BMJ Open

Dose-response relationship between cigarette smoking and site-specific cancer risk: protocol for a systematic review with an original design combining umbrella and traditional reviews

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-018930
Article Type:	Protocol
Date Submitted by the Author:	01-Aug-2017
Complete List of Authors:	Lugo, Alessandra; IRCCS — Istituto di Ricerche Farmacologiche "Mario Negri", Milan, Italy, Department of Environmental Health Sciences Bosetti, Cristina; Istituto di Ricerche Farmacologiche "Mario Negri", Department of Oncology Peveri, Giulia; IRCCS – Istituto di Ricerche Farmacologiche "Mario Negri", Department of Environmental Health Sciences Rota, Matteo; Università degli Studi di Milano, Department of Biomedical and Clinical Sciences Bagnardi, Vincenzo; Università degli Studi di Milano-Bicocca, Department of Statistics and Quantitative Methods Gallus, Silvano; IRCCS – Istituto di Ricerche Farmacologiche "Mario Negri", Department of Environmental Health Sciences
Primary Subject Heading:	Smoking and tobacco
Secondary Subject Heading:	Epidemiology, Oncology, Public health, Research methods
Keywords:	Cancer, Risk, Cigarette smoking, Dose-response relationship, Systematic review, Meta-analysis

Dose-response relationship between cigarette smoking and site-specific cancer risk: protocol for a systematic review with an original design combining umbrella and traditional reviews

Alessandra LUGO[1], Cristina BOSETTI[2], Giulia PEVERI[1],

Matteo ROTA[3], Vincenzo BAGNARDI[4], Silvano GALLUS[1]

[1] Department of Environmental Health Sciences, IRCCS – Istituto di Ricerche Farmacologiche "Mario Negri", Milan, Italy

[2] Department of Oncology, IRCCS – Istituto di Ricerche Farmacologiche "Mario Negri", Milan, Italy

[3] Department of Biomedical and Clinical Sciences, Università degli Studi di Milano, Milan, Italy

[4] Department of Statistics and Quantitative Methods, Università degli Studi di Milano-Bicocca, Milan, Italy

Corresponding author

Alessandra Lugo, PhD

Department of Environmental Health Sciences, Laboratory of Lifestyle Epidemiology,

IRCCS – Istituto di Ricerche Farmacologiche "Mario Negri", Milan, Italy

Via G. La Masa 19, 20156 Milan, Italy

tel: +390239014653 – fax: +390233200231

e-mail: alessandra.lugo@marionegri.it

Word count: 2838

ABSTRACT

Introduction. Only a limited number of meta-analyses providing risk curve functions of dose-response relationships between various smoking variables and cancer-specific risk are available.

Methods and analysis. To identify all relevant original publications on the issue, we will conduct a series of comprehensive systematic reviews based on three subsequent literature searches: 1) an umbrella review, to identify published meta-analyses, pooled-analyses, and systematic reviews on the association between cigarette smoking and cancer risk; 2) for each cancer site, an updated systematic review of original publications on the association between cigarette smoking and site-specific cancer risk, starting from the last available comprehensive review identified through the umbrella review; 3) a review of all original articles on the association between cigarette smoking and site-specific cancer risk included in the publications identified through the umbrella review or through the update of the reviews. The primary outcomes of interest will be: i) the excess incidence/mortality of various cancers for smokers compared to never smokers; ii) the dose-response curves describing the association between smoking intensity, duration, and time since stopping and incidence/mortality for various cancers. For each cancer site, we will perform a meta-analysis by pooling all the study-specific estimates for smoking status. We will also estimate the dose-response curves for other smoking variables through random-effects meta-regression models based on a nonlinear dose-response relationship framework.

Ethics and dissemination. Ethics approval is not required for this study. The main results of this study will be published in peer-reviewed journals and will also be included in a publicly available website. We will provide therefore the most complete and updated estimates on the association between various measures of cigarette smoking and site-specific cancer risk. This will allow us to obtain precise estimates on the cancer burden attributable to cigarette smoking.

Systematic review registration. This protocol was registered in PROSPERO (CRD42017063991).

Keywords: cancer; risk; cigarette smoking; dose-response relationship; systematic review; meta-analysis.

STRENGTHS AND LIMITATIONS OF THE STUDY

- We will use an original and innovative approach combining an umbrella review and traditional systematic reviews.
- Although we will not conduct a systematic review of the entire scientific literature on the issue, our study is highly feasible, since we consider published reviews.
- We will carry out dose-response analyses using two-steps random-effects meta-regression models in order to examine potential nonlinear relationships between smoking-related variables and the risk of cancer.
- We will provide the most complete and updated estimates on the association between cigarette smoking and site-specific cancer risk.

INTRODUCTION

In 1950, Ernst Wynder and Evards Graham[1] and Richard Doll and Bradford Hill[2] first reported an association between tobacco smoking and lung cancer risk. Over the subsequent 65 years, thousands of studies systematically confirmed this association and found that tobacco smoking also increases the risk of several other neoplasms.[3] In 2004, tobacco smoking has been classified as carcinogenic to humans (Group 1) by the International Agency for Research on Cancer (IARC)[3], which provided evidence on the causal relationship between cigarette smoking and cancer of the lung, oral cavity, nasal cavity and paranasal sinuses, nasopharynx, oropharynx, hypopharynx, larynx, oesophagus, stomach, pancreas, liver, kidney, ureter, urinary bladder, cervix and myeloid leukaemia[3]. The results on the association between (cigarette) smoking and other cancer sites, including cancers of the breast and endometrium, remain conflicting.[3]

Since late 1980s, publications reporting results from systematic reviews and meta-analyses, the large majority of which were unnecessary and misleading, rapidly increased.[4] Although the effects of cigarette smoking on cancer incidence and/or mortality have been largely investigated, only a limited number of systematic reviews/meta-analyses are available on the quantification of the dose-response relationship between selected smoking variables, including smoking intensity, duration, pack-years and time since quitting, and the risk of cancer. More importantly, just a few meta-analyses, if any, provided the cancer-specific risk curve functions of the dose-response relationships. Dose-response data, however, are crucial to provide reliable and accurate estimates of the cancer burden due to smoking, both at individual - absolute cancer incidence/mortality obtained through lung cancer risk-assessment models and risk charts[5 6] - and population level - smoking attributable deaths.[7 8] Currently, the methods developed to quantify the cancer burden due to smoking use cancer-specific estimates of relative risks (RR) according to tobacco smoking derived from a few large cohorts, mainly from the USA.[9] The use of these cohorts to derive RRs may lead to validity problems when applying these estimates to other populations with different smoking patterns.[7 10] Furthermore, these RRs are often estimated after allowance for a limited number of

socio-demographic characteristics, excluding potentially important confounding variables, such as alcohol drinking.

Using an original and innovative approach, which combines an umbrella review and traditional systematic reviews, we aim at providing a comprehensive and updated picture of the association between various smoking-related variables and site-specific cancer risk. We will be able to estimate the most robust cancer-specific RRs, obtained from the existing scientific literature, possibly derived after adjustment for relevant covariates. Moreover, for each cancer site, we will be able to provide the risk curves which best describe the dose-response effect of smoking intensity, smoking duration, pack-years, and time since stopping smoking on cancer risk.

METHODS

The present cancer-specific systematic reviews/meta-analyses will be based on the following three subsequent literature searches on the association between cigarette smoking and cancer risk:

1. Umbrella review: a systematic review to identify published meta-analyses, pooled-analyses, and systematic reviews providing data on the association between cigarette smoking and cancer risk;

2. Update of available cancer-specific reviews: for each cancer site, the conduction of a systematic review of studies providing original data (including pooled-analyses on individual participants data) on the association between cigarette smoking and cancer-specific risk, starting from the last available comprehensive review publication identified through the umbrella review (point 1);

3. Review of all original publications: a review of all original articles on the association between cigarette smoking and site-specific cancer risk included in the cancer-specific review publications identified through the umbrella review (point 1) or identified through the update of the available reviews (point 2).

The design of the present systematic reviews was developed following the Preferred Reporting Items for Systematic review and Meta-Analyses (PRISMA) guidelines[11] and its extension for

protocols (PRISMA-P).[12 13] This protocol was registered in the International Prospective Register of Systematic Reviews (PROSPERO) on 4 May 2017 (registration number: CRD42017063991).

1. Umbrella review

Search strategy

We will conduct a systematic literature search to identify all published meta-analyses, pooled-analyses, and systematic reviews providing data on the association between cigarette smoking and the risk of various cancers. Literature search strategy will include combinations of Medical Subjects Headings (MeSH) and text words related to *cancer* and *tobacco* or *smoking*, and will be restricted to the following publication types: meta-analyses, pooled-analyses, and systematic reviews. No restriction on cancer site or on publication date will be applied. The following databases will be used: MEDLINE, Embase, ISI Web of Science (WoS), and Cochrane Database of Systematic Reviews (CDSR). The search strings to be used in various databases are reported in **Appendix 1**. To ensure literature saturation, reference lists of selected relevant publications identified through the search will also be checked. Besides the publications found through the databases searches, we will also consider the reviews of the literature on tobacco smoking provided within the IARC monographs vol. 83 [3] and vol. 100E [14] and the Surgeon General Report [15], three reports of known importance providing data on the association between tobacco smoking and various cancer sites.

Eligibility criteria

Study design

We will include meta-analyses, pooled-analyses and systematic reviews of observational studies providing measures of RRs between cigarette smoking and cancer risk. Original observational studies (e.g., case-control, cohort or cross-sectional studies) will be excluded. Reports, letters to the editor, book chapters, conference proceedings, dissertations, and theses will be not considered.

Conditions

We will considers publications providing data on the following 28 (namely all) malignant neoplasms: cancer of lip, oral cavity and pharynx (ICD-10: C00-C14), nasopharynx (C11), oesophagus (C15), stomach (C16), colon (C18), rectum and anus (C19-C21), liver (C22), gallbladder (C23-C24), pancreas (C25), larynx (C32), lung trachea and bronchus (C33-C34), bone (C40-C41), melanoma of skin (C43), mesothelioma (C45), breast (C50), cervix uteri (C53), corpus uteri (C54), ovary (C56), prostate (C61), testis (C62), kidney (C64), bladder (C67), brain, central nervous system (C70-C72), thyroid (C73), Hodgkin lymphoma (C81), non-Hodgkin lymphoma (C82-C86, C96), multiple myeloma (C88-C90) and leukaemia (C91-C95). We will also consider review publications on groups of cancers (e.g., head and neck, upper aero-digestive tract, or intestinal cancers). Studies specifically based on benign neoplasms, such as colorectal polyps, acoustic neuroma and meningioma, and neuroendocrine tumours will be excluded.

Participants

We will include review publications providing data on humans, in the general population. We will therefore exclude review publications based on patients with cancer or other diseases (i.e., reporting data on the effect of smoking on the prognosis of the disease), or on subgroups of the population with selected lifestyle habits or other characteristics (e.g., populations limited to alcohol drinkers or tobacco smokers). No restriction will be applied according to age of participants at cancer incidence or mortality given that practically all studies providing data on the association between cigarette smoking and cancer risk are based on adults, only.

Exposures

We will include all review publications providing data on the use of cigarettes in the general population. Publications focused on the use of tobacco products other than cigarettes (e.g., pipe, smokeless tobacco, cigar, water pipe, electronic cigarettes) or on the exposure to second-hand smoke will be excluded.

Outcomes

The primary outcomes of interest will be: i) the excess incidence and/or mortality of various cancers in current/ever smokers compared to non/never smokers; ii) the dose-response curves describing the association between cigarette smoking duration, intensity and time since stopping and incidence and/or mortality for various cancers.

Languages

We will include only articles published in English language.

Study selection

All the review publications found in various electronic databases through the above mentioned search strategy will be uploaded in an EndNote library (EN1), and duplicate records will be deleted. Titles and/or abstracts of the meta-analyses, pooled-analyses and systematic reviews will be screened independently by two reviewers (AL, GP) to exclude publications which will not meet the eligibility criteria. The full text of the remaining review publications will be retrieved and independently assessed for eligibility by the two reviewers. Discrepancies between the two reviewers will be discussed and solved. In case of disagreement, a third reviewer (SG) will help to find a final decision. Data from other available reports will also be integrated in the same EN1 library.

Quality assessment

Assessment of the quality of various review publications is out of the scope of the present systematic review. Thus, no quality score will be assigned to the publications. No review publication will be excluded *a priori* for weakness of design or data quality.

Data extraction and management

A standardised form in Microsoft (MS) Excel will be used to extract data from each identified review publication. Relevant information will include: first author, year of publication, type of study (i.e., meta-analysis, pooled-analysis, or systematic review), cancer site(s) and/or subsite(s), endpoint (i.e., incidence and/or mortality), and other information about the methodology of studies included in the reviews (e.g., study design, country(ies) of study conduction, number of studies considered in, and overall population size). Data will be extracted from each included meta-analysis, pooled analysis or systematic review by one reviewer and verified by a second reviewer. Any disagreement between the two reviewers will be solved by consensus. Otherwise, a third reviewer will help to find a final decision. After data extraction, publications will be grouped according to the considered cancer site.

Preliminary umbrella review

We already conducted a preliminary umbrella review through which we identified the meta-analyses, pooled-analyses and systematic reviews on the association between smoking and cancer risk published before 28th April 2017. Within the comprehensive literature search, we found a total of 1430 publications from the four considered databases (726 from Medline, 316 from Embase, 376 from ISI WoS, and 12 from CDSR; **Figure 1**). After the exclusion of duplicates (n=542) and ineligible papers (n=716), and after the inclusion of three important reports[3 14 15], we obtained a total of 175 relevant publications (i.e., 107 meta-analyses, 52 pooled-analyses, and 16 systematic reviews) on the association between smoking and risk of various neoplasms.

2. Update of the available cancer-specific reviews

Search strategy

For each of the 28 cancer sites previously described in the umbrella review, we will identify the last available comprehensive systematic review or meta-analysis on the association with cigarette

smoking. We will then update the identified cancer-specific review through the conduction of a systematic literature search on all observational studies (e.g., case-control, cohort, and nested case-control studies) providing original data on the association between cigarette smoking and site-specific cancer risk, and published after the year of publication of the most recent article included in the last comprehensive review. Only studies published in English language will be considered.

The literature search strategy will be conducted in MEDLINE and Embase, and will include combinations of MeSH terms and text words related to *site-specific cancer* and *tobacco* or *smoking* (**Appendix 2**).

Eligibility criteria

Study design

We will include original observational studies (e.g., case-control, cohort, nested case-control studies, or pooled-analysis of individual participant data) providing measures of RRs of the association between cigarette smoking and site-specific cancer risk. Reports, book chapters, conference proceedings, dissertations, and theses will not be considered. We will exclude case-control studies using patients with cancer or other chronic diseases as comparison group.

Comparator

Never smokers are our comparators. We will in fact consider, as the measure of association, the RRs of smoking variables compared to never smokers.

Eligibility criteria for *conditions*, *participants*, *exposures*, and *outcomes* are those reported also for the umbrella review.

Study selection

For each cancer site, we will upload all the original publications found using the above mentioned search strategy in cancer-specific EndNote libraries (EN2_1-EN2_28) and we will delete the identified duplicates. Titles and/or abstracts of original articles will be screened independently by

two reviewers (AL, GP) to exclude publications that do not meet the inclusion criteria outlined above. The full text of the remaining original publication will be retrieved and independently assessed for eligibility by the two reviewers. Discrepancies on the assessment between the two reviewers will be discussed and solved. In case of disagreement, a third reviewer (SG) will help to find a final decision. Being out of the scope of the present systematic review, no quality assessment will be considered. No original publication will be excluded *a priori* for weakness of design or data quality.

3. Review of all original publications

For each cancer site, we will upload in an EndNote library (EN3_1-EN3_28) all the original publications obtained from the 28 cancer-specific reviews identified in the umbrella review (point 1). In the same EndNote libraries, we will add the original publications obtained from the corresponding updates of the reviews (point 2), and duplicate publications will be deleted. The full text of all the retained original publications will be retrieved. Non-English reports, unpublished studies, conference proceedings, dissertations and theses will be excluded.

Figure 2 shows the flowchart we will consider for each of the 28 cancer-specific reviews. For each cancer site, original publications from both the umbrella review and the update of the reviews will be considered.

Data extraction and management

For each cancer site, two standardised forms in MS Excel will be used to collect relevant information on the study design and the risk estimates from the original publications. A first form will be used to extract data related to the study design including: first author, year of publication, journal, country, study name, period and design of study, outcome and sample size. In the second form we will collect the exposure categories (i.e., smoking status, intensity, duration, pack-years,

age at starting, time since stopping), and corresponding RR estimates (or other estimates, such as odds ratios and hazard ratios) and 95% confidence intervals (CIs). The number of cases in each exposure category and covariates used in the model will be also collected. When the results of the same study have been published in more than one original publication, only data from the most recent and/or complete article will be retained and reported in the second Excel form.

Data analysis

For each cancer site, we will pool all the RRs or other risk estimates (e.g., hazard ratios and odds ratios) for current, former or ever compared to never smokers. Because cancer is a relatively rare outcome, we assume that ORs, risk ratios and rate ratios are all comparable estimates of the RR. Heterogeneity between studies will be assessed using the Cochran Q test and the I-squared statistics, i.e. the proportion of total variation contributed by between-study heterogeneity.[16] As we anticipated between study heterogeneity, we will present pooled RRs from random-effects models using the DerSimonian and Laird moment estimator of the between-study variance component.[17] However, if no heterogeneity between-study is detected, pooled estimates from the random-effects model will correspond to those deriving from the fixed-effect model.

When RRs are reported separately for current and former compared to never smokers, we will combine them into a single estimate for ever smokers using the method for pooling non-independent estimates described by Hamling et al.[18] This method uses the number of subjects exposed to different levels of smoking and non-exposed subjects, and the corresponding risk estimates, to derive a set of pseudo-numbers of cases and controls/subjects at risk, by taking into account the correlation between the original estimates due to the common reference group. The obtained pseudo-numbers will be used to compute the new adjusted RR. This methodology will also be used to convert RR estimates when the reference category considered in the study is not represented by never smokers.

For smoking intensity, duration and time since stopping smoking, we will compute pooled RRs according to various categories of the considered smoking-related variables (e.g, low, intermediate, high cigarette consumption). Moreover, we will carry out dose-response analyses using two-steps random-effects meta-regression models in order to examine potential nonlinear relationships between those variables and the risk of cancer. In particular, we will consider a method providing the best fitting two-term fractional polynomial model[19], and a method modelling the considered smoking-related variables using restricted cubic splines.[20]

If the necessary data are available, we will consider to further conduct separate analyses by sex, study period, geographic area, and income (defined on the basis of per-capita Gross Domestic Product).

Analyses will be conducted using SAS software version 9.4 (SAS Institute Inc., Cary, NC, USA), and R software version 3.3.0 (R Development Core Team, 2008), in particular *meta* and *dosresmeta* packages.

ETHICS AND DISSEMINATION

Ethics

This review does not require approval from an ethics committee, since no individual-level patients' data will be collected.

Implications and dissemination

Through these systematic reviews and meta-analyses, we will provide the most complete and updated estimates on the association between cigarette smoking and site-specific cancer risk. These estimates will be used to quantify the cancer burden due to cigarette smoking at both individual and population level. This information is essential to guide policy decisions to control tobacco smoking and improve cancer prevention.

Given the relevance and originality of this project, we plan to publish results from the meta-analyses in peer-reviewed journals, considering either single cancer sites or various apparatus or tracts. A final publication will provide the summary results for all cancers. We will also include the main results of our systematic reviews and meta-analyses in a publicly available website. Readers will have the possibility to contact us to communicate possible lacks or updates.

Contributions

SG and AL had the original idea of the work. AL, GP, and SG structured the study protocol, designed the search strategies, and conducted the umbrella review. MR, VB, and CB provided statistical and epidemiological supervision. All authors critically revised the draft of the manuscript and approved its final version.

Funding

This work was partially supported by the Italian League Against Cancer (Milan). AL was supported by a fellowship from the Italian Association for Cancer Research (AIRC). MR was supported by a fellowship from the Italian Foundation for Cancer Research (FIRC).

Competing interests

None declared.

REFERENCES

1. Wynder EL, Graham EA. Tobacco smoking as a possible etiologic factor in bronchiogenic carcinoma; a study of 684 proved cases. *J Am Med Assoc* 1950;143(4):329-36.
2. Doll R, Hill AB. Smoking and carcinoma of the lung; preliminary report. *Br Med J* 1950;2(4682):739-48.
3. IARC. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Volume 83. Tobacco Smoke and Involuntary Smoking. Lyon, France: International Agency for Research on Cancer 2004.
4. Ioannidis JP. The Mass Production of Redundant, Misleading, and Conflicted Systematic Reviews and Meta-analyses. *Milbank Q* 2016;94(3):485-514.
5. Spitz MR, Hong WK, Amos CI, *et al.* A risk model for prediction of lung cancer. *J Natl Cancer Inst* 2007;99(9):715-26.
6. Tammemagi MC, Katki HA, Hocking WG, *et al.* Selection criteria for lung-cancer screening. *N Engl J Med* 2013;368(8):728-36.
7. Perez-Rios M, Montes A. Methodologies used to estimate tobacco-attributable mortality: a review. *BMC Public Health* 2008;8:22.
8. Tachfouti N, Raherison C, Obtel M, *et al.* Mortality attributable to tobacco: review of different methods. *Arch Public Health* 2014;72(1):22.
9. ACS. Cancer Prevention Study II (CPS II). Available at: http://www.cancer.org/research/researchtopreventcancer/currentcancerpreventionstudies/cancer-prevention-study 2016
10. Gallus S, Muttarak R, Martinez-Sanchez JM, *et al.* Smoking prevalence and smoking attributable mortality in Italy, 2010. *Prev Med* 2011;52(6):434-8.
11. Moher D, Liberati A, Tetzlaff J, *et al.* Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;339:b2535.
12. 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.
13. Shamseer L, Moher D, Clarke M, *et al.* Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ* 2015;349:g7647.
14. IARC. IARC monograph on the evaluation of carcinogenic risks to humans. Volume 100E. A review of human carcinogens - Personal habits and indoor combustions. Available at: http://monographs.iarc.fr/ENG/Monographs/vol100E/mono100E.pdf. Lyon, France 2012.

15. U.S. Department of Health and Human Services. The Health Consequences of Smoking - 50 Years of Progress A Report of the Surgeon General. Available at: https://www.surgeongeneral.gov/library/reports/50-years-of-progress/full-report.pdf. Atlanta, GA 2014.
16. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002;21(11):1539-58.
17. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7(3):177-88.
18. Hamling J, Lee P, Weitkunat R*, et al.* Facilitating meta-analyses by deriving relative effect and precision estimates for alternative comparisons from a set of estimates presented by exposure level or disease category. *Stat Med* 2008;27(7):954-70.
19. Rota M, Bellocco R, Scotti L*, et al.* Random-effects meta-regression models for studying nonlinear dose-response relationship, with an application to alcohol and esophageal squamous cell carcinoma. *Stat Med* 2010;29(26):2679-87.
20. Orsini N, Li R, Wolk A*, et al.* Meta-analysis for linear and nonlinear dose-response relations: examples, an evaluation of approximations, and software. *Am J Epidemiol* 2012;175(1):66-73.

FIGURES LEGEND

Figure 1. Flowchart for the selection of papers in the umbrella review.

ISI WoS: ISI Web of Science

CDSR: Cochrane Database of Systematic Reviews

Figure 2. Flowchart for each of the 28 cancer-specific reviews.

Appendix 1. Literature search strings for the umbrella review used in various databases.

MEDLINE

(neoplasms [MeSH Terms] OR cancer [title] OR cancers [title] OR neoplasm [title] OR neoplasms [title] OR leukaemia [title] OR leukemia [title] OR leukaemias [title] OR leukemias [title] OR carcinoma[title] OR maligna*[title]) AND (meta-analysis OR "meta analysis" OR pooled-analysis OR "pooled analysis" OR "systematic review"[tiab] OR "systematic revision"[tiab]) AND (cigarette [title] OR cigarettes [title] OR tobacco [title] OR smoking [title] OR smokers [title] OR smoking [MeSH Terms])

Embase

(cancer:ti OR cancers:ti OR neoplasm:ti OR neoplasms:ti OR leukaemia:ti OR leukemia:ti OR leukaemias:ti OR leukemias:ti OR carcinoma:ti OR maligna*:ti) AND ('meta analysis' OR 'pooled analysis' OR 'meta analysis':it) AND (cigarette:ti OR cigarettes:ti OR tobacco:ti OR smoking:ti OR smokers:ti)

ISI Web of Science

TITLE: ((cancer OR cancers OR neoplasm OR neoplasms OR leukaemia OR leukemia OR leukaemias OR leukemias OR carcinoma OR maligna*) AND (cigarette OR cigarettes OR tobacco OR smoking OR smokers)) AND TOPIC: ((meta-analysis OR "meta analysis" OR pooled-analysis OR "pooled analysis" OR "systematic review" OR "systematic revision")) NOT DOCUMENT TYPES: (Letter OR Meeting Abstract OR Meeting Summary)

Cochrane Database of Systematic Reviews

(cancer:ti or cancers:ti or neoplasm:ti or neoplasms:ti or leukaemia:ti or leukemia:ti or leukaemias:ti or leukemias:ti or carcinoma:ti or maligna*:ti) and (meta-analysis or "meta analysis" or pooled-

analysis or "pooled analysis") and (cigarette:ti or cigarettes:ti or tobacco:ti or smoking:ti or smokers:ti)

Appendix 2. Literature search strings for the update of the available reviews used in various databases.

MEDLINE

((((*cancer site*) AND (cancer OR neoplasm OR carcinoma OR Neoplasms [MeSH Terms]) AND (cigarette OR cigarettes OR tobacco OR smoking OR smokers OR smoking [MeSH Terms]))) AND English[Language]) AND ("*last review year*"[Date - Publication]: "2017"[Date - Publication])

Embase

cigarette:ti OR cigarettes:ti OR tobacco:ti OR smoking:ti OR smokers:ti AND (*cancer site*:ab,ti) AND (cancer:ab,ti OR neoplasm:ab,ti OR carcinoma:ab,ti) AND (article:it OR review:it) AND [english]/lim AND [*last review year*-2017]/py NOT [medline]/lim

Cancer site stands for the terms used to identify the site-specific cancer (e.g., for liver cancer, the combination of the following terms will be considered: liver, hepatocellular, hepatic).

Last review year refers to the year of publication of the most recent article included in the last comprehensive review.

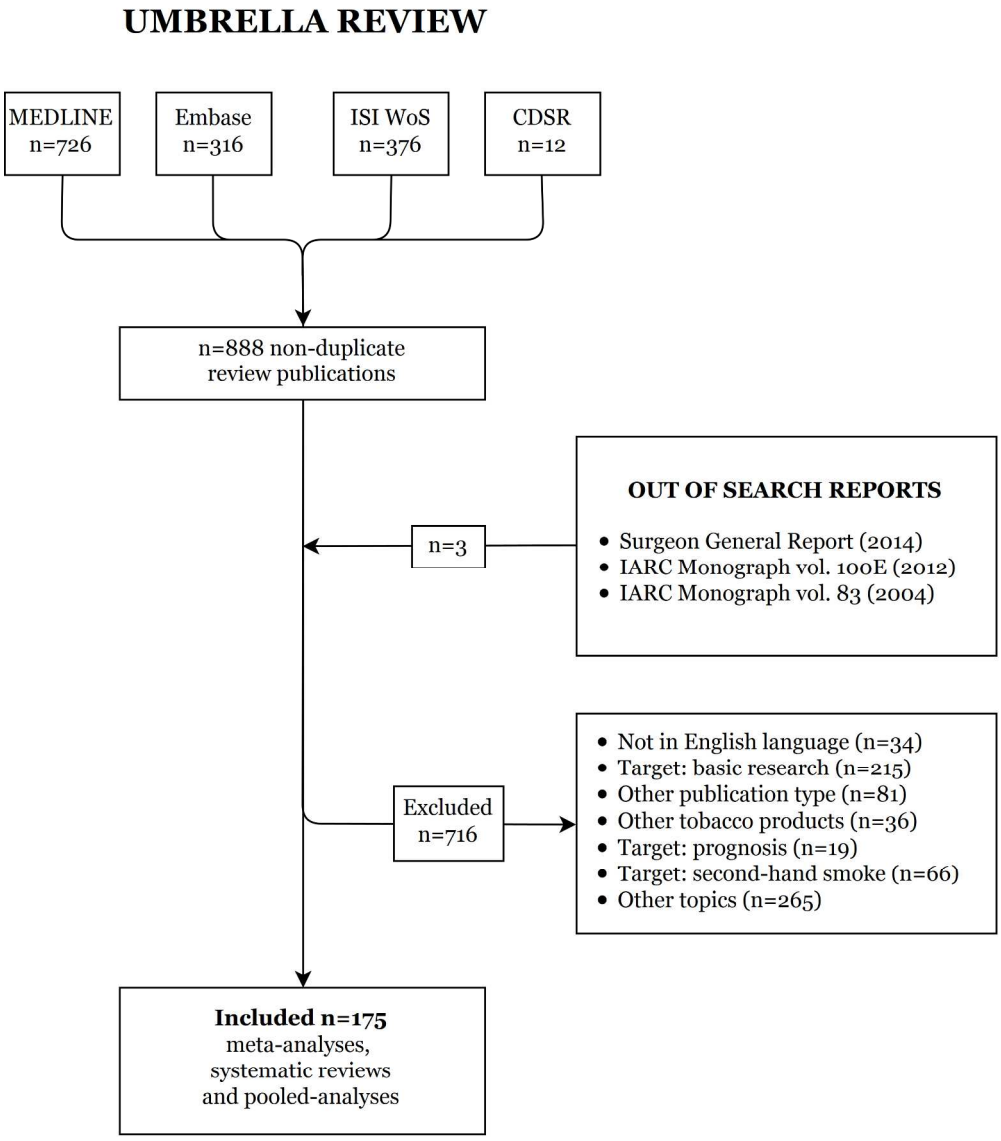

Figure 1. Flowchart for the selection of papers in the umbrella review.

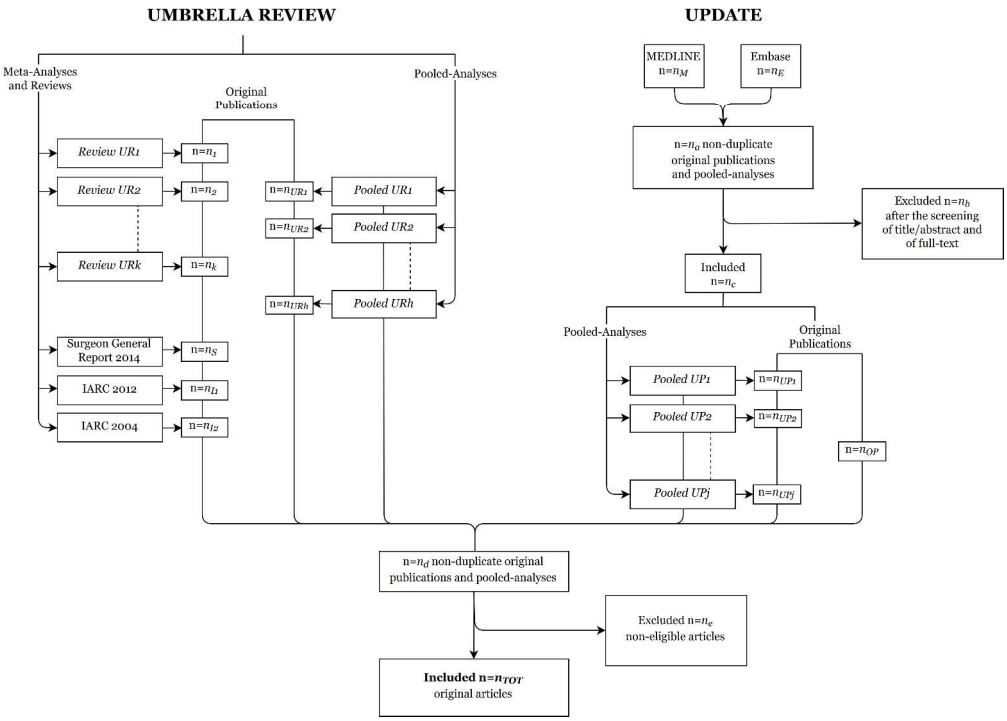

Figure 2. Flowchart for each of the 28 cancer-specific reviews.

PRISMA-P (Preferred Reporting Items for Systematic review and Meta-Analysis Protocols) 2015 checklist: recommended items to address in a systematic review protocol*

Section and topic	Item No	Checklist item	Page N°
ADMINISTRATIVE INFORMATION			
Title:			
Identification	1a	Identify the report as a protocol of a systematic review	1
Update	1b	If the protocol is for an update of a previous systematic review, identify as such	NA
Registration	2	If registered, provide the name of the registry (such as PROSPERO) and registration number	2
Authors:			
Contact	3a	Provide name, institutional affiliation, e-mail address of all protocol authors; provide physical mailing address of corresponding author	1
Contributions	3b	Describe contributions of protocol authors and identify the guarantor of the review	15
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	NA
Support:			
Sources	5a	Indicate sources of financial or other support for the review	15
Sponsor	5b	Provide name for the review funder and/or sponsor	NA
Role of sponsor or funder	5c	Describe roles of funder(s), sponsor(s), and/or institution(s), if any, in developing the protocol	NA
INTRODUCTION			
Rationale	6	Describe the rationale for the review in the context of what is already known	5
Objectives	7	Provide an explicit statement of the question(s) the review will address with reference to participants, interventions, comparators, and outcomes (PICO)	6
METHODS			
Eligibility criteria	8	Specify the study characteristics (such as PICO, study design, setting, time frame) and report characteristics (such as years considered, language, publication status) to be used as criteria for eligibility for the review	7-9, 11
Information sources	9	Describe all intended information sources (such as electronic databases, contact with study authors, trial registers or other grey literature sources) with planned dates of coverage	7,10

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	19-21
Study records:			
Data management	11a	Describe the mechanism(s) that will be used to manage records and data throughout the review	9-12
Selection process	11b	State the process that will be used for selecting studies (such as two independent reviewers) through each phase of the review (that is, screening, eligibility and inclusion in meta-analysis)	9-12
Data collection process	11c	Describe planned method of extracting data from reports (such as piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators	9-12
Data items	12	List and define all variables for which data will be sought (such as PICO items, funding sources), any pre-planned data assumptions and simplifications	9-12
Outcomes and prioritization	13	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale	9
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	NA
Data synthesis	15a	Describe criteria under which study data will be quantitatively synthesised	13-14
	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 (such as I^2, Kendall's τ)	13-14
	15c	Describe any proposed additional analyses (such as sensitivity or subgroup analyses, meta-regression)	13-14
	15d	If quantitative synthesis is not appropriate, describe the type of summary planned	NA
Meta-bias(es)	16	Specify any planned assessment of meta-bias(es) (such as publication bias across studies, selective reporting within studies)	13-14
Confidence in cumulative evidence	17	Describe how the strength of the body of evidence will be assessed (such as GRADE)	NA

*** It is strongly recommended that this checklist be read in conjunction with the PRISMA-P Explanation and Elaboration (cite when available) for important clarification on the items. Amendments to a review protocol should be tracked and dated. The copyright for PRISMA-P (including checklist) is held by the PRISMA-P Group and is distributed under a Creative Commons Attribution Licence 4.0.**

From: Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart L, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ. 2015 Jan 2;349(jan02 1):g7647.

BMJ Open

Dose-response relationship between cigarette smoking and site-specific cancer risk: protocol for a systematic review with an original design combining umbrella and traditional reviews

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-018930.R1
Article Type:	Protocol
Date Submitted by the Author:	01-Sep-2017
Complete List of Authors:	Lugo, Alessandra; IRCCS — Istituto di Ricerche Farmacologiche "Mario Negri", Milan, Italy, Department of Environmental Health Sciences Bosetti, Cristina; Istituto di Ricerche Farmacologiche "Mario Negri", Department of Oncology Peveri, Giulia; IRCCS – Istituto di Ricerche Farmacologiche "Mario Negri", Department of Environmental Health Sciences Rota, Matteo; Università degli Studi di Milano, Department of Biomedical and Clinical Sciences Bagnardi, Vincenzo; Università degli Studi di Milano-Bicocca, Department of Statistics and Quantitative Methods Gallus, Silvano; IRCCS – Istituto di Ricerche Farmacologiche "Mario Negri", Department of Environmental Health Sciences
Primary Subject Heading:	Smoking and tobacco
Secondary Subject Heading:	Epidemiology, Oncology, Public health, Research methods
Keywords:	Cancer, Risk, Cigarette smoking, Dose-response relationship, Systematic review, Meta-analysis

Dose-response relationship between cigarette smoking and site-specific cancer risk: protocol for a systematic review with an original design combining umbrella and traditional reviews

Alessandra LUGO[1], Cristina BOSETTI[2], Giulia PEVERI[1],

Matteo ROTA[3], Vincenzo BAGNARDI[4], Silvano GALLUS[1]

[1] Department of Environmental Health Sciences, IRCCS – Istituto di Ricerche Farmacologiche "Mario Negri", Milan, Italy

[2] Department of Oncology, IRCCS – Istituto di Ricerche Farmacologiche "Mario Negri", Milan, Italy

[3] Department of Biomedical and Clinical Sciences, Università degli Studi di Milano, Milan, Italy

[4] Department of Statistics and Quantitative Methods, Università degli Studi di Milano-Bicocca, Milan, Italy

Corresponding author

Alessandra Lugo, PhD

Department of Environmental Health Sciences, Laboratory of Lifestyle Epidemiology,

IRCCS – Istituto di Ricerche Farmacologiche "Mario Negri", Milan, Italy

Via G. La Masa 19, 20156 Milan, Italy

tel: +390239014653 – fax: +390233200231

e-mail: alessandra.lugo@marionegri.it

Word count: 2912

ABSTRACT

Introduction. Only a limited number of meta-analyses providing risk curve functions of dose-response relationships between various smoking variables and cancer-specific risk are available.

Methods and analysis. To identify all relevant original publications on the issue, we will conduct a series of comprehensive systematic reviews based on three subsequent literature searches: 1) an umbrella review, to identify meta-analyses, pooled-analyses, and systematic reviews published before 28th April 2017 on the association between cigarette smoking and the risk of 28 (namely all) malignant neoplasms; 2) for each cancer site, an updated review of original publications on the association between cigarette smoking and cancer risk, starting from the last available comprehensive review identified through the umbrella review; 3) a review of all original articles on the association between cigarette smoking and site-specific cancer risk included in the publications identified through the umbrella and the updated reviews. The primary outcomes of interest will be: i) the excess incidence/mortality of various cancers for smokers compared to never smokers; ii) the dose-response curves describing the association between smoking intensity, duration, and time since stopping and incidence/mortality for various cancers. For each cancer site, we will perform a meta-analysis by pooling study-specific estimates for smoking status. We will also estimate the dose-response curves for other smoking variables through random-effects meta-regression models based on a nonlinear dose-response relationship framework.

Ethics and dissemination. Ethics approval is not required for this study. Main results will be published in peer-reviewed journals and will also be included in a publicly available website. We will provide therefore the most complete and updated estimates on the association between various measures of cigarette smoking and site-specific cancer risk. This will allow us to obtain precise estimates on the cancer burden attributable to cigarette smoking.

Systematic review registration. This protocol was registered in PROSPERO (CRD42017063991).

Keywords: cancer; risk; cigarette smoking; dose-response relationship; systematic review; meta-analysis.

STRENGTHS AND LIMITATIONS OF THE STUDY

- This study represents the most complete and updated review on the association between cigarette smoking and site-specific cancer risk.
- We will not conduct a systematic review of the entire scientific literature on the issue, but we will rather use an original and innovative approach combining an umbrella review and traditional systematic reviews.
- We will carry out dose-response analyses using two-steps random-effects meta-regression models in order to examine potential nonlinear relationships between smoking-related variables and the risk of cancer.
- We will not systematically consider the assignment of a quality score to all original publications for each cancer site.

INTRODUCTION

In 1950, Ernst Wynder and Evards Graham[1] and Richard Doll and Bradford Hill[2] first reported an association between tobacco smoking and lung cancer risk. Over the subsequent 65 years, thousands of studies systematically confirmed this association and found that tobacco smoking also increases the risk of several other neoplasms.[3] In 2004, tobacco smoking has been classified as carcinogenic to humans (Group 1) by the International Agency for Research on Cancer (IARC)[3], which provided evidence on the causal relationship between cigarette smoking and cancer of the lung, oral cavity, nasal cavity and paranasal sinuses, nasopharynx, oropharynx, hypopharynx, larynx, oesophagus, stomach, pancreas, liver, kidney, ureter, urinary bladder, cervix and myeloid leukaemia.[3] The results on the association between (cigarette) smoking and other cancer sites, including cancers of the breast and endometrium, remain conflicting.[3]

Since late 1980s, publications reporting results from systematic reviews and meta-analyses, the large majority of which were unnecessary and misleading, rapidly increased.[4] Although the effects of cigarette smoking on cancer incidence and/or mortality have been largely investigated, only a limited number of systematic reviews/meta-analyses are available on the quantification of the dose-response relationship between selected smoking variables, including smoking intensity, duration, pack-years and time since quitting, and the risk of cancer. More importantly, just a few meta-analyses, if any, provided the cancer-specific risk curve functions of the dose-response relationships. Dose-response data, however, are crucial to provide reliable and accurate estimates of the cancer burden due to smoking, both at individual - absolute cancer incidence/mortality obtained through lung cancer risk-assessment models and risk charts[5 6] - and population level - smoking attributable deaths.[7 8] Currently, the methods developed to quantify the cancer burden due to smoking use cancer-specific estimates of relative risks (RR) according to tobacco smoking derived from a few large cohorts, mainly from the USA.[9] The use of these cohorts to derive RRs may lead to validity problems when applying these estimates to other populations with different smoking patterns.[7 10] Furthermore, these RRs are often estimated after allowance for a limited number of

socio-demographic characteristics, excluding potentially important confounding variables, such as alcohol drinking.

Using an original and innovative approach, which combines an umbrella review and traditional systematic reviews, we aim at providing a comprehensive and updated picture of the association between various smoking-related variables and the risk of all cancers. We will be able to estimate the most robust cancer-specific RRs, obtained from the existing scientific literature, possibly derived after adjustment for relevant covariates. Moreover, for each cancer site, we will be able to provide the risk curves which best describe the dose-response effect of smoking intensity, smoking duration, pack-years, and time since stopping smoking on cancer risk.

METHODS

The present cancer-specific systematic reviews/meta-analyses will be based on the following three subsequent literature searches on the association between cigarette smoking and cancer risk:

1. Umbrella review: a systematic review to identify published meta-analyses, pooled-analyses, and systematic reviews providing data on the association between cigarette smoking and cancer risk;

2. Update of available cancer-specific reviews: for each cancer site, the conduction of a systematic review of studies providing original data (including pooled-analyses on individual participants data) on the association between cigarette smoking and cancer-specific risk, starting from the last available comprehensive review publication identified through the umbrella review (point 1);

3. Review of all original publications: a review of all original articles on the association between cigarette smoking and site-specific cancer risk included in the cancer-specific review publications identified through the umbrella review (point 1) or identified through the update of the available reviews (point 2).

The design of the present systematic reviews was developed following the Preferred Reporting Items for Systematic review and Meta-Analyses (PRISMA) guidelines[11] and its extension for

protocols (PRISMA-P).[12 13] This protocol was registered in the International Prospective Register of Systematic Reviews (PROSPERO) on 4 May 2017 (registration number: CRD42017063991).

1. Umbrella review

Search strategy

We will conduct a systematic literature search to identify all published meta-analyses, pooled-analyses, and systematic reviews providing data on the association between cigarette smoking and the risk of various cancers. Literature search strategy will include combinations of Medical Subjects Headings (MeSH) and text words related to *cancer* and *tobacco* or *smoking*, and will be restricted to the following publication types: meta-analyses, pooled-analyses, and systematic reviews. No restriction on cancer site or on publication date will be applied. The following databases will be used: MEDLINE, Embase, ISI Web of Science (WoS), and Cochrane Database of Systematic Reviews (CDSR). The search strings to be used in various databases are reported in **Appendix 1**. To ensure literature saturation, reference lists of selected relevant publications identified through the search will also be checked. Besides the publications found through the databases searches, we will also consider the reviews of the literature on tobacco smoking provided within the IARC monographs vol. 83[3] and vol. 100E[14] and the Surgeon General Report[15], three reports of known importance providing data on the association between tobacco smoking and various cancer sites.

Eligibility criteria

Study design

We will include meta-analyses, pooled-analyses and systematic reviews of observational studies providing measures of RRs between cigarette smoking and cancer risk. Original observational studies (e.g., case-control, cohort or cross-sectional studies) will be excluded. Reports, letters to the editor, book chapters, conference proceedings, dissertations, and theses will be not considered.

Conditions

We will considers publications providing data on the following 28 (namely all) malignant neoplasms: cancer of lip, oral cavity and pharynx (ICD-10: C00-C14), nasopharynx (C11), oesophagus (C15), stomach (C16), colon (C18), rectum and anus (C19-C21), liver (C22), gallbladder (C23-C24), pancreas (C25), larynx (C32), lung trachea and bronchus (C33-C34), bone (C40-C41), melanoma of skin (C43), mesothelioma (C45), breast (C50), cervix uteri (C53), corpus uteri (C54), ovary (C56), prostate (C61), testis (C62), kidney (C64), bladder (C67), brain, central nervous system (C70-C72), thyroid (C73), Hodgkin lymphoma (C81), non-Hodgkin lymphoma (C82-C86, C96), multiple myeloma (C88-C90) and leukaemia (C91-C95). We will also consider review publications on groups of cancers (e.g., head and neck, upper aero-digestive tract, or intestinal cancers). Studies specifically based on benign neoplasms, such as colorectal polyps, acoustic neuroma and meningioma, and neuroendocrine tumours will be excluded.

Participants

We will include review publications providing data on humans, in the general population. We will therefore exclude review publications based on patients with cancer or other diseases (i.e., reporting data on the effect of smoking on the prognosis of the disease), or on subgroups of the population with selected lifestyle habits or other characteristics (e.g., populations limited to alcohol drinkers or tobacco smokers). No restriction will be applied according to age of participants at cancer incidence or mortality given that practically all studies providing data on the association between cigarette smoking and cancer risk are based on adults, only.

Exposures

We will include all review publications providing data on the use of cigarettes in the general population. Publications focused on the use of tobacco products other than cigarettes (e.g., pipe, smokeless tobacco, cigar, water pipe, electronic cigarettes) or on the exposure to second-hand smoke will be excluded.

Outcomes

The primary outcomes of interest will be: i) the excess incidence and/or mortality of various cancers in current/ever smokers compared to non/never smokers; ii) the dose-response curves describing the association between cigarette smoking duration, intensity and time since stopping and incidence and/or mortality for various cancers.

Languages

We will include only articles published in English language.

Study selection

All the review publications found in various electronic databases through the above mentioned search strategy will be uploaded in an EndNote library (EN1), and duplicate records will be deleted. Titles and/or abstracts of the meta-analyses, pooled-analyses and systematic reviews will be screened independently by two reviewers (AL, GP) to exclude publications which will not meet the eligibility criteria. The full text of the remaining review publications will be retrieved and independently assessed for eligibility by the two reviewers. Discrepancies between the two reviewers will be discussed and solved. In case of disagreement, a third reviewer (SG) will help to find a final decision. Data from other available reports will also be integrated in the same EN1 library.

Quality assessment

Assessment of the quality of various review publications is out of the scope of the present systematic review. Thus, no quality score will be assigned to the publications. No review publication will be excluded *a priori* for weakness of design or data quality.

Data extraction and management

A standardised form in Microsoft (MS) Excel will be used to extract data from each identified review publication. Relevant information will include: first author, year of publication, type of study (i.e., meta-analysis, pooled-analysis, or systematic review), cancer site(s) and/or subsite(s), endpoint (i.e., incidence and/or mortality), and other information about the methodology of studies included in the reviews (e.g., study design, country(ies) of study conduction, number of studies considered in, and overall population size). Data will be extracted from each included meta-analysis, pooled analysis or systematic review by one reviewer and verified by a second reviewer. Any disagreement between the two reviewers will be solved by consensus. Otherwise, a third reviewer will help to find a final decision. After data extraction, publications will be grouped according to the considered cancer site.

Preliminary umbrella review

In April 2017, we already conducted a preliminary umbrella review through which we identified the meta-analyses, pooled-analyses and systematic reviews on the association between smoking and cancer risk published before 28th April 2017. Within the comprehensive literature search, we found a total of 1430 publications from the four considered databases (726 from Medline, 316 from Embase, 376 from ISI WoS, and 12 from CDSR; **Figure 1**). After the exclusion of duplicates (n=542) and ineligible papers (n=716), and after the inclusion of three important reports[3 14 15], we obtained a total of 175 relevant publications (i.e., 107 meta-analyses, 52 pooled-analyses, and 16 systematic reviews) on the association between smoking and risk of various neoplasms.

2. <u>Update of the available cancer-specific reviews</u>

Search strategy

For each of the 28 cancer sites previously described in the umbrella review, we will identify the last available comprehensive systematic review or meta-analysis on the association with cigarette

smoking. We will then update the identified cancer-specific review through the conduction of a systematic literature search on all observational studies (e.g., case-control, cohort, and nested case-control studies) providing original data on the association between cigarette smoking and site-specific cancer risk, and published after the year of publication of the most recent article included in the last comprehensive review. Only studies published in English language will be considered.

The literature search strategy will be conducted in MEDLINE and Embase, and will include combinations of MeSH terms and text words related to *site-specific cancer* and *tobacco* or *smoking* (**Appendix 2**).

Eligibility criteria

Study design

We will include original observational studies (e.g., case-control, cohort, nested case-control studies, or pooled-analysis of individual participant data) providing measures of RRs of the association between cigarette smoking and site-specific cancer risk. Reports, book chapters, conference proceedings, dissertations, and theses will not be considered. We will exclude case-control studies using patients with cancer or other chronic diseases as comparison group.

Comparator

Never smokers are our comparators. We will in fact consider, as the measure of association, the RRs of smoking variables compared to never smokers.

Eligibility criteria for *conditions*, *participants*, *exposures*, and *outcomes* are those reported also for the umbrella review.

Study selection

For each cancer site, we will upload all the original publications found using the above mentioned search strategy in cancer-specific EndNote libraries (EN2_1-EN2_28) and we will delete the identified duplicates. Titles and/or abstracts of original articles will be screened independently by

two reviewers (AL, GP) to exclude publications that do not meet the inclusion criteria outlined above. The full text of the remaining original publication will be retrieved and independently assessed for eligibility by the two reviewers. Discrepancies on the assessment between the two reviewers will be discussed and solved. In case of disagreement, a third reviewer (SG) will help to find a final decision.

3. Review of all original publications

For each cancer site, we will upload in an EndNote library (EN3_1-EN3_28) all the original publications obtained from the 28 cancer-specific reviews identified in the umbrella review (point 1). In the same EndNote libraries, we will add the original publications obtained from the corresponding updates of the reviews (point 2), and duplicate publications will be deleted. The full text of all the retained original publications will be retrieved. Non-English reports, unpublished studies, conference proceedings, dissertations and theses will be excluded.

Figure 2 shows the flowchart we will consider for each of the 28 cancer-specific reviews. For each cancer site, original publications from both the umbrella review and the update of the reviews will be considered.

Quality assessment

Although quality assessment of original publications is out of the scope of the present systematic review, we do not exclude the possibility, at least for selected neoplasms, to assign a quality score to the original publications in order to conduct sensitivity analyses excluding the publications with a relatively low quality. In this case, the quality (risk of bias) for each observational study will be assessed using the Newcastle-Ottawa scale for observational studies.[16]

Data extraction and management

For each cancer site, two standardised forms in MS Excel will be used to collect relevant information on the study design and the risk estimates from the original publications. A first form will be used to extract data related to the study design including: first author, year of publication, journal, country, study name, period and design of study, outcome and sample size. In the second form we will collect the exposure categories (i.e., smoking status, intensity, duration, pack-years, age at starting, time since stopping), and corresponding RR estimates (or other estimates, such as odds ratios and hazard ratios) and 95% confidence intervals (CIs). The number of cases in each exposure category and covariates used in the model will be also collected. When the results of the same study have been published in more than one original publication, only data from the most recent and/or complete article will be retained and reported in the second Excel form.

Data analysis

For each cancer site, we will pool all the RRs or other risk estimates (e.g., hazard ratios and odds ratios) in order to obtain the association between smoking status (separately for current, former and ever compared to never smokers) and the risk of cancer. Because cancer is a relatively rare outcome, we assume that ORs, risk ratios and rate ratios are all comparable estimates of the RR. Heterogeneity between studies will be assessed using the Cochran Q test and the I-squared statistics, i.e. the proportion of total variation contributed by between-study heterogeneity.[17] As we anticipated between study heterogeneity, we will present pooled RRs from random-effects models using the DerSimonian and Laird moment estimator of the between-study variance component.[18] However, if no heterogeneity between-study is detected, pooled estimates from the random-effects model will correspond to those deriving from the fixed-effect model.

When RRs are not reported for ever smokers, but only separately for current and former smokers, we will use the method for pooling non-independent estimates described by Hamling et al.[19] to obtain RRs for ever smokers besides those for current and former compared to never smokers.

This method uses the number of subjects exposed to different levels of smoking and non-exposed subjects, and the corresponding risk estimates, to derive a set of pseudo-numbers of cases and controls/subjects at risk, by taking into account the correlation between the original estimates due to the common reference group. The obtained pseudo-numbers will be used to compute the new adjusted RR. This methodology will also be used to convert RR estimates when the reference category considered in the study is not represented by never smokers.

For smoking intensity, duration and time since stopping smoking, we will compute pooled RRs according to various categories of the considered smoking-related variables (e.g, low, intermediate, high cigarette consumption). Moreover, we will carry out dose-response analyses using two-steps random-effects meta-regression models in order to examine potential nonlinear relationships between those variables and the risk of cancer. In particular, we will consider a method providing the best fitting two-term fractional polynomial model[20], and a method modelling the considered smoking-related variables using restricted cubic splines.[21]

If the necessary data are available, we will consider to further conduct separate analyses by sex, study period, geographic area, and income (defined on the basis of per-capita Gross Domestic Product).

Analyses will be conducted using SAS software version 9.4 (SAS Institute Inc., Cary, NC, USA), and R software version 3.3.0 (R Development Core Team, 2008), in particular *meta* and *dosresmeta* packages.

ETHICS AND DISSEMINATION

Ethics

This review does not require approval from an ethics committee, since no individual-level patients' data will be collected.

Implications and dissemination

Through these systematic reviews and meta-analyses, we will provide the most complete and updated estimates on the association between cigarette smoking and site-specific cancer risk. These estimates will be used to quantify the cancer burden due to cigarette smoking at both individual and population level. This information is essential to guide policy decisions to control tobacco smoking and improve cancer prevention.

Given the relevance and originality of this project, we plan to publish results from the meta-analyses in peer-reviewed journals, considering either single cancer sites or various apparatus or tracts. A final publication will provide the summary results for all cancers. We will also include the main results of our systematic reviews and meta-analyses in a publicly available website. Readers will have the possibility to contact us to communicate possible lacks or updates.

Contributions

SG and AL had the original idea of the work. AL, GP, and SG structured the study protocol, designed the search strategies, and conducted the umbrella review. MR, VB, and CB provided statistical and epidemiological supervision. All authors critically revised the draft of the manuscript and approved its final version.

Funding

This work was partially supported by the Italian League Against Cancer (Milan). AL was supported by a fellowship from the Italian Association for Cancer Research (AIRC). MR was supported by a fellowship from the Italian Foundation for Cancer Research (FIRC).

Competing interests

None declared.

REFERENCES

1. Wynder EL, Graham EA. Tobacco smoking as a possible etiologic factor in bronchiogenic carcinoma; a study of 684 proved cases. *J Am Med Assoc* 1950;143(4):329-36.
2. Doll R, Hill AB. Smoking and carcinoma of the lung; preliminary report. *Br Med J* 1950;2(4682):739-48.
3. IARC. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Volume 83. Tobacco Smoke and Involuntary Smoking. Lyon, France: International Agency for Research on Cancer 2004.
4. Ioannidis JP. The Mass Production of Redundant, Misleading, and Conflicted Systematic Reviews and Meta-analyses. *Milbank Q* 2016;94(3):485-514.
5. Spitz MR, Hong WK, Amos CI, *et al.* A risk model for prediction of lung cancer. *J Natl Cancer Inst* 2007;99(9):715-26.
6. Tammemagi MC, Katki HA, Hocking WG, *et al.* Selection criteria for lung-cancer screening. *N Engl J Med* 2013;368(8):728-36.
7. Perez-Rios M, Montes A. Methodologies used to estimate tobacco-attributable mortality: a review. *BMC Public Health* 2008;8:22.
8. Tachfouti N, Raherison C, Obtel M, *et al.* Mortality attributable to tobacco: review of different methods. *Arch Public Health* 2014;72(1):22.
9. ACS. Cancer Prevention Study II (CPS II). Available at: http://www.cancer.org/research/researchtopreventcancer/currentcancerpreventionstudies/cancer-prevention-study 2016
10. Gallus S, Muttarak R, Martinez-Sanchez JM, *et al.* Smoking prevalence and smoking attributable mortality in Italy, 2010. *Prev Med* 2011;52(6):434-8.
11. Moher D, Liberati A, Tetzlaff J, *et al.* Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;339:b2535.

12. 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.

13. Shamseer L, Moher D, Clarke M, *et al.* Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ* 2015;349:g7647.

14. IARC. IARC monograph on the evaluation of carcinogenic risks to humans. Volume 100E. A review of human carcinogens - Personal habits and indoor combustions. Available at: http://monographs.iarc.fr/ENG/Monographs/vol100E/mono100E.pdf. Lyon, France2012.

15. U.S. Department of Health and Human Services. The Health Consequences of Smoking - 50 Years of Progress A Report of the Surgeon General. Available at: https://www.surgeongeneral.gov/library/reports/50-years-of-progress/full-report.pdf. Atlanta, GA2014.

16. Wells GA, Shea B, O'Connell D, *et al.* The Newcastle-Ottawa Scale (NOS) for assessing the quality if nonrandomized studies in meta-analyses. Available online at: http://www.ohri.ca/programs/clinical_epidemiology/oxford.htm. 2014

17. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002;21(11):1539-58.

18. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7(3):177-88.

19. Hamling J, Lee P, Weitkunat R, *et al.* Facilitating meta-analyses by deriving relative effect and precision estimates for alternative comparisons from a set of estimates presented by exposure level or disease category. *Stat Med* 2008;27(7):954-70.

20. Rota M, Bellocco R, Scotti L, *et al.* Random-effects meta-regression models for studying nonlinear dose-response relationship, with an application to alcohol and esophageal squamous cell carcinoma. *Stat Med* 2010;29(26):2679-87.

21. Orsini N, Li R, Wolk A, *et al.* Meta-analysis for linear and nonlinear dose-response relations: examples, an evaluation of approximations, and software. *Am J Epidemiol* 2012;175(1):66-73.

FIGURES LEGEND

Figure 1. Flowchart for the selection of papers (published before 28th April 2017) in the umbrella review.

ISI WoS: ISI Web of Science

CDSR: Cochrane Database of Systematic Reviews

Figure 2. Flowchart for each of the 28 cancer-specific reviews.

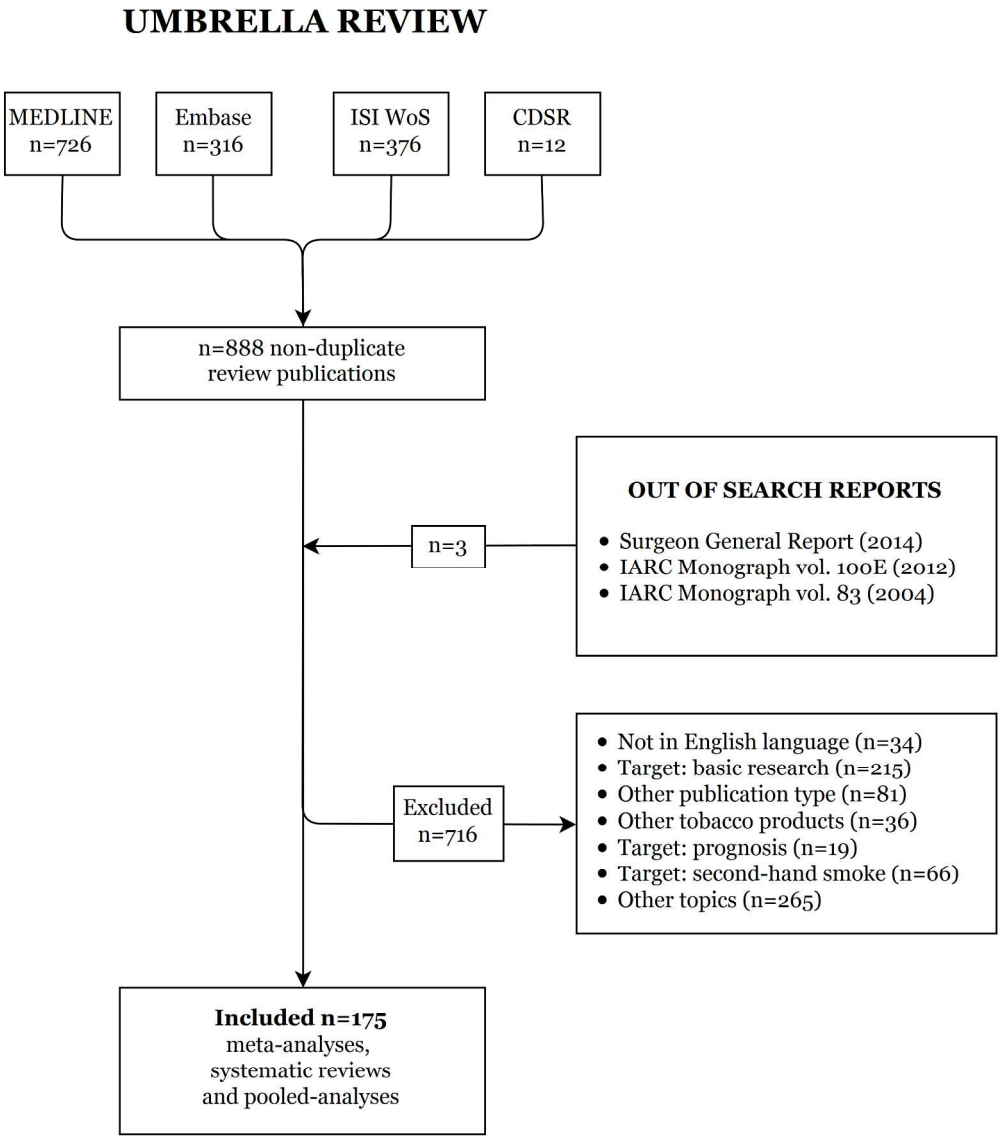

Figure 1. Flowchart for the selection of papers (published before 28th April 2017) in the umbrella review.

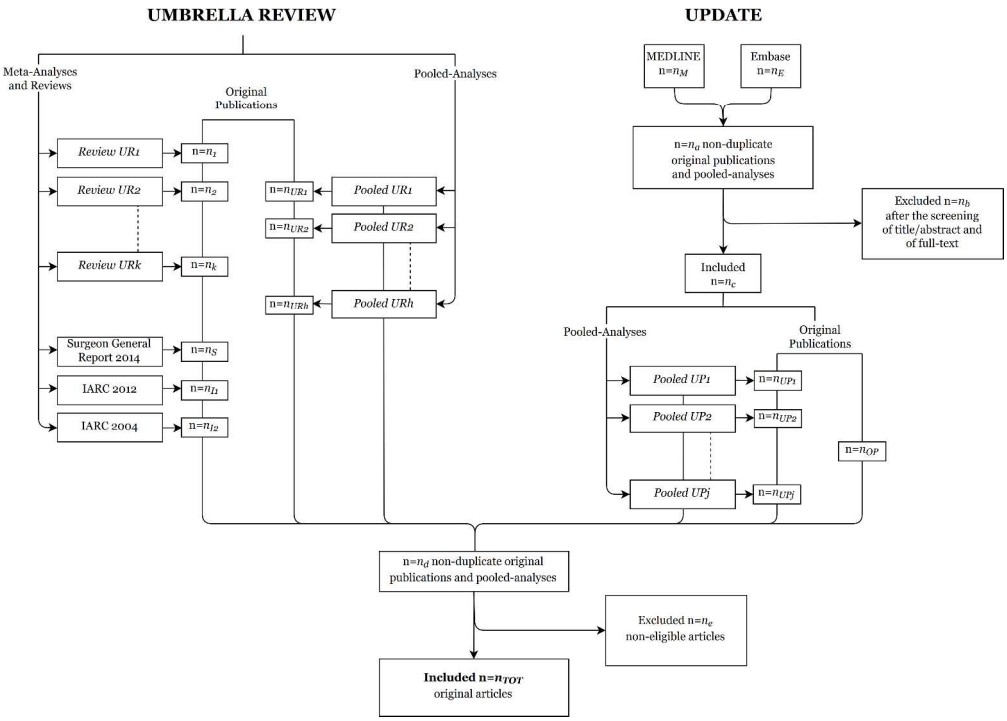

Figure 2. Flowchart for each of the 28 cancer-specific reviews.

Appendix 1. Literature search strings for the umbrella review used in various databases.

MEDLINE

(neoplasms [MeSH Terms] OR cancer [title] OR cancers [title] OR neoplasm [title] OR neoplasms [title] OR leukaemia [title] OR leukemia [title] OR leukaemias [title] OR leukemias [title] OR carcinoma[title] OR maligna*[title]) AND (meta-analysis OR "meta analysis" OR pooled-analysis OR "pooled analysis" OR "systematic review"[tiab] OR "systematic revision"[tiab]) AND (cigarette [title] OR cigarettes [title] OR tobacco [title] OR smoking [title] OR smokers [title] OR smoking [MeSH Terms])

Embase

(cancer:ti OR cancers:ti OR neoplasm:ti OR neoplasms:ti OR leukaemia:ti OR leukemia:ti OR leukaemias:ti OR leukemias:ti OR carcinoma:ti OR maligna*:ti) AND ('meta analysis' OR 'pooled analysis' OR 'meta analysis':it) AND (cigarette:ti OR cigarettes:ti OR tobacco:ti OR smoking:ti OR smokers:ti)

ISI Web of Science

TITLE: ((cancer OR cancers OR neoplasm OR neoplasms OR leukaemia OR leukemia OR leukaemias OR leukemias OR carcinoma OR maligna*) AND (cigarette OR cigarettes OR tobacco OR smoking OR smokers)) AND TOPIC: ((meta-analysis OR "meta analysis" OR pooled-analysis OR "pooled analysis" OR "systematic review" OR "systematic revision")) NOT DOCUMENT TYPES: (Letter OR Meeting Abstract OR Meeting Summary)

Cochrane Database of Systematic Reviews

(cancer:ti or cancers:ti or neoplasm:ti or neoplasms:ti or leukaemia:ti or leukemia:ti or leukaemias:ti or leukemias:ti or carcinoma:ti or maligna*:ti) and (meta-analysis or "meta analysis" or pooled-

analysis or "pooled analysis") and (cigarette:ti or cigarettes:ti or tobacco:ti or smoking:ti or smokers:ti)

Appendix 2. Literature search strings for the update of the available reviews used in various databases.

MEDLINE

((((*cancer site*) AND (cancer OR neoplasm OR carcinoma OR Neoplasms [MeSH Terms]) AND (cigarette OR cigarettes OR tobacco OR smoking OR smokers OR smoking [MeSH Terms]))) AND English[Language]) AND ("*last review year*"[Date - Publication]: "2017"[Date - Publication])

Embase

cigarette:ti OR cigarettes:ti OR tobacco:ti OR smoking:ti OR smokers:ti AND (*cancer site*:ab,ti) AND (cancer:ab,ti OR neoplasm:ab,ti OR carcinoma:ab,ti) AND (article:it OR review:it) AND [english]/lim AND [*last review year*-2017]/py NOT [medline]/lim

Cancer site stands for the terms used to identify the site-specific cancer (e.g., for liver cancer, the combination of the following terms will be considered: liver, hepatocellular, hepatic).

Last review year refers to the year of publication of the most recent article included in the last comprehensive review.

PRISMA-P (Preferred Reporting Items for Systematic review and Meta-Analysis Protocols) 2015 checklist: recommended items to address in a systematic review protocol*

Section and topic	Item No	Checklist item	Page N°
ADMINISTRATIVE INFORMATION			
Title:			
Identification	1a	Identify the report as a protocol of a systematic review	1
Update	1b	If the protocol is for an update of a previous systematic review, identify as such	NA
Registration	2	If registered, provide the name of the registry (such as PROSPERO) and registration number	2
Authors:			
Contact	3a	Provide name, institutional affiliation, e-mail address of all protocol authors; provide physical mailing address of corresponding author	1
Contributions	3b	Describe contributions of protocol authors and identify the guarantor of the review	15
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	NA
Support:			
Sources	5a	Indicate sources of financial or other support for the review	15
Sponsor	5b	Provide name for the review funder and/or sponsor	NA
Role of sponsor or funder	5c	Describe roles of funder(s), sponsor(s), and/or institution(s), if any, in developing the protocol	NA
INTRODUCTION			
Rationale	6	Describe the rationale for the review in the context of what is already known	5
Objectives	7	Provide an explicit statement of the question(s) the review will address with reference to participants, interventions, comparators, and outcomes (PICO)	6
METHODS			
Eligibility criteria	8	Specify the study characteristics (such as PICO, study design, setting, time frame) and report characteristics (such as years considered, language, publication status) to be used as criteria for eligibility for the review	7-9, 11
Information sources	9	Describe all intended information sources (such as electronic databases, contact with study authors, trial registers or other grey literature sources) with planned dates of coverage	7,10

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	19-21
Study records:			
Data management	11a	Describe the mechanism(s) that will be used to manage records and data throughout the review	9-12
Selection process	11b	State the process that will be used for selecting studies (such as two independent reviewers) through each phase of the review (that is, screening, eligibility and inclusion in meta-analysis)	9-12
Data collection process	11c	Describe planned method of extracting data from reports (such as piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators	9-12
Data items	12	List and define all variables for which data will be sought (such as PICO items, funding sources), any pre-planned data assumptions and simplifications	9-12
Outcomes and prioritization	13	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale	9
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	NA
Data synthesis	15a	Describe criteria under which study data will be quantitatively synthesised	13-14
	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 (such as I^2, Kendall's τ)	13-14
	15c	Describe any proposed additional analyses (such as sensitivity or subgroup analyses, meta-regression)	13-14
	15d	If quantitative synthesis is not appropriate, describe the type of summary planned	NA
Meta-bias(es)	16	Specify any planned assessment of meta-bias(es) (such as publication bias across studies, selective reporting within studies)	13-14
Confidence in cumulative evidence	17	Describe how the strength of the body of evidence will be assessed (such as GRADE)	NA

*** It is strongly recommended that this checklist be read in conjunction with the PRISMA-P Explanation and Elaboration (cite when available) for important clarification on the items. Amendments to a review protocol should be tracked and dated. The copyright for PRISMA-P (including checklist) is held by the PRISMA-P Group and is distributed under a Creative Commons Attribution Licence 4.0.**

From: Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart L, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ. 2015 Jan 2;349(jan02 1):g7647.