, Ghaffari M, Mohammadi S, Naghavi-Behzad M. Association of lung cancer and tuberculosis: A cross sectional study from northwest of Iran. Internat Jour Canc Manag [Internet]. 2019;12(6 PG-).
Schubauer-Berigan et al (2009)	Exposure of active TB cannot be determined Schubauer-Berigan MK, Daniels RD, Pinkerton LE. Radon exposure and mortality among white and American Indian uranium miners: An update of the Colorado Plateau cohort. Am J Epidemiol [Internet]. 2009;169(6 PG-718–730):718–30.
Shen et al (2021)	Wrong study design Shen B-J, Lin H-H. Time-dependent association between cancer and risk of tuberculosis: A population-

	based cohort study. Int J Infect Dis [Internet]. 2021;108(PG-340-346):340–6.
Shuldiner et al (2016)	Lung cancer as outcome cannot be determined Shuldiner J, Leventhal A, Chemtob D, Mor Z. Mortality after anti-tuberculosis treatment completion: Results of long-term follow-up. Int J Tuberc Lung Dis [Internet]. 2016;20(1 PG-43–48):43–8.
Su et al (2014)	Language other than English, French or Spanish Su M, Zhou B. Association of Genetic Polymorphisms in IL-6 and IL-1β gene with Risk of Lung Cancer in Female Non-Smokers. Chin J Lung Cancer [Internet]. 2014;17(8 PG-612–617):612–7.
Su et al (2016)	Exposure of active TB cannot be determined Su VY-F, Yen Y-F, Pan S-W, Chuang P-H, Feng J-Y, Chou K-T, et al. Latent tuberculosis infection and the risk of subsequent cancer. Medicine (Baltimore) [Internet]. 2016;95(4 PG-).
Tse et al (2014)	Data provided was not enough to extract or calculate a risk estimate Tse LA, Lin X, Li W, Qiu H, Chan CK, Wang F, et al. Smoking cessation sharply reduced lung cancer mortality in a historical cohort of 3185 Chinese silicotic workers from 1981 to 2014. Br J Cancer [Internet]. 2018;119(12 PG-1557–1562):1557–62.
Vento et al (2011)	Wrong type of publication: Comment Vento S, Lanzafame M. Tuberculosis and cancer: A complex and dangerous liaison. Lancet Oncol [Internet]. 2011;12(6 PG-520–522):520–2.
Wiwanitkit et al (2014)	Wrong type of publication: Letter to the editor Wiwanitkit S, Wiwanitkit V. Tuberculosis and lung cancer. South Asian J Cancer [Internet]. 2014;3(2 PG-141):141.