Supplementary Section

Technical Appendix

METHODS

Study eligibility criteria

The PICO framework was adopted in our search strategy, and studies were selected based on pre-defined inclusion and exclusion criteria which were summarised in Table 1. The population was never-smokers. The exposure of interest was indoor or residential radon exposure, and the comparator was no or low radon exposure. The outcome was lung cancer diagnosis or death. Published articles were included if they reported full or subgroup analyses of the association between residential radon exposure and LCINS. Editorials, conference proceedings, abstracts, posters, narrative reviews, commentaries or grey literature (referring to studies that are either unpublished or have been published in non-commercial form) were excluded. This review was restricted to cohort and case-control studies and pooled data analyses, systematic reviews or meta-analyses thereof. If the same data were reported in individual cohort or case–control studies and also included in a pooled collaborative study, to avoid duplication of results only the results of the pooled collaborative analysis were included. Where data from a primary study was included in more than one pooled collaborative analysis, we obtained directly from the study investigators a re-analysis of the pooled collaborative results excluding that particular study.

Data extraction and management

For the meta-analysis two reviewers (EC and SE) extracted effect estimates and standard errors for never-smokers and categories of ever-smokers (i.e., ever-smoker; or current-smoker and ex-smoker; or ever-smoker with lifetime exposure divided into tertiles) in each study with discrepancies resolved by consensus or adjudication from another reviewer (XQY).

Preferably, estimates of the adjusted excess relative risk (aERR) per 100 Bq/m^3 were extracted if available. The common confounders included in the adjustment were age, sex, education, occupations with high risk of lung cancer and exposure to environmental tobacco smoke. aERR per 100 Bq/m^3 are typically

approximated in case-control studies from linear excess odds ratio (LEOR) models of the form OR = 1 + Bx, where x is a continuous covariate representing an individual's average radon exposure during an exposure time window and B is the ERR parameter per 100 Bq/m^3. Estimated aERR from LEOR models typically have right skewed confidence intervals (CIs) on the excess relative risk (ERR), relative risk (RR) and log(RR) scales (with RR=ERR+1). If estimates of the aERR per 100 Bq/m^3 from LEOR models were not available, adjusted relative risks (aRRs) for categories of radon exposure relative to a reference category were extracted if available. These aRRs are typically estimated in case-control studies using logistic regression with RRs approximated by odds ratios. Estimated RRs for categories of exposure from logistic regressions have right skewed CIs on the ERR and RR scales, but symmetric CIs on the log(RR) scale. The category-based aRRs were then used to estimate aERR per 100 Bq/m^3 using the methods outlined by Greenland et al. [1, 2]. For studies that reported estimates for more than one category of ever-smoker (i.e., current-smoker and ex-smoker; or ever-smoker with lifetime exposure divided into tertiles), estimated aERRs per 100 Bq/m^3 were pooled across the categories corresponding to ever-smokers using the generic inverse variance methods, thus obtaining a single estimated aERRs per 100 Bq/m^3 for ever-smokers for further pooling. For cohort studies, the reported measures of effect were either the adjusted incidence rate ratio (aIRR) or hazard ratio (aHR), and these were extracted where available. Stata 14 was used for statistical analyses.

Sensitivity analysis

The inverse variance method for pooling tends to perform better when component effect estimates are normally distributed, but aERR estimates from LEOR models typically have right skewed confidence intervals (CIs) on the ERR, RR and log(RR) scales (with RR=ERR+1). However, ERR estimates from LEOR models tend to be less skewed on the log(RR) scale than on the ERR scale, and ERR estimates derived from estimated RRs for categories of exposure from logistic regression have symmetric CIs in the log(RR) scale. Hence, given that component estimated ERRs are either symmetric on the log(RR) scale or at least more symmetric on the log(RR) scale than on the ERR scale, we performed sensitivity analyses in which the component effect estimates were pooled on the log(RR) scale using the standard generic inverse variance method. For this analysis, effect estimates were exponentiated and displayed on the RR scale.

Supplementary Table 1. **PRISMA-P 2015 Checklist**

This checklist has been adapted for use with protocol submissions to *Systematic Reviews* from Table 3 in Moher D et al. [3]**:** Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews* 2015 **4**:1

Section/topic	#	Checklist item	Information reported		Page number(s)
			Yes	No	
ADMINISTRATIVE INFORMATION					
Title					
Identification	1a	Identify the report as a protocol of a systematic review	☒	☐	(p.1)
Update	1b	If the protocol is for an update of a previous systematic review, identify as such	☐	☒	N/A
Registration	2	If registered, provide the name of the registry (e.g., PROSPERO) and registration number in the Abstract	☒	☐	In accordance with the guidelines, our systematic review protocol was registered with the International Prospective Register of Systematic Reviews PROSPERO and the registration number is CRD42020154551.
Authors					
Contact	3a	Provide name, institutional affiliation, and e-mail address of all protocol authors; provide physical mailing address of corresponding author	☐	☒	
Contributions	3b	Describe contributions of protocol authors and identify the guarantor of the review	☐	☒	
Amendments	4	If the protocol represents an amendment of a previously completed or published protocol, identify as such and list changes; otherwise, state plan for documenting important protocol amendments	☐	☒	N/A
Support					
Sources	5a	Indicate sources of financial or other support for the review	☐	☒	N/A
Sponsor	5b	Provide name for the review funder and/or sponsor	☐	☒	N/A

Section/topic	#	Checklist item	Information reported		Page number(s)
			Yes	No	
Role of sponsor/funder	5c	Describe roles of funder(s), sponsor(s), and/or institution(s), if any, in developing the protocol	☐	☒	N/A
INTRODUCTION					
Rationale	6	Describe the rationale for the review in the context of what is already known	☒	☐	(p.3,4)
Objectives	7	Provide an explicit statement of the question(s) the review will address with reference to participants, interventions, comparators, and outcomes (PICO)	☒	☐	(p.4)
METHODS					
Eligibility criteria	8	Specify the study characteristics (e.g., PICO, study design, setting, time frame) and report characteristics (e.g., years considered, language, publication status) to be used as criteria for eligibility for the review	☒	☐	(p.5,6 and Table 1)
Information sources	9	Describe all intended information sources (e.g., electronic databases, contact with study authors, trial registers, or other grey literature sources) with planned dates of coverage	☒	☐	(Table 1)
Search strategy	10	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated	☒	☐	(p.5 and Supplementary Table 3)
STUDY RECORDS					
Data management	11a	Describe the mechanism(s) that will be used to manage records and data throughout the review	☒	☐	(p.5,6 and Figure 1)
Selection process	11b	State the process that will be used for selecting studies (e.g., two independent reviewers) through each phase of the review (i.e., screening, eligibility, and inclusion in meta-analysis)	☒	☐	(p.5,6 and Figure 1)
Data collection process	11c	Describe planned method of extracting data from reports (e.g., piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators	☒	☐	(p.6,7)
Data items	12	List and define all variables for which data will be sought (e.g., PICO items, funding sources), any pre-planned data assumptions and simplifications	☒	☐	(p.6,7)

Section/topic	#	Checklist item	Information reported		Page number(s)
			Yes	No	
Outcomes and prioritization	13	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale	☒	☐	(p.6,7)
Risk of bias in individual studies	14	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis	☒	☐	(p.6 and Supplementary Table 1)
DATA					
Synthesis	15a	Describe criteria under which study data will be quantitatively synthesized	☒	☐	(p.8)
	15b	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data, and methods of combining data from studies, including any planned exploration of consistency (e.g., I^2, Kendall's tau)	☒	☐	(p.8)
	15c	Describe any proposed additional analyses (e.g., sensitivity or subgroup analyses, meta-regression)	☒	☐	(p.8,9)
	15d	If quantitative synthesis is not appropriate, describe the type of summary planned	☐	☒	N/A
Meta-bias(es)	16	Specify any planned assessment of meta-bias(es) (e.g., publication bias across studies, selective reporting within studies)	☐	☒	N/A
Confidence in cumulative evidence	17	Describe how the strength of the body of evidence will be assessed (e.g., GRADE)	☒	☐	(p.16)

Supplementary Table 2. **Initial search - search terms & databases**

Database(s): Embase Classic+Embase 1947 to 2017 November 28, Ovid MEDLINE(R) Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) 1946 to Present

KEYWORDS	#	Search terms
Lung cancer	1	exp Lung Neoplasms/
	2	(pulmon$ adj3 (cancer$ or carcinoma$ or malignan$ or tumo?r$ or neoplas$ or metast$ or adeno$)).mp.
	3	(lung adj3 (cancer$ or carcinoma$ or malignan$ or tumo?r$ or neoplas$ or metast$ or adeno$)).mp.
Radon	4	Radon Daughters/ or Radon/
	5	radon.tw.

Supplementary Table 3. **Final search - search strategy, databases & results**

Database(s): **Embase Classic+Embase** 1947 to 2020 March 05**, Ovid MEDLINE(R) ALL** 1946 to March 05, 2020 (Run on 6 Mar 2020)

Search Strategy:

#	Searches	Results
1	exp Lung Neoplasms/	616120
2	(pulmon$ adj3 (cancer$ or carcinoma$ or malignan$ or tumo?r$ or neoplas$ or metast$ or adeno$)).mp.	76020
3	(lung adj3 (cancer$ or carcinoma$ or malignan$ or tumo?r$ or neoplas$ or metast$ or adeno$)).mp.	775413
4	1 or 2 or 3	811282
5	Radon Daughters/ or Radon/	14599
6	radon.tw.	16413
7	5 or 6	19195
8	4 and 7	3933
9	limit 8 to english language	3610
10	limit 9 to human	2936
11	limit 10 to yr="1990 -Current"	2521
12	remove duplicates from 11	1628

Supplementary Table 4. Risk of bias assessment for included studies

ROB assessment[a] for Pooled collaborative studies

Criteria	1. Study eligibility	2. Identification/selection	3. Data collection	4. Synthesis/findings	Risk of bias
Signaling questions	1.1 Adhere to predefined objectives and eligibility criteria	2.1 Appropriate range of databases for published and unpublished reports	3.1 Minimizing error in data collection	4.1 Inclusion all studies that it should be	All the concerns in 4 domains addressed
	1.2 eligibility criteria appropriate	2.2 additional studies searched	3.2 characteristics of studies provided	4.2 all predefined analysis reported	Relevant questions considered
	1.3 eligibility criteria unambiguous	2.3 search strategy comprehensive	3.3 all relevant study results collected	4.3 the synthesis appropriate	Avoid emphasising significant results only
	1.4 restrictions appropriate (study characteristics)	2.4 Restrictions appropriate (pub date, language)	3.4 ROB assessed	4.4 between-study variation minimal	
	1.5 restrictions appropriate (sources of information)	2.5 minimizing error in collecting studies	3.5 minimizing error in ROB assessment	5.4 finding robust? Bias in primary studies minimal	
Study	Specification of study eligibility criteria	Methods used to identify and/or select studies	Methods used to collect data	The synthesis	ROB in the review
Darby et al. (2005) [4]	Low concern	Low concern	Low concern	Low concern	Low concern
Krewski et al. (2005) [5]	Low concern	Low concern	Low concern	Low concern	Low concern
Lubin et al. (2004) [6]	Low concern	Low concern	Low concern	Low concern	Low concern
Lorenzo-Gonzalez et al. (2019) [7]	Low concern	Low concern	Unclear concern[¶]	Low concern	Low concern
Summary for pooled collaborative studies: Three studies had low concern in all domains and one study had unclear concern in one domain. Overall ROB was judged to be low.					

ROB assessment[b] for Cohort study (Turner et al. 2011)			
#	Domain	Rating	Risk of bias
(I)	Bias in selection of participants into study		
	Selection of the exposed and non-exposed cohorts	1. Drawn from the same population	Low
(II)	Bias due to error in exposure measurement		
	Measurement of exposure	1. Objective measurements from pre-existing records or baseline physical or biological assessment blind to outcome status	Low
(III)	Bias due to error in outcome measurement		
a	Measurement of outcome	1. Outcome measurement unlikely to be influenced by exposure	Low
b	Was outcome of interest absent at the time to which the exposure refers?	1. Yes	Low
c	Was follow-up long enough for outcome to occur as a consequence of measured exposure?	1. Yes	Low
(IV)	Bias due to non-participation		
	Participation rate	1. Participation rate in exposed cohort is ≤10 percentage points different from non-exposed cohort OR exposed and non-exposed are from the same cohort	Low
(V)	Bias due to missing data		
a	Completeness of follow-up	2. There is a plausible estimate of 70-90% follow-up	Moderate
b	Accuracy of dates of outcome or censoring	1. Dates of outcome or censoring ascertained to within one year	Low
c	Difference in follow-up between exposed and non-exposed	1. Follow-up methods are the same and likely to achieve the same completeness of follow-up in exposed and unexposed participants	Low
d	Difference in missing data for exposure between those with or without the outcome	1. Difference in missing data for exposure < 10 percentage points	Low

(VI)	Bias due to confounding		
	Comparability of exposed and non-exposed cohorts with respect to potentially important confounding variables (prior specification of potentially important confounders)	1. Age and other potentially important confounders measured and controlled by design or in analysis	Low
(VII)	Analysis bias		
	Covariates are appropriately included in statistical analysis models	1. Variables measuring the same underlying concept or lying in the same causal pathway ARE NOT included together as covariates in statistical analysis models	Low

Summary for cohort study: All domains except one had low ROB, and one domain had moderate ROB.

Overall ROB was judged to be moderate.

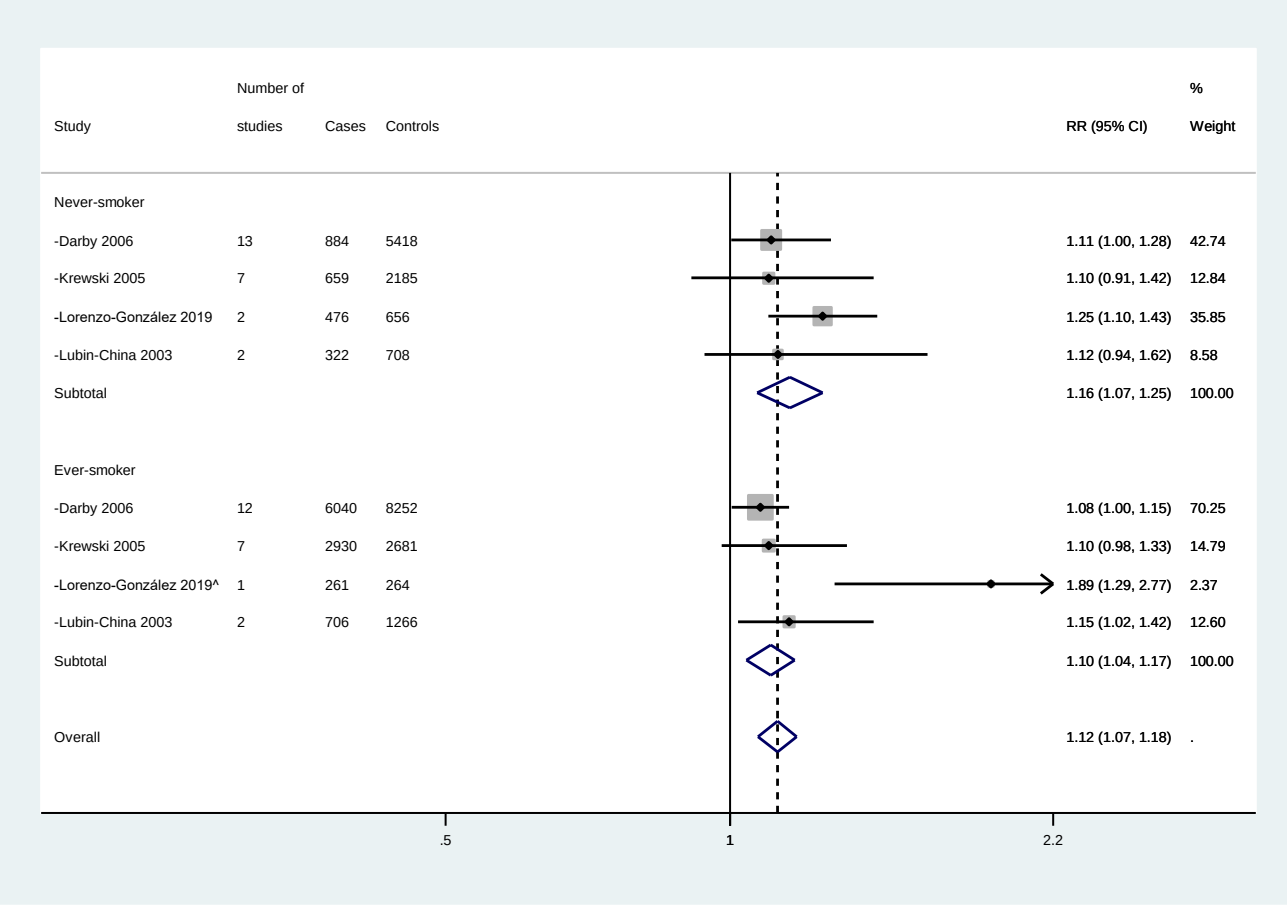
ROB assessment[b] for Case-control study (Wilcox et al. 2008)

#	Domain	Rating	Risk of bias
(I)	Definition and selection of cases and controls		
a	Definition of cases	1. Outcome precisely specified and with pathological or other objective confirmation	Low
b	Definition of controls	2. Self-report of no past history of outcome of interest OR insufficient information to tell	Moderate
c	Selection of cases and controls	1. Drawn from the same population	Low
(II)	Participation (response) rates		
a	Participation (response) rate of cases	2. ≥50 to <70% participation rate (≥60 to <80% response rate)	Moderate
b	Participation (response) rate of controls	1. ≥60% participation rate (≥70% response rate)	Low
c	Difference in participation rate (response rate) between cases and controls	1. Participation (response) rate in cases ≤10 percentage points different from controls	Low
(III)	Measurement of exposure		

a	1. Objective measurements from biological assessment blind to case or control status		Low
b	Was the same method used to measure exposure in cases and controls?	1. Yes	Low
(IV)	Temporality of exposure		
	3. Exposure does not precede onset of disease in cases OR insufficient information to tell		High
(V)	Missing exposure data		
	Difference in missing data of exposure between cases and controls	1. Difference in missing data of exposure < 10 percentage points	Low
(VI)	Control of confounding		
a	Comparability of cases and controls with respect to potentially important confounding variables (Requires prior specification of potentially important confounders)	2. Age and some but not all other potentially important confounders controlled by design or in analysis	Moderate
b	Matching variables are appropriately included in the analysis	1. When controls are frequency matched to cases, matching variables are controlled in the analysis	Low
c	Other covariates are appropriately included in the analysis	1. NO variable measuring the same underlying concept or lying in the same causal pathway as the exposure variable under study IS included as a covariate in the statistical analysis models	Low
(VII)	Conflict of interest		
	Conflict of interest	1. No conflict of interest declared	Low

Summary for case-control study: One domain had high ROB, three domains had moderate ROB, and others had low ROB. Overall ROB was judged to be high.

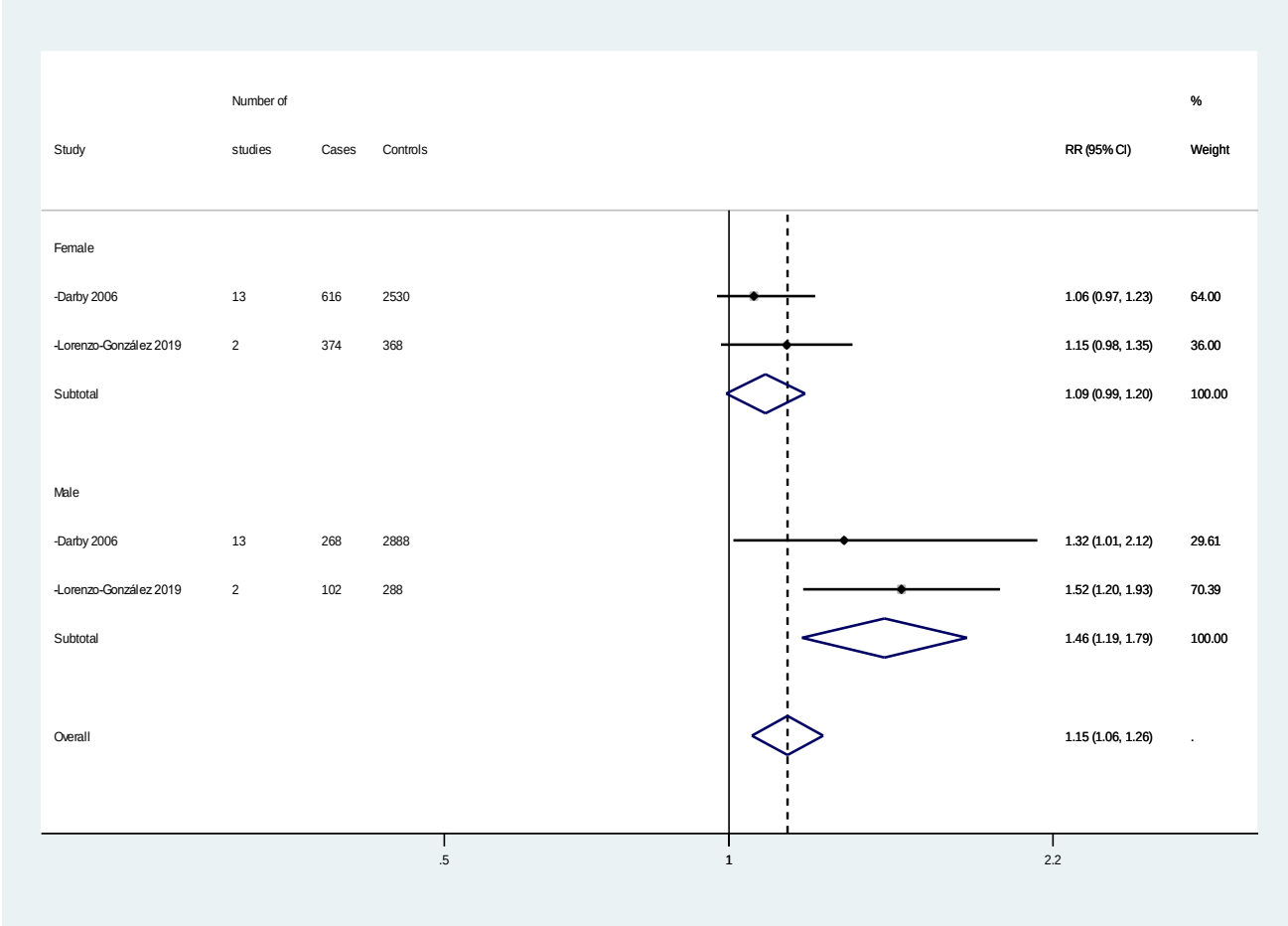
Supplementary Figure 1. Adjusted relative risk (aRR) and 95% confidence interval (CI) at 100 Bq/m^3 (radon exposure) for diagnosis of lung cancer



p=0.35 for test of difference between pooled aRRs for never-smokers and ever-smokers.

^ Barros-Dios et al. (2012) [8] in this pooled study contributed to the 'ever-smoker' meta-analysis.

Supplementary Figure 2. Adjusted relative risk (aRR) and 95% confidence interval (CI) at 100 Bq/m^3 (radon exposure) for diagnosis of lung cancer in never-smokers stratified by sex



p=0.011 for difference between male and female pooled aRRs.

References

1. Greenland S, Longnecker MP. Methods for trend estimation from summarized dose-response data, with applications to meta-analysis. Am J Epidemiol 1992;135(11):1301-9.
2. Orsini N, Bellocco R, Greenland S. Generalized least squares for trend estimation of summarized dose–response data. Stata J 2006;6(1):40-57.
3. Moher D, Shamseer L, Clarke M*, et al.* Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015;4:1.
4. Darby S, Hill D, Auvinen A*, et al.* Radon in homes and risk of lung cancer: collaborative analysis of individual data from 13 European case-control studies. BMJ 2005;330(7485):223.
5. Krewski D, Lubin JH, Zielinski JM*, et al.* Residential radon and risk of lung cancer: a combined analysis of 7 North American case-control studies. Epidemiology 2005;16(2):137-45.
6. Lubin JH, Wang ZY, Boice JD, Jr.*, et al.* Risk of lung cancer and residential radon in China: pooled results of two studies. Int J Cancer. 2004;109(1):132-7.
7. Lorenzo-Gonzalez M, Ruano-Ravina A, Torres-Duran M*, et al.* Lung cancer and residential radon in never-smokers: A pooling study in the Northwest of Spain. Environ Res 2019;172:713-718.
8. Barros-Dios JM, Ruano-Ravina A, Perez-Rios M*, et al.* Residential radon exposure, histologic types, and lung cancer risk. A case-control study in Galicia, Spain. Cancer Epidemiol Biomarkers Prev 2012;21(6):951-8.