Table of Contents

Details on methodology

Determining topics of interest, formulation of PICO questions and outcomes, rating the importance of outcome parameters

During the initial task force meeting, eight topics of interest within the scope of this task force on various aspects of quality in lung cancer care were developed and consented based on an unanimous decision by the task force members. Subsequently, well-framed questions were formulated for each topic of interest by the co-chairs and the lead methodologist using the PICO (Population, Intervention, Comparator and Outcomes) format [1]. Equally, outcome parameters for the questions were defined and consented by the Task Force group (see ***online supplement section B*** for detailed PICO questions including outcome parameters).

For each question, all outcome parameters were rated individually regarding their importance for clinical decision making for the respective question by the accountable task force members applying a rating scale from 1 to 9 (of limited importance: 1-3, important: 4-6, critical: 7-9). Rounded means of the scorings rates were calculated and further discussed, leading to final importance score for each outcome (not important, important, critical) (see ***online supplement section C*** for results of the outcome parameter scoring).

Literature searches

The literature search was designed by V. Durieux and T. Berghmans and reviewed by T. Blum. Searches were performed in April 2016 and updated in September 2017, September 2018, May 2019, December 2019, May 2020 and January 2021 in the Medline database by one medical librarian (V. Durieux), experienced in searching for medical and scientific publications, and supervised by the two co-chairs (T. Berghmans, T. Blum).

Ovid Medline was searched using the OvidSP interface (Online supplement 3. Search strategies). Unless otherwise stated, search terms were MeSH terms (medical subject headings). MeSH terms were also combined with relevant free-text terms that were searched for in titles and abstracts.

The corresponding PICO search criteria were translated into MeSH terms and free-text keywords which were usable as search equations by the OvidSP interface. Completed search strategies included P and I criteria, further limited by O criteria only when the number of retrieved citations for the P and I criteria combinations exceeded 5,000 citations. This cut-off was chosen arbitrary to ensure a meaningful, but manageable basic set of studies per PICO for the selection process by avoiding excessive noise around the evidence of interest. Although a possible risk of study selection bias could not be completely ruled out, initial validity controls of the search equations demonstrated that this differentiated approach recorded all the reference articles known to the authors in advance in the context of all the PICO questions. Subsequently, topic-related review articles and systematic reviews which were identified during the periodic literature searches served as external validity controls for the completeness of search results. The Ovid Medline search strategies are provided in the ***online supplement section D***.

For each of the PICO searches, citations were separately exported from Medline into reference management software (Endnote®) to allow the removal of duplicates and to facilitate the selection process performed by reviewers (V. Durieux).

Study selection

Studies were eligible for selection if fulfilling pre-defined inclusion criteria relating to study type (randomized controlled trials, non-randomized controlled trials as well as observational studies with cohort and case-control studies), publication language (English, French, Dutch, German, and Spanish) and according to the PICO-questions themselves. Detailed eligibility criteria are listed in ***online supplement table 4***.

In a first step, exported references were screened for relevance. Articles were selected or rejected on initial screening by two independent reviewers depending on whether titles and abstracts met or did not meet the inclusion criteria, respectively (PICO 1: B. Grigoriu, T. Berghmans; PICO 2: A.-P. Meert, T. Berghmans; PICO 3: J.-P. Sculier, T. Berghmans; PICO 4: P. Knaut, T. Blum; PICO 5: D. Subotic, T. Blum; PICO 6: D. Jovanovic, T. Blum; PICO 7: R. Muhr, T. Blum; PICO 8: P. Knaut, T. Blum). In case of discrepancy of the evidence synthesis results for one PICO question, consensus was sought by the two reviewers. Full paper publications were requested if at least one reviewer selected the reference.

Full paper publications for all articles selected in step 1 were collected and linked to the reference management databases (T. Blum, P. Knaut, R. Muhr, V. Durieux). In a second step, for each question, the remaining articles were then evaluated based on the full paper publications for final inclusion in the current systematic reviews. Final selections were done in agreement of the two allocated reviewers for the respective PICO questions, again discrepancies being resolved by consensus. The same methodology was applied by the two co-chairs for the literature update searches (T. Berghmans, T. Blum).

These selections were supplemented by the above-mentioned update searches as well as additional full paper publications derived from screening the references of the selected articles as well as other related literature known to the task force group. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagrams were utilized to report the search results for each of the eight questions (V. Durieux, T. Blum) [2].

All results were provided to the task force members to check for missing evidence.

Data extraction and risk of bias assessment

Two task force members were allocated to collaboratively perform the data extraction from the studies and the evidence assessment per search question (PICO 1: B. Grigoriu, T. Berghmans; PICO 2: A.-P. Meert, T. Berghmans; PICO 3: J.-P. Sculier, T. Berghmans; PICO 4: P. Knaut, T. Blum; PICO 5: D. Subotic, T. Blum; PICO 6: D. Jovanovic, T. Blum; PICO 7: R. Muhr, T. Blum; PICO 8: P. Knaut, T. Blum). The two task force co-chairs T. Berghmans and T. Blum took the lead in performing the subsequent tasks for PICO-question 1-3 and 4-8, respectively, supported by the named other task force member for each PICO. First, the respective pairs per PICO extracted study characteristics, types of participants, interventions, outcomes measured, and results from each of the selected studies. Second, they assessed the risk of bias for the individual outcomes within every selected study. Randomized controlled trials were analysed for specific study limitations according to *the Cochrane risk of bias tool for assessing risk of bias in randomized trials* consisting of five discrete items: 1. lack of randomization (selection bias), 2. lack of allocation concealment (selection bias), 3. lack of blinding (performance bias), 4. incomplete accounting of patients and outcome events (attrition bias), and 5. selective outcome reporting (reporting bias) [3]. In addition to these domains, observational studies were also assessed based on four distinct items: 1. failure to develop and apply appropriate eligibility criteria, 2. flawed measurement of both exposure and outcome, 3. failure to adequately control confounding, and 4. incomplete follow-up in observational studies [4].

Specific data collection forms were designed for a standardised handling of the extracted data and the risk of bias assessment for the individual outcomes within the studies (T. Berghmans, D. Rigau, T. Tonia). In case of discrepancy of the extraction or assessment results, consensus was sought by the two task force members allocated to each PICO.

All results were provided to the task force members for validation.

Data synthesis, meta-analyses, own individual four-stage effect strength evaluation scheme as well as assessing the effect direction, effect size and certainty of the evidence

Just like in the narrative review by the preceding task force [5], substantial heterogeneity of evidence regarding study designs and methodologies was detected for all PICO questions. To ensure meaningful evidence synthesis and conclusion of guideline recommendations, the three task force chairs together with one co-lead methodologist (R. Morgan) thoroughly discussed and agreed upon the formation of representative subgroups (based on type of cancer and cancer staging) out of the selected full publications to ensure the most direct evidence to answer the PICO questions. These were selected posteriori to the risk of bias assessment of individual studies.

Whenever clinically meaningful, the available evidence per outcome was synthesized quantitatively with the calculation of aggregated effects using a meta-analytic technique and personal programming for calculating these combined effects. For binary outcomes, the individual effect of the covariate of interest was reported as odds ratio with a 95% confidence interval (CI) and for continuous outcomes, it was reported as a mean difference with a 95% CI. Heterogeneity between individual effects was assessed using the I^2 statistic. Heterogeneity was suspected and explored when I^2 was greater than 60%. For survival outcomes, individual effects were

summarized with hazard ratios (HR) and 95% CI. Fixed-effects model was used when fewer than 3 studies were pooled, otherwise, random effects models were used. Findings from the meta-analysis were presented in a forest plot using the Metaplot®-software. If needed data to allow a quantitative synthesis were not stated or incalculable out of the publications, results were summarized narratively.

In situations in which statistically pooling the results of studies was inappropriate, i. e. studies were too different to reasonable synthesize in a meta-analysis, the evidence were presented narratively [6]. For each outcome of interest, the two task force co-chairs predefined an individual four-stage evaluation scheme to discriminate the effect size into trivial, small, moderate and large which contains self-selected thresholds. In view of a lack of evidence-based recommendations, the absolute and relative threshold values for each outcome were chosen based on clinical experience providing a clinical meaningful distinction between a trivial and small effect as well as between small, moderate, and large effects. The evaluation scheme which was agreed upon by all task force members is depicted in **Table 1**. To assess the effect direction and effect size of the evidence per outcome, first, all single studies were rated according to this evaluation scheme. Second (if meta-analysis was not feasible), individual studies were grouped according to their estimated effect sizes and respective included patient figures per study were summed up for each of the four effect size categories. Finally, the effect direction for each outcome was determined depending on whether the total number of patients predominated in the group of studies with small to large effects for the intervention or the group with trivial effects or those even opposite to the intervention. Likewise, the overall effect size per outcome was selected based on the effect size category with the largest number of included patients.

Outcome	**Absolute difference**	**Relative difference (if absolute difference incalculable/not stated)**
1. Overall survival	5-year overall survival rate benefit:	HR point estimate benefit:
-large	- >5%	- >10%
-moderate	- >2.5-5%	- >5-10%
-small	- >1-2.5%	- >1-5%
-trivial	- ≤1%	- ≤1%
2. Disease-free survival (DSF)	DFS benefit:	HR point estimate benefit:
-large	- >12 months	- >10%
-moderate	- >6-12 months	- >5-10%
-small	- >1-6 months	- >1-5%
-trivial	- ≤1%	- ≤1%
3. Progression-free survival (PFS)	PFS benefit:	HR point estimate benefit:
-large	- >6 months	- >10%
-moderate	- >3-6 months	- >5-10%
-small	- >1-3 months	- >1-5%
-trivial	- ≤1 month	- ≤1%
4. Mortality	Mortality rate benefit:	OR point estimate benefit:
-large	- >1.5%	- >30%
-moderate	- >1.0-1.5%	- >20-30%
-small	- >0.5-1.0%	- >10-20%
-trivial	- ≤0.5%	- ≤10%
5. Morbidity	Morbidity rate benefit:	OR point estimate benefit:
-large	- >5%	- >40%
-moderate	- >2.5-5%	- >20-40%
-small	- >1-2.5%	- >10-20%
-trivial	- ≤1%	- ≤10%
6. Accuracy of staging	Accuracy of staging rate benefit:	OR point estimate benefit:
-large	- >10%	- >20%
-moderate	- >5-10%	- >10-20%
-small	- >1-5%	- >1-10%
-trivial	- ≤1%	- ≤1%
7. Pathological confirmation	Pathological confirmation rate benefit:	OR point estimate benefit:
-large	- >10%	- >20%
-moderate	- >5-10%	- >10-20%
-small	- >1-5%	- >5-10%
-trivial	- ≤1%	- ≤5%
8. Receipt of curative treatment	Curative treatment rate benefit:	OR point estimate benefit:
-large	- >10%	- >20%
-moderate	- >5-10%	- >10-20%
-small	- >1-5%	- >5-10%
-trivial	- ≤1%	- ≤5%

9. Receipt of any tumour-specific treatment -large -moderate -small -trivial	Any tumour-specific treatment rate benefit: - >10% - >5-10% - >1-5% - ≤1%	OR point estimate benefit: - >20% - >10-20% - >5-10% - ≤5%
10. Quality of Life -large -moderate -small -trivial	Quality of Life improvement: a) Center for Epidemiological Study-Depression Scale (score: 0-60 points) - >12 points - >6-12 points - >3-6 points - ≤3 points b) Chronic Respiratory Disease Questionnaire Health-related Quality of Life (20-140 points) - >28 points - >14-28 points - >7-14 points - ≤7 points c) City of Hope Quality of Life Instruments (score: 0-100 points) - >20 points - >10-20 points - >5-10 points - ≤5 points d) Edmonton Symptom Assessment Scale (0-900 points) - >180 points - >90-180 points - >45-90 points - ≤45 points e) EORTC QLQ-C30 (0-100 points) - >20 points - >10-20 points - >5-10 points - ≤5 points f) EQ-5D (0-100 points) - >20 points - >10-20 points - >5-10 points - ≤5 points g) FACIT-Pal (0-184 points) - >36 points - >18-36 points - >9-18 points - ≤9 points h) FACIT-Spiritual Well-Being (score: 0-156 points) - >32 points - >16-32 points - >8-16 points - ≤8 points i) FACT-G (0-108 points) - >22 points - >11-22 points - >6-11 points - ≤6 points j) FACT-L (0-140 points) - >28 points - >14-28 points - >7-14 points	OR point estimate benefit: a)-m) All - >20% - >10-20% - >5-10% - ≤5%

	- ≤7 points k) Multidimensional Quality of Life Scale-Cancer Version (score: 0-100 points) 100 points) - >20 points - >10-20 points - >5-10 points - ≤5 points l) Quality of Life at End of Life (4-20 points) - >4 points - >2-4 points - >1-2 points - ≤1 points m) Reid-Gundlach Satisfaction with Services instrument (0-48 points) - >10 points - >5-10 points - >3-5 points - ≤3 points	
11. Patient satisfaction -large -moderate -small -trivial	Patient satisfaction improvement: a) FAMCARE-P16 scale (score 16-80) - >16 points - >8-16 points - >4-8 points - ≤4 points b) Group Health Association of America Consumer Satisfaction Survey (score:20-100) - >20 points - >10-20 points - >5-10 points - ≤5 points	OR point estimate benefit: a)-b) All - >20% - >10-20% - >5-10% - ≤5%
12. Performance status -large -moderate -small -trivial	Depending on specific performance status-measure	

Table 1: Self-selected evaluation to estimate the effect sizes of single studies per outcome.

The level of certainty of evidence was then assessed per outcome across studies as per GRADE approach [7-9], which grades the certainty of the evidence across 8 domains. Evidence may be rated down for 1) risk of bias [4], 2) imprecision [10], 3) inconsistency [11], 4) indirectness [12], 5) publication bias [13]. If there are no concerns for rating down, the body of evidence informed by observational studies may be rated up due to 1) Large or very large magnitude of effect, 2) dose-response, or 3) opposing residual confounding. The two task force co-chairs T. Berghmans and T. Blum collaboratively conducted the evidence assessment for all PICO questions.

GRADEpro Guideline Development Tool evidence profile forms (available online: www.gradepro.org) were used to present the quality of evidence per outcome across studies for each PICO question (T. Blum, R. Morgan) [8].

All results were discussed with the task force members.

Determining strength and direction of guideline recommendations

The GRADE approach for evidence to decision-making was used to determine the strength and direction of the recommendations collaboratively by the three task force co-chairs for each search question [14]. Accordance was found among the three co-chairs if their initial recommendation proposals related to one search

question differed. Utilizing GRADE Evidence to Decision (EtD) frameworks (available online: www.gradepro.org) [8, 9], in addition to the results of the evidence assessment, the task force members considered the balance of benefits and harms, values and preferences, resource use, health equity, acceptability and feasibility when making recommendations. The task force group members discussed and formulated the guideline recommendations for each of the PICO questions during five virtual task force meetings in March/April 2021 which were used to collect individual feedback on the revised guideline manuscript as well as to formulate and consent the recommendations. The framework for the interpretation of strong and conditional recommendations is depicted in

Table 2.

Target group	Strong recommendations (*)	Conditional recommendations
Patients	All or almost all informed people would choose the recommended choice for or against an intervention.	Most informed people would choose the recommended course of action, but a substantial number would not.
Clinicians	Most patients should receive the recommended course of action.	Recognise that different choices will be appropriate for different patients. Clinicians and other healthcare providers need to devote more time to the process of shared decision-making, by which they ensure that the informed choice reflects individual values and preferences; decision aids and shared decision-making are particularly useful.
Policymakers	The recommendation can be adopted as a policy in most situations.	Policy making will require substantial debate and involvement of many stakeholders.

Table 2: Framework for interpretation of recommendations

(*) strong recommendations based on high or moderate quality evidence will apply to most patients for whom these recommendations are made, but they may not apply to all patients in all conditions; no recommendation can take into account all of the unique features of individual patients and clinical circumstances. [15, 16].

Paradigmatic situations according to GRADE

In accordance with GRADE methodology, the task force group considered strong recommendations despite low or very low quality of evidence in the following five phrased constellations, so-called ***paradigmatic situations*** [16]:

1. "When low quality evidence suggests benefit in a life-threatening situation (evidence regarding harms can be low or high)"

2. "When low quality evidence suggests benefit and high quality evidence suggests harm or a very high cost"

3. "When low quality evidence suggests equivalence of two alternatives, but high quality evidence of less harm for one of the competing alternatives"

4. "When high quality evidence suggests equivalence of two alternatives and low quality evidence suggests harm in one alternative"

5. "When high quality evidence suggests modest benefits and low/ very low quality evidence suggests

possibility of catastrophic harm”

Exceptionally, the task force group also considered good practice statements as an alternative for GRADE-derived recommendations for individual search questions in presence of high-certainty indirect evidence that would be onerous and time-consuming to formally accumulate and review yet supporting the recommendation [16, 17].

Manuscript preparation

The initial draft of the manuscript and the online supplements were prepared by the three co-chairs (T. Berghmans, T. Blum, J. Chorostowska-Wynimko) and two methodologists (R. Morgan, T. Tonia). Both the manuscript and the online supplement were reviewed, edited and approved by all panel members prior to submission.

Detailed description of search questions based on PICO format

1. Do waiting times have an impact on outcome in lung cancer?

PICO question 1: *In patients with lung cancer (or those suspected of having lung cancer), should shorter rather than longer cancer care time intervals be used (e.g., time from diagnosis to treatment)?*

Population: Adult patients with suspected or clinically confirmed lung cancer (e.g., NSCLC, SCLC)

Intervention:

- ***patient interval:*** time from first symptom to first presentation/clinical appearance [regular if ≤ 7-14 days]
- ***doctor interval:*** time from first presentation/clinical appearance to first investigation, primary care responsible for the patient [regular if ≤ 7-14 days]
- ***system interval:*** time from first investigation, primary care responsible for the patient to treatment start [regular if ≤ 28-42 days]
- ***primary care interval:*** time from first presentation/clinical appearance to first referral to secondary care/refer responsibility [regular if ≤ 14-21 days]
- ***secondary care interval:*** time from first referral to secondary care/refer responsibility to treatment start [regular if ≤ 21-35 days]
- ***diagnostic interval:*** time from first presentation/clinical appearance to diagnosis [regular if ≤ 14-28 days]
- ***treatment interval:*** time from diagnosis to treatment start [regular if ≤ 14-28 days]
- ***total interval:*** time from first symptom to treatment start [regular if ≤ 56-84 days]

Remarks: All listed time points and waiting time intervals within the lung cancer care continuum from first symptom to treatment start were adopted from the internationally well-accepted Aarhus statement paper1. So far, several varying timelines of lung cancer care have been introduced, yet all by national bodies only. At this stage, no evidence-based recommendations regarding waiting time cut-off-values can be made from an international perspective. Thus, we arbitrarily defined one individual upper time limit interval for each of the defined waiting time intervals related to lung cancer care in regular patients.
Yet, we were aware that special treatment situations (not considered as regular lung cancer care) might require different waiting time limits (i. e. urgent admissions during standard working times; emergency admissions anytime 24/7).

Comparison: Longer time of diagnosis to treatment (e.g. exceeding the time period specified by Aarhus staging

Outcomes considered important or critical for decision-making and included in the GRADE evidence profile: overall survival, progression-free survival, disease-free survival, mortality, morbidity, accuracy of staging, pathological confirmation, receipt of curative treatment, receipt of any active tumour-specific treatment, quality of life, patient satisfaction, performance status

Setting: Health care systems

2. Does the involvement of MDT or certain discipline in lung cancer care have an impact on the outcome in lung cancer?

PICO question 2: *In patients with lung cancer (or those suspected of having lung cancer), should a multi-disciplinary team (MDT) or certain disciplines be involved during lung cancer care rather than no involvement of an MDT or certain disciplines during lung cancer care?*

Population: adult lung cancer patients or those suspected of having lung cancer

Intervention: involvement of an MDT or oncology nurses during lung cancer care

Remarks: We have defined multi-disciplinary team (MDT) care according to the statement paper of the Metropolitan Health and Aged Care Services Division (Melbourne, Victoria, Australia) broadly as 'an integrated team approach to health care in which medical and allied health care professionals consider all relevant treatment options and develop an individual treatment plan for each patient collaboratively'. Specifically, the Task Force members opted for the following disciplines as essential constituents of an MDT in lung cancer care:

- respiratory medicine
- pathology
- radiology
- thoracic surgery
- radiotherapy
- oncology
- oncology nurse

The Task Force panel has adopted the definition of the National Cancer Institute (USA) for an Oncology Nurse: 'nurse who specializes in treating and caring for people who have cancer.'

Comparison: non-involvement of an MDT or oncology nurses during lung cancer care

Outcomes considered important or critical for decision-making and included in the GRADE evidence profile: overall survival, progression-free survival, disease-free survival, mortality, morbidity, accuracy of staging, pathological confirmation, receipt of curative treatment, receipt of any active tumour-specific treatment, quality of life, patient satisfaction, performance status

Setting: Health care system

3. Should guidelines or standard operating procedures (SOP) be used in lung cancer care?

PICO question 3: *In patients with lung cancer (or those suspected of having lung cancer), should guidelines or standard operating procedures (SOP) for lung cancer care be implemented or adhered to rather than non-implementation of or non-adherence to these guidelines or standard operating procedures?*

Population: adult lung cancer patients or those suspected of having lung cancer

Intervention: implementation of or adherence to guidelines or standard operating procedures (SOP) for lung cancer care

Comparison: non-implementation of or non-adherence to guidelines or standard operating procedures (SOP) for lung cancer care

Outcomes considered important or critical for decision-making and included in the GRADE evidence profile: overall survival, progression-free survival, disease-free survival, mortality, morbidity, accuracy of staging, pathological confirmation, receipt of curative treatment, receipt of any active tumour-specific treatment, quality of life, patient satisfaction, performance status

Setting: Health care system

4. Does hospital/professional volume of care/specialization have an impact in lung cancer diagnostics or therapy?

PICO question 4: *Should patients with lung cancer patients (or those suspected of having lung cancer) receive lung cancer-specific diagnostic or therapeutic procedures in hospitals/from professionals with higher volumes of activity/with a higher grade of specialization for these procedures rather than (compared to receiving them in hospitals/from professionals with lower volumes of activity/with lower grade of specialization for these procedures)?*

4a) Does hospital volume of activity have an impact in lung cancer diagnostics or therapy?

PICO question 4a: *Should patients with lung cancer (or those suspected of having lung cancer) receive lung cancer-specific diagnostic or therapeutic procedures in hospitals with higher volumes of activity for these procedures rather than receiving them in hospitals with lower volumes of activity for these procedures?*

Population: adult lung cancer patients or those suspected of having lung cancer
Subgroups*:* according to diagnostic or therapeutic procedures
- surgical resection of lung cancer
- other diagnostic and therapeutic modalities

Intervention: lung cancer-specific diagnostic or therapeutic procedure received in hospitals with higher volumes of activity for this procedure

Comparison: lung cancer-specific diagnostic or therapeutic procedure received in hospitals with lower volumes of activity for this procedure

Outcomes considered important or critical for decision-making and included in the GRADE evidence profile: overall survival, progression-free survival, disease-free survival, mortality, morbidity, accuracy of staging, pathological confirmation, receipt of curative treatment, receipt of any active tumour-specific treatment, quality of life, patient satisfaction, performance status

Setting: Health care system

4b) Does hospital specialization have an impact in lung cancer diagnostics or therapy?

PICO question 4b: *Should patients with lung cancer (or those suspected of having lung cancer) receive lung cancer-specific diagnostic or therapeutic procedures in hospitals with a higher grade of specialization for these procedures rather than receiving them hospitals with lower grade of specialization for these procedures?*

Population: adult lung cancer patients or those suspected of having lung cancer
Subgroups: according to diagnostic or therapeutic procedures
- surgical resection of lung cancer
- other diagnostic and therapeutic modalities

Intervention: lung cancer-specific diagnostic or therapeutic procedure received in hospitals with a higher grade of specialization for this procedure

Comparison: lung cancer-specific diagnostic or therapeutic procedure received in hospitals with a lower grade of specialization for this procedure

Outcomes considered important or critical for decision-making and included in the GRADE evidence profile: overall survival, progression-free survival, disease-free survival, mortality, morbidity, accuracy of staging, pathological confirmation, receipt of curative treatment, receipt of any active tumour-specific treatment, quality of life, patient satisfaction, performance status

Setting: Health care system

4c) Does surgeon and other professional volume of activity have an impact in lung cancer diagnostics or therapy?

PICO question 4c: *Should patients with lung cancer (or those suspected of having lung cancer) receive lung cancer-specific diagnostic or therapeutic procedures from surgeons and other professionals with higher volumes of activity for these procedures rather than receiving them from surgeons and other professionals with lower volumes of activity for these procedures?*

Population: adult lung cancer patients or those suspected of having lung cancer
Subgroups: according to diagnostic or therapeutic procedures

- surgical resection of lung cancer
- other diagnostic and therapeutic modalities

Intervention: lung cancer-specific diagnostic or therapeutic procedure received from surgeons and other professionals with higher volumes of activity for this procedure

Comparison: lung cancer-specific diagnostic or therapeutic procedure received from surgeons and other professionals with higher volumes of activity for this procedure

Outcomes considered important or critical for decision-making and included in the GRADE evidence profile: overall survival, progression-free survival, disease-free survival, mortality, morbidity, accuracy of staging, pathological confirmation, receipt of curative treatment, receipt of any active tumour-specific treatment, quality of life, patient satisfaction, performance status

Setting: Health care system

4d) Does surgeon and other professional specialization have an impact in lung cancer diagnostics or therapy?

PICO question 4d: *Should patients with lung cancer (or those suspected of having lung cancer) receive lung cancer-specific diagnostic or therapeutic procedures from surgeons and other professionals with a higher grade of specialization for these procedures rather than receiving them from surgeons and other professionals with a lower grade of specialization for these procedures?*

Population: adult lung cancer patients or those suspected of having lung cancer
Subgroups: according to diagnostic or therapeutic procedures

- surgical resection of lung cancer
- other diagnostic and therapeutic modalities

Intervention: lung cancer-specific diagnostic or therapeutic procedure received from surgeons with a higher grade of specialization for these procedures

Comparison: lung cancer-specific diagnostic or therapeutic procedure received from surgeons with a higher grade of specialization for these procedures

Outcomes considered important or critical for decision-making and included in the GRADE evidence profile: overall survival, progression-free survival, disease-free survival, mortality, morbidity, accuracy of staging, pathological confirmation, receipt of curative treatment, receipt of any active tumour-specific treatment, quality of life, patient satisfaction, performance status

Setting: Health care system

5. Should pathological confirmation of tumours or subtyping of lung cancers be obtained in lung cancer patients?

PICO question 5: *In lung cancer patients (or those suspected of having lung cancer), should pathological confirmation of tumours or subtyping of lung cancers be obtained rather than (compared to no attempted pathological confirmation of tumours or subtyping of lung cancers?*

Population: adult lung cancer patients or those suspected of having lung cancer

Intervention: pathological confirmation of tumours or subtyping of lung cancers
Subgroups: according to kind of lung cancer subtyping

- SCLC vs. NSCLC
- Subtyping of NSCLC
- Application of new WHO lung cancer classification for adenocarcinoma

Comparison: no attempted pathological confirmation tumours or subtyping of lung cancers

Outcomes considered important or critical for decision-making and included in the GRADE evidence profile: overall survival, progression-free survival, disease-free survival, mortality, morbidity, accuracy of staging, pathological confirmation, receipt of curative treatment, receipt of any active tumour-specific treatment, quality of life, patient satisfaction, performance status

Setting: Health care system

6. Should palliative care or palliative care specialists be included early in lung cancer care?

Search question: *In patients with lung cancer(or those suspected of having lung cancer), should palliative care or its delivery by specialists be integrated into lung cancer care already early during the course of the disease rather than no integration of palliative care or no palliative care delivery by specialists?*

Population: adult lung cancer patients

Intervention: integration of palliative care or its deliverance by specialists into lung cancer care early during the disease course

Comparison: no integration of palliative care or no palliative care delivery by specialists early during the disease course

Outcomes considered important or critical for decision-making and included in the GRADE evidence profile: overall survival, progression-free survival, disease-free survival, mortality, morbidity, accuracy of staging, pathological confirmation, receipt of curative treatment, receipt of any active tumour-specific treatment, quality of life, patient satisfaction, performance status

Setting: Health care system

7. Should quality improvement measures be applied in lung cancer care?

Search question: *In patients with lung cancer (or those suspected of having lung cancer), should quality improvement measures be applied in lung cancer care rather than no application of these methods in lung cancer care?*

Population: adult lung cancer patients

Intervention: application of quality assurance methods in lung cancer care
Subgroups: according to specification of quality improvement measures

- cancer registries and quality indicators
- specialized lung cancer services
- individual quality improvement measures
- audits/quality indicator systems

Comparison: no application of quality assurance methods in lung cancer care

Outcomes considered important or critical for decision-making and included in the GRADE evidence profile: overall survival, progression-free survival, disease-free survival, mortality, morbidity, accuracy of staging, pathological confirmation, receipt of curative treatment, receipt of any active tumour-specific treatment, quality of life, patient satisfaction, performance status

Setting: Health care system

8. Should patient decision tools be involved in the decision making in lung cancer?

Search question: *In patients with lung cancer (or those suspected of having lung cancer), should patient decision tools be involved in the decision-making and -sharing process rather than not involving them?*

Population: adult lung cancer patients

Intervention: involving patients in the decision-making process

Comparison: not involving patients in the decision-making process

Outcomes considered important or critical for decision-making and included in the GRADE evidence profile: overall survival, progression-free survival, disease-free survival, mortality, morbidity, accuracy of staging, pathological confirmation, receipt of curative treatment, receipt of any active tumour-specific treatment, quality of life, patient satisfaction, performance status

Setting: Health care system

Rating of outcomes for PICO questions 1-8

For the rating of outcomes by the task force members the following rating scale was used to assess the importance for clinical decision making of outcome parameters:
1-3 points: limited importance
4-6 points: important
7-9 points: critical

The rating results for each of the eight PICO questions are listed in **Table 3**.

	Outcome parameters	**PICO 1**	**PICO 2**	**PICO 3**	**PICO 4**	**PICO 5**	**PICO 6**	**PICO 7**	**PICO 8**
1	Overall survival	**9***	**9***	**9***	**8***	**9***	**7***	**8***	**7**
2	Progression-free survival	**8**	**5**	**6**	**7**	**8***	**5**	**6**	**7**
3	Disease-free survival	**8**	**5**	**6**	**7**	**8**	**5**	**7**	**7**
4	Mortality	**9***	**9***	**9***	**8***	**9**	**7**	**8***	**6**
5	Morbidity	**6**	**8**	**8**	**8**	**7**	**6**	**5**	**6**
6	Accuracy of staging	**5***	**6***	**8**	**7***	**7**	**6**	**7***	**5**
7	Pathological confirmation	**6**	**7***	**6**	**6***	**7**	**6**	**8***	**5**
8	Receipt of curative treatment	**8**	**8***	**8**	**7***	**7***	**6**	**9***	**7**
9	Receipt of any active tumour-specific treatment	**7**	**4***	**7**	**5**	**7**	**7***	**8***	**8**
10	Quality of Life	**6**	**5***	**7**	**6**	**7**	**9***	**7**	**9**
11	Patient satisfaction	**6**	**7***	**6**	**7**	**7**	**9***	**7**	**9***
12	Performance status	**7**	**8**	**8**	**7**	**6**	**9**	**5**	**8**

Table 3: Results of rating of outcomes for each PICO question (rating scale: 1-3 points - limited importance; 4-6 points – important; 7-9 points – critical; *outcomes that were actually selected for respective PICOs based on appropriate evidence)

Search strategy Medline for PICO questions 1-8

Database : Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) <1946 to January 5th, 2021

Search for P [lung neoplasms]

lung neoplasms/ or bronchial neoplasms/ or carcinoma, bronchogenic/ or carcinoma, non-small-cell lung/ or small cell lung carcinoma/ or pancoast syndrome/ or pulmonary blastoma/ or lung neoplasm*.ti,ab. or lung cancer*.ti,ab. or lung carcinoma*.ti,ab. or lung tumour*.ti,ab. or lung tumor*.ti,ab. or pulmonary neoplasm*.ti,ab. or pulmonary cancer*.ti,ab. or pulmonary carcinoma*.ti,ab. or pulmonary tumour*.ti,ab. or pulmonary tumor*.ti,ab. or bronchial neoplasm*.ti,ab. or bronchial cancer*.ti,ab. or bronchial carcinoma*.ti,ab. or bronchial tumour*.ti,ab. or bronchial tumor*.ti,ab. or bronchogenic neoplasm*.ti,ab. or bronchogenic cancer*.ti,ab. or bronchogenic carcinoma*.ti,ab. or bronchogenic tumour*.ti,ab. or bronchogenic tumor*.ti,ab. or pancoast* syndrome*.ti,ab. or pancoast* tumor*.ti,ab. or pancoast* tumour*.ti,ab. or ((lung.ti,ab or pulmonary.ti,ab) and (cancer*.ti,ab OR neoplasms/))

1. Do waiting times have an impact on outcome in lung cancer?

I

time factor*.ti,ab OR diagnosis delay*.ti,ab OR diagnostic delay*.ti,ab OR care delay*.ti,ab OR referral delay*.ti,ab OR treatment delay*.ti,ab OR therapeutic delay*.ti,ab OR delay* in diagnos*.ti,ab OR delay* of diagnosis.ti,ab OR wait time*.ti,ab OR time to diagnosis.ti,ab OR delayed initiation*.ti,ab OR consultation delay*.ti,ab OR travel time*.ti,ab OR delay* of treatment.ti,ab OR delay* to surgery.ti,ab OR delay* diagnosis.ti,ab OR doctor* delay*.ti,ab OR timeliness of diagnosis.ti,ab OR delay* in the diagnosis.ti,ab OR timing of referral.ti,ab OR waiting.ti,ab OR delay* in the referral.ti,ab OR timely care.ti,ab OR delay* cancer treatment*.ti,ab OR timeliness of care.ti,ab OR time before consulting.ti,ab OR delay* in assessment.ti,ab

O

Hospital mortality/ OR Mortality/ OR Survival/ OR Survival rate/ OR Disease-Free Survival/ OR "Quality of Life"/ OR Patient Satisfaction/ OR exp Intraoperative Complications/ OR exp Postoperative Complications/ OR complication*.ti,ab OR survival.ti,ab OR mortality.ti,ab OR quality of life.ti,ab OR patient* satisfaction.ti,ab OR morbidit*.ti,ab OR Treatment Outcome/ OR treatment outcome*.ti,ab

2. Does the involvement of MDT or certain discipline in lung cancer care have an impact on the outcome in lung cancer?

I

interdisciplinary communication/ OR multidisciplinary lung cancer team*.ti,ab OR multidisciplinary participation*.ti,ab OR multidisciplinary team*.ti,ab OR interdisciplinary team*.ti,ab OR interdisciplinary perspective*.ti,ab OR multidisciplinary perspective*.ti,ab OR interdisciplinary care.ti,ab OR multidisciplinary care.ti,ab OR multidisciplinary approach*.ti,ab OR interdisciplinary approach*.ti,ab OR multidisciplinary management.ti,ab OR interdisciplinary management.ti,ab OR multidisciplinary meeting*.ti,ab OR multidisciplinary clinic*.ti,ab OR interdisciplinary end of life care*.ti,ab OR multidisciplinary conference*.ti,ab OR multisciplinary oncology.ti,ab OR interdisciplinary collaboration*.ti,ab OR multidisciplinary lung cancer clinic*.ti,ab OR integrative practice*.ti,ab OR integrative medicine.ti,ab OR nursing-led intervention*.ti,ab OR nurse led follow up.ti,ab OR educational intervention*.ti,ab OR educational session*.ti,ab OR Nurse role/ OR Nurse-Patient Relations/ OR psycho-oncological.ti,ab OR Nutritionists/ OR nutritionist*.ti,ab OR dietician*.ti,ab OR psychologist*.ti,ab OR social workers/ OR social worker*.ti,ab OR Pastoral care/ OR Spirituality/ OR spiritual care worker*.ti,ab

O

Hospital mortality/ OR Mortality/ OR Survival/ OR Survival rate/ OR Disease-Free Survival/ OR "Quality of Life"/ OR Patient Satisfaction/ OR exp Intraoperative Complications/ OR exp Postoperative Complications/ OR complication*.ti,ab OR survival.ti,ab OR mortality.ti,ab OR quality of life.ti,ab OR patient* satisfaction.ti,ab OR morbidit*.ti,ab OR accuracy of staging.ti,ab OR accurate staging.ti,ab OR histological confirmation*.ti,ab OR histology confirmation*.ti,ab OR Treatment Outcome/ OR treatment outcome*.ti,ab

3. Should guidelines or standard operating procedures (SOP) be used in lung cancer care?
I
Practice guidelines as topic/ OR standard operating procedure*.ti,ab OR clinical recommendations.ti,ab OR practice guideline*.ti,ab OR management guideline*.ti,ab OR care guideline*.ti,ab OR treatment guideline*.ti,ab
O
Hospital mortality/ OR Mortality/ OR Survival/ OR Survival rate/ OR Disease-Free Survival/ OR "Quality of Life"/ OR Patient Satisfaction/ OR exp Intraoperative Complications/ OR exp Postoperative Complications/ OR complication*.ti,ab OR survival.ti,ab OR mortality.ti,ab OR quality of life.ti,ab OR patient* satisfaction.ti,ab OR morbidit*.ti,ab OR accuracy of staging.ti,ab OR accurate staging.ti,ab OR histological confirmation*.ti,ab OR histology confirmation*.ti,ab OR Treatment Outcome/ OR treatment outcome*.ti,ab

4. Does hospital/individual volume of activity/specialization have an impact in lung cancer diagnostics or therapy?
I
hospital volume*.ti,ab or high* volume hospital*.ti,ab or low volume hospital*.ti,ab or hospital procedure volume*.ti,ab or surgeon volume*.ti,ab or volume-outcome relationship*.ti,ab or operative volume.ti,ab or high* volume center*.ti,ab or low volume center*.ti,ab or surgical volume*.ti,ab or number of procedures performed.ti,ab or institutional experience.ti,ab

5. Should pathological confirmation of tumours or subtyping of lung cancers be obtained in lung cancer patients?
I
(histological confirmation.ti,ab OR histology confirmation.ti,ab OR histological classification.ti,ab OR histology diagnosis.ti,ab OR pathological confirmation.ti,ab OR diagnosed histologically.ti,ab) OR ((EGFR OR epidermal growth factor receptor OR EGF receptor* OR erbB 1).ti,ab. AND (Mutation/ OR mutation*.ti,ab.) AND (guideline* or documentation* or recommendation*).ti,ab) OR (ALK Translocation*.ti,ab OR ALK rearrangement*.ti,ab OR ALK fusion*.ti,ab OR ALK testing.ti,ab)
O
Hospital mortality/ or Mortality/ or Survival/ or Survival rate/ or Disease-Free Survival/ or survival.ti,ab. or mortality.ti,ab. or "Quality of Life"/ or quality of life.ti,ab

6. Should palliative care or palliative care specialists be included early in lung cancer care?
I
(Palliative care/ OR palliative care.ti,ab OR Terminal care/) AND (integration.ti,ab OR integrating.ti,ab OR integrated.ti,ab OR introducing.ti,ab OR general ward*.ti,ab OR early palliative care.ti,ab OR interdisciplinary palliative care.ti,ab OR palliative care intervention.ti,ab)

7. Should quality improvement measures be applied for lung cancer patients?
I
"Quality of Health Care"/ OR Quality Assurance, Health Care/ OR Quality Indicators, Health Care/ OR Patient Care Management/ OR Benchmarking/ OR Clinical audit/ OR Medical audit/ OR Certification/ OR "Outcome and Process Assessment (Health Care)"/ OR Peer Review, Health Care/ OR "Organization and administration"/ OR logistics.ti,ab OR supervision.ti,ab OR

administrative technics.ti,ab OR administrative technique*.ti,ab OR quality of healthcare.ti,ab OR quality of health care.ti,ab OR healthcare quality.ti,ab OR health care quality.ti,ab OR assessment* of quality.ti,ab OR quality assurance*.ti,ab OR quality assessment*.ti,ab OR quality measure*.ti,ab OR quality evaluation*.ti,ab OR quality apprais*.ti,ab OR performance measure*.ti,ab. OR quality indicator*.ti,ab OR certification program*.ti,ab OR benchmarking.ti,ab OR audits.ti,ab OR audit.ti,ab

O

Hospital mortality/ OR Mortality/ OR Survival/ OR Survival rate/ OR Disease-Free Survival/ OR "Quality of Life"/ OR Patient Satisfaction/ OR exp Intraoperative Complications/ OR exp Postoperative Complications/ OR complication*.ti,ab OR survival.ti,ab OR mortality.ti,ab OR quality of life.ti,ab OR patient* satisfaction.ti,ab OR morbidit*.ti,ab OR Treatment Outcome/ OR treatment outcome*.ti,ab

8. Should patient decision tools be involved in the decision making in lung cancer?

I

patient decision making.ti,ab OR shared decision making.ti,ab OR ((Patient participation/ OR patient* participation.ti,ab OR patient* involvement.ti,ab OR patient* engagement.ti,ab OR patient* empowerment.ti,ab OR engaging patient*.ti,ab OR involving patient*.ti,ab) AND (decision*.ti,ab OR choice*.ti,ab))

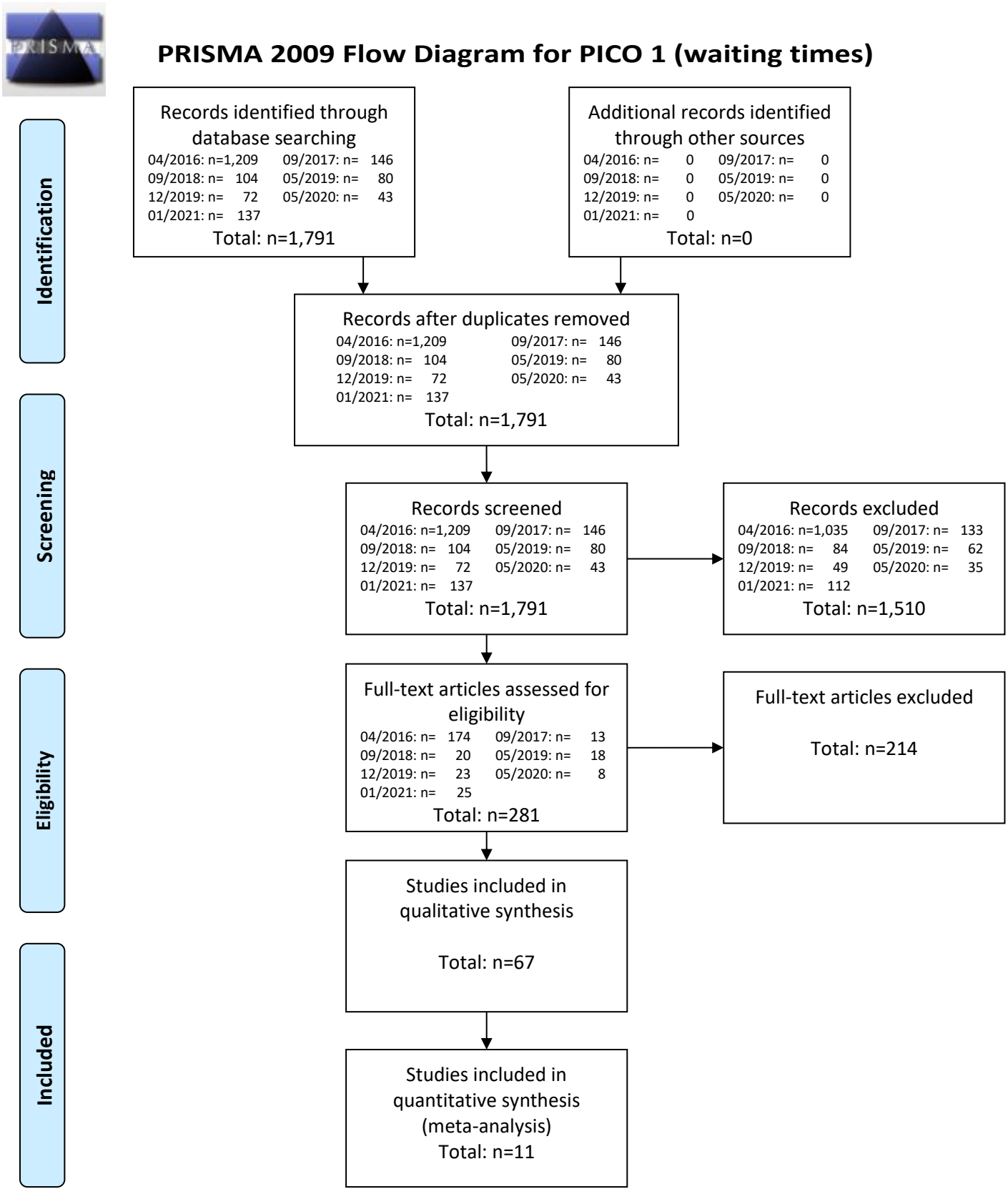
Eligibility criteria for included studies to inform PICO questions 1-8

Eligibility criteria for inclusion of studies to inform PICOs 1-8:

- study design: randomized controlled trials and comparative non-randomized studies including trials, observational cohort, and case-control studies
- study population: within the scope of the respective PICO question, yet mixed study populations are allowed if the pre-defined population of interest is included and separate data for lung cancer patients are available
- study interventions/controls: within the scope of the respective search question, yet additional study interventions/controls are allowed if the pre-defined interventions/controls of interest are included and separate data are available for the latter
- publication language: only languages fluently spoken by Task Force members which are English, French, Dutch, German, and Spanish
- publication period: no restrictions by Task Force panel, yet technically limited from 1946 (due to accessibility through the OvidSP interface) to January 5th, 2021 (latest search date)

PRISMA flow charts for PICO questions 1-8

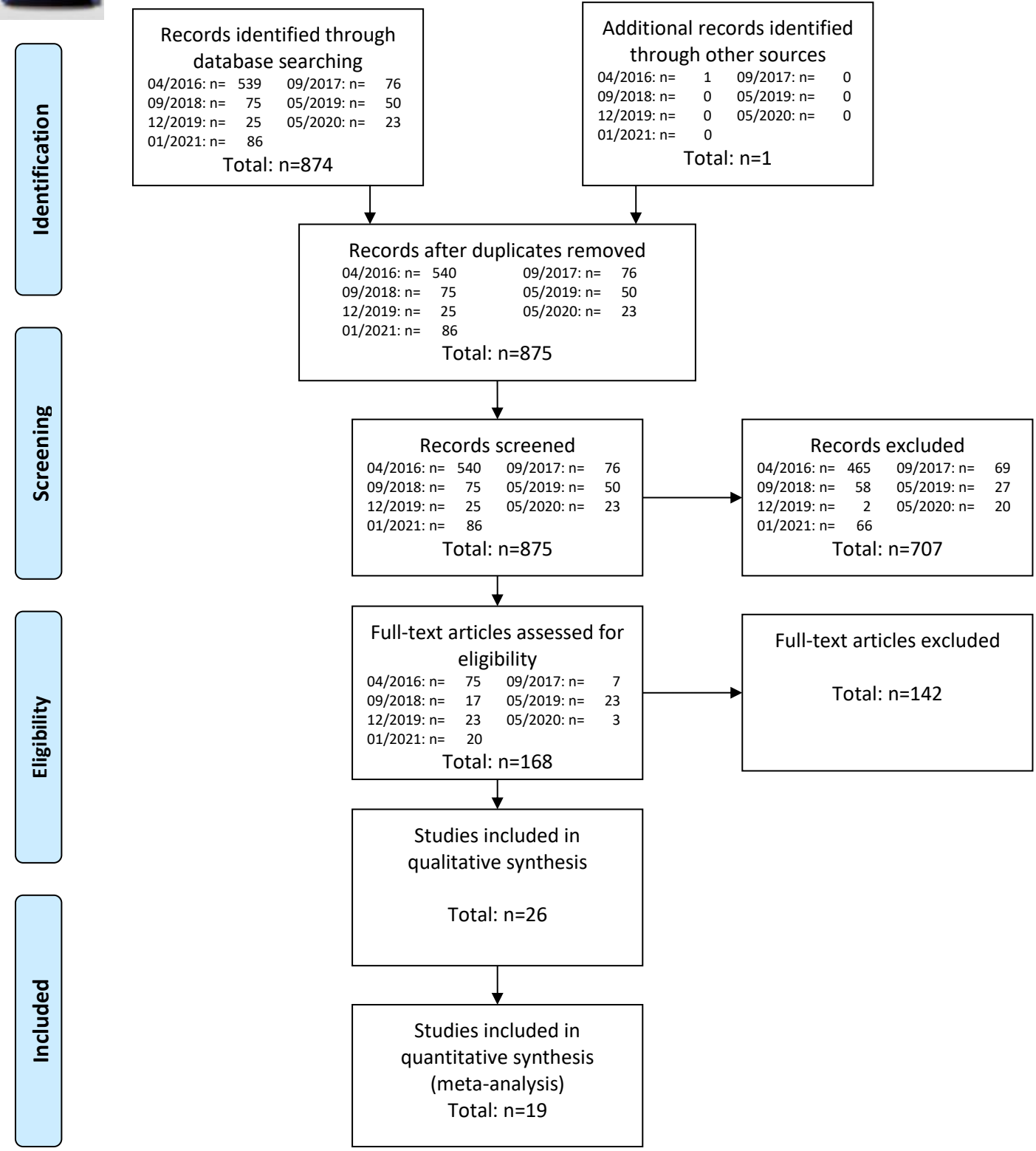
This sections includes the PRISMA flow diagrams for each of the eight PICO question



From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). *P*referred *R*eporting *I*tems for *S*ystematic Reviews and *M*eta-*A*nalyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit www.prisma-statement.org.

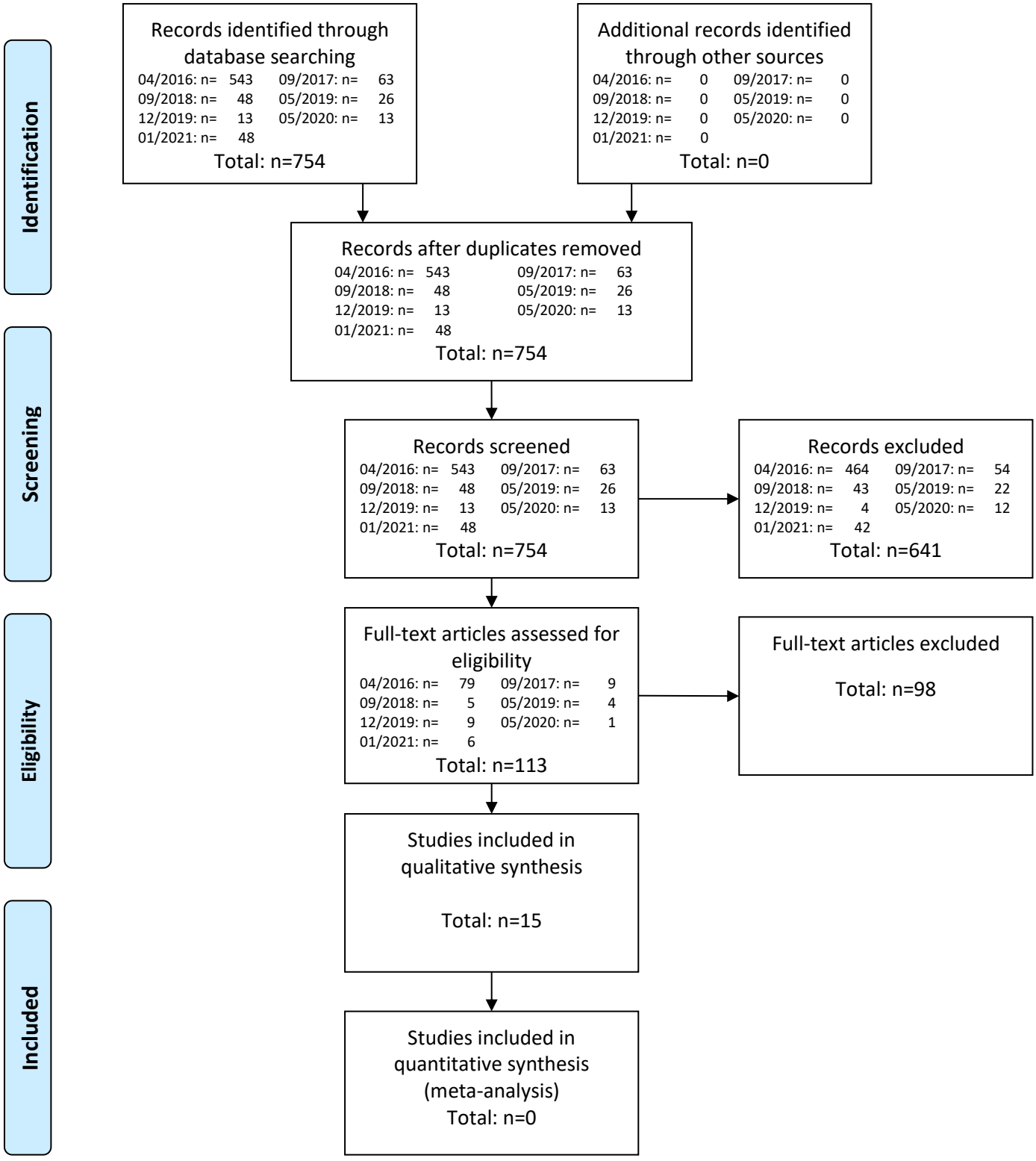
PRISMA 2009 Flow Diagram for PICO 2 (MDT)



From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). *P*referred *R*eporting *I*tems for *S*ystematic Reviews and *M*eta-*A*nalyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

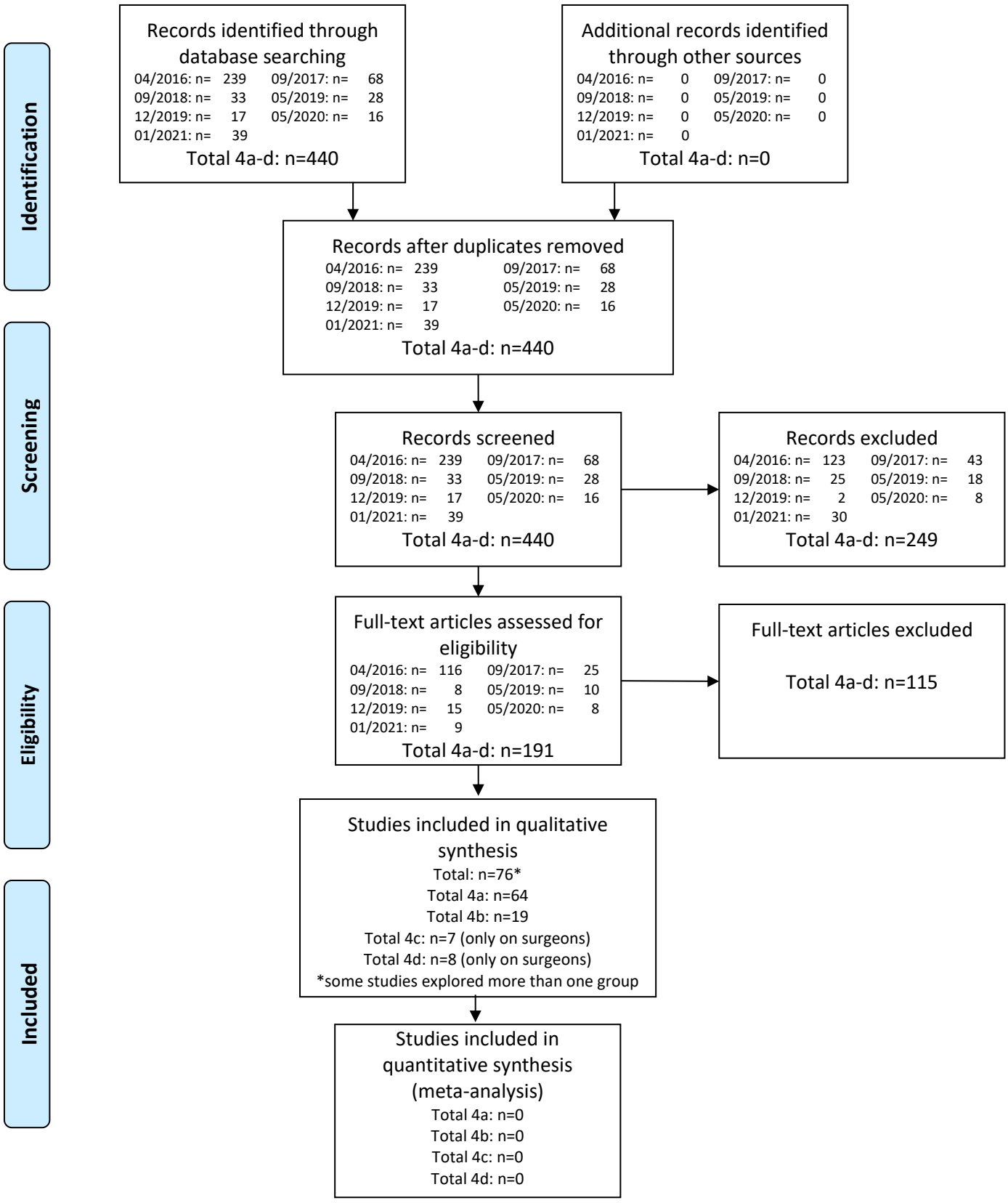
For more information, visit www.prisma-statement.org.

PRISMA 2009 Flow Diagram for PICO 3 (guideline implementation/adherence)



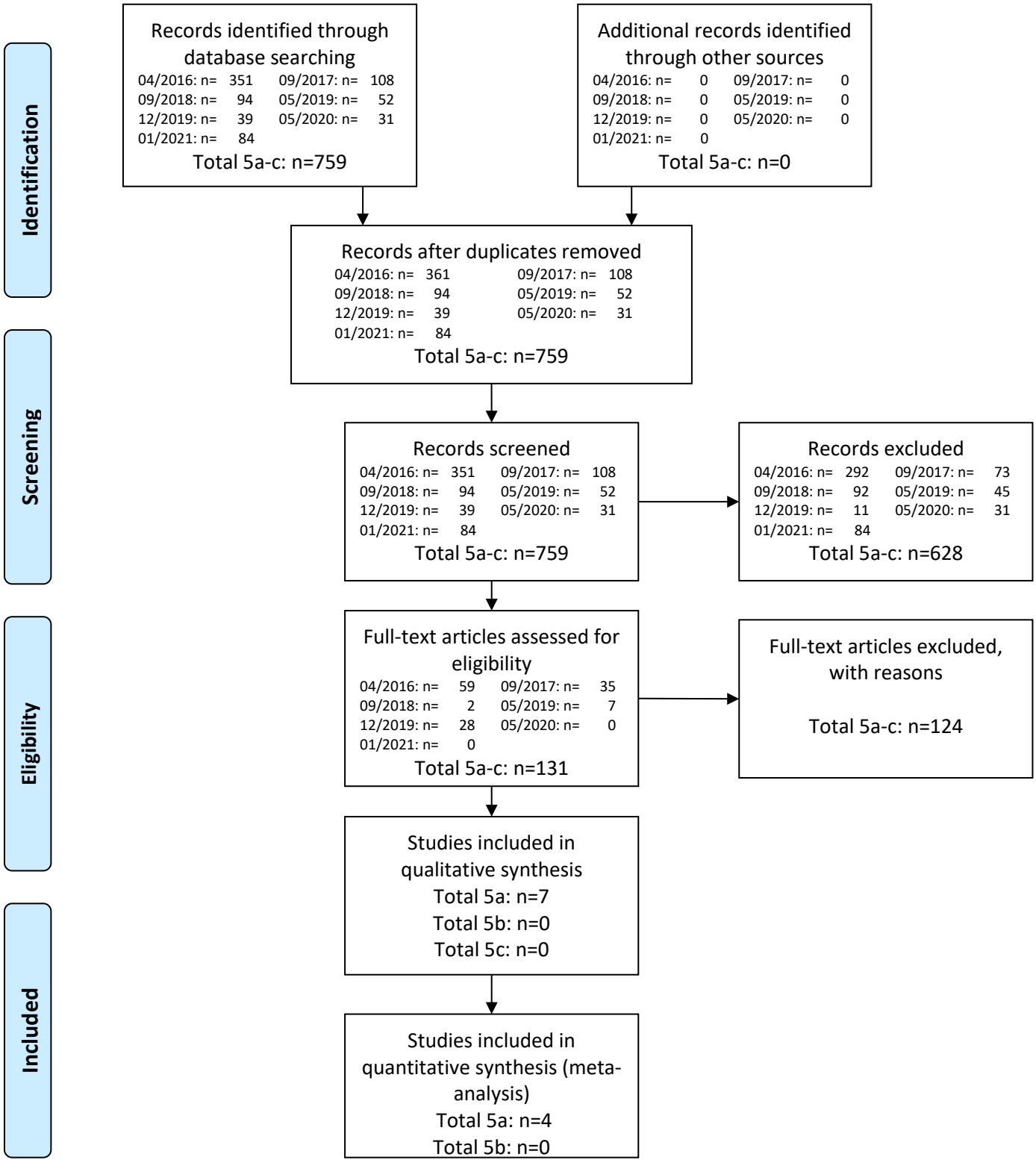
From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). *P*referred *R*eporting *I*tems for *S*ystematic Reviews and *M*eta-*A*nalyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

PRISMA 2009 Flow Diagram for PICO 4a-d
(hospital/professional volume of care [4a+c] and specialization [4b+d])



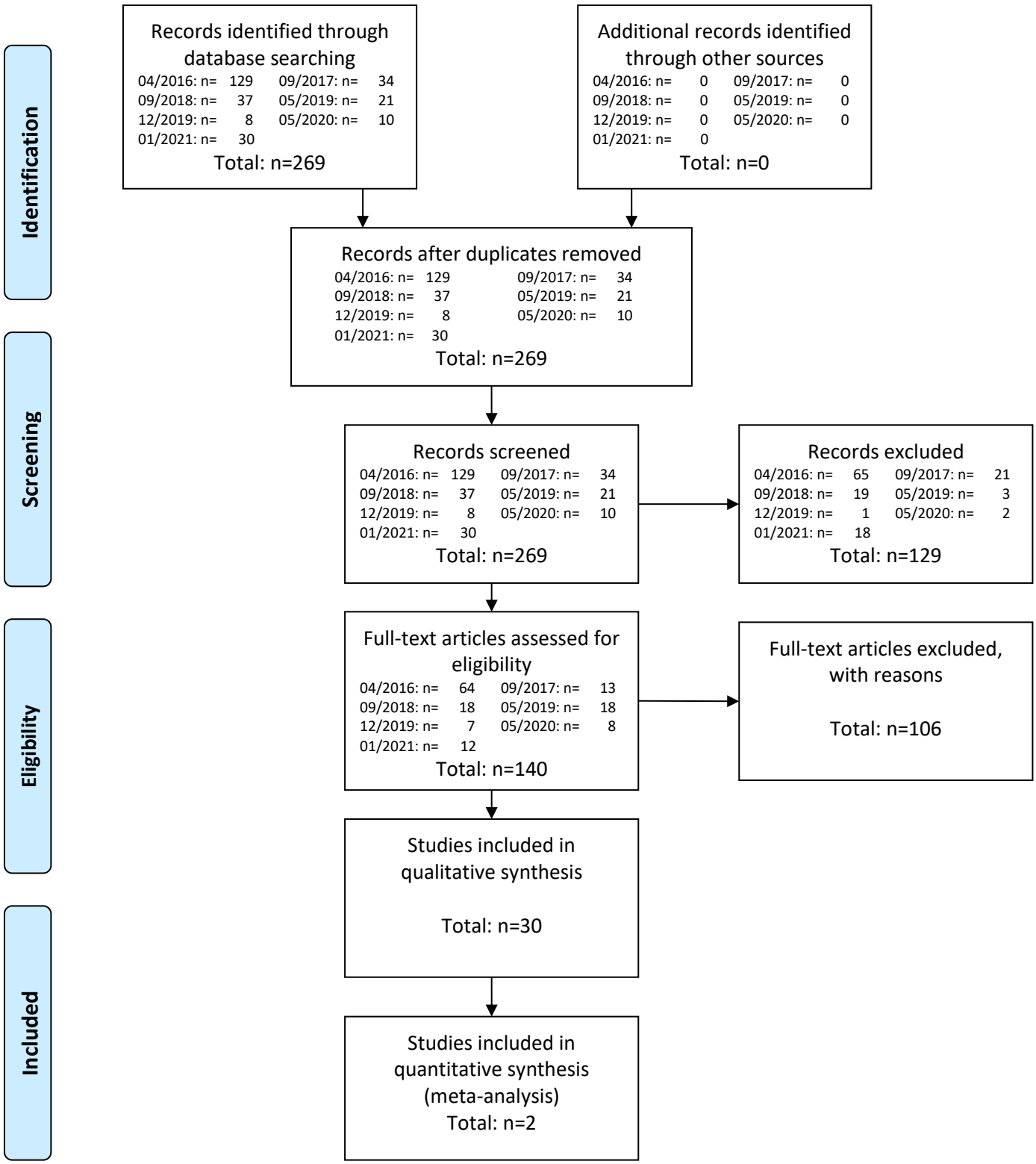
From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). *P*referred *R*eporting *I*tems for *S*ystematic Reviews and *M*eta-*A*nalyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

PRISMA 2009 Flow Diagram for PICO 5a-c (pathological confirmation of tumours [5a] and subtyping of lung cancers [5b+c])



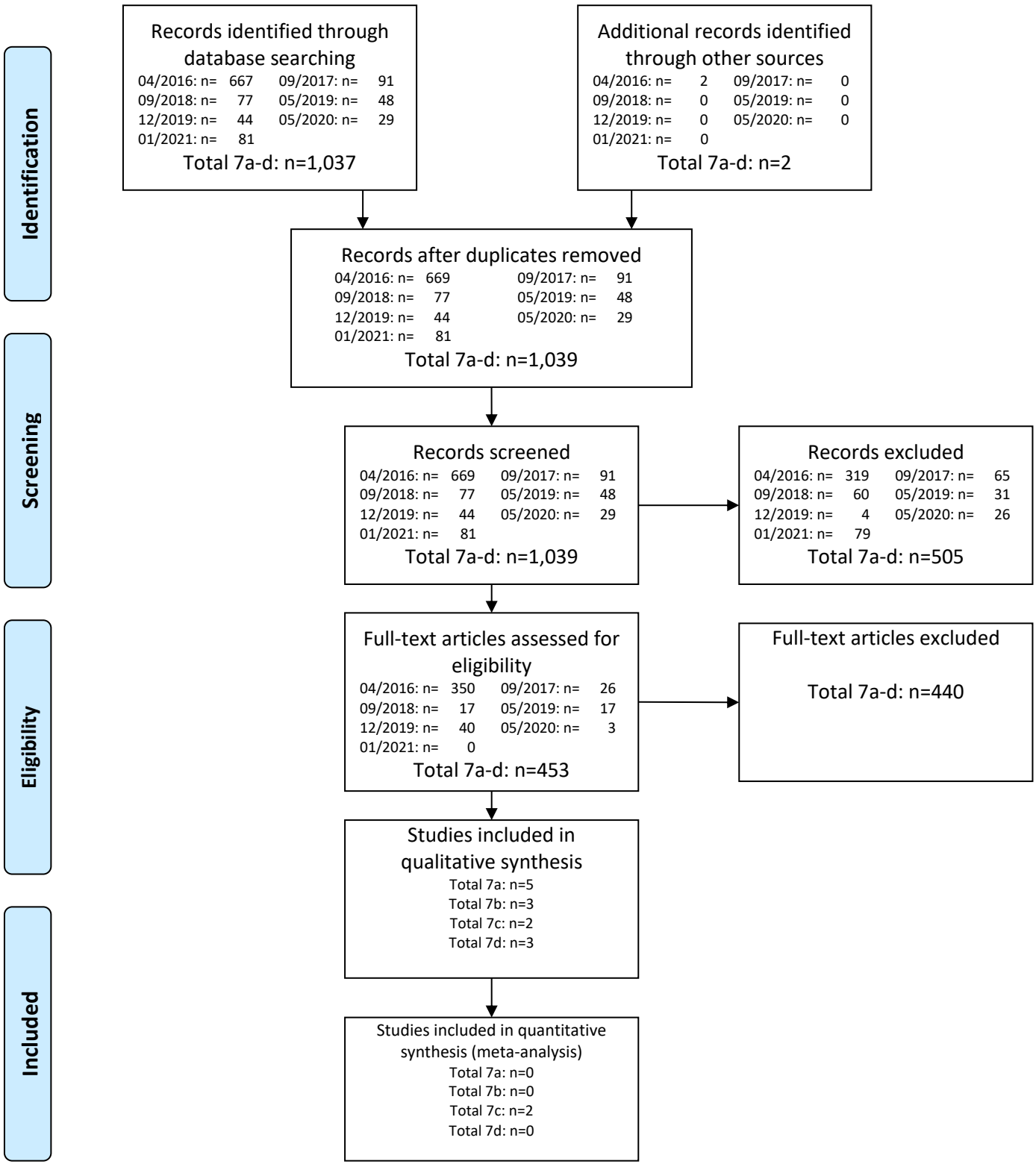
From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). *P*referred *R*eporting *I*tems for *S*ystematic Reviews and *M*eta-*A*nalyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

PRISMA 2009 Flow Diagram for PICO 6 (early integration of palliative care)



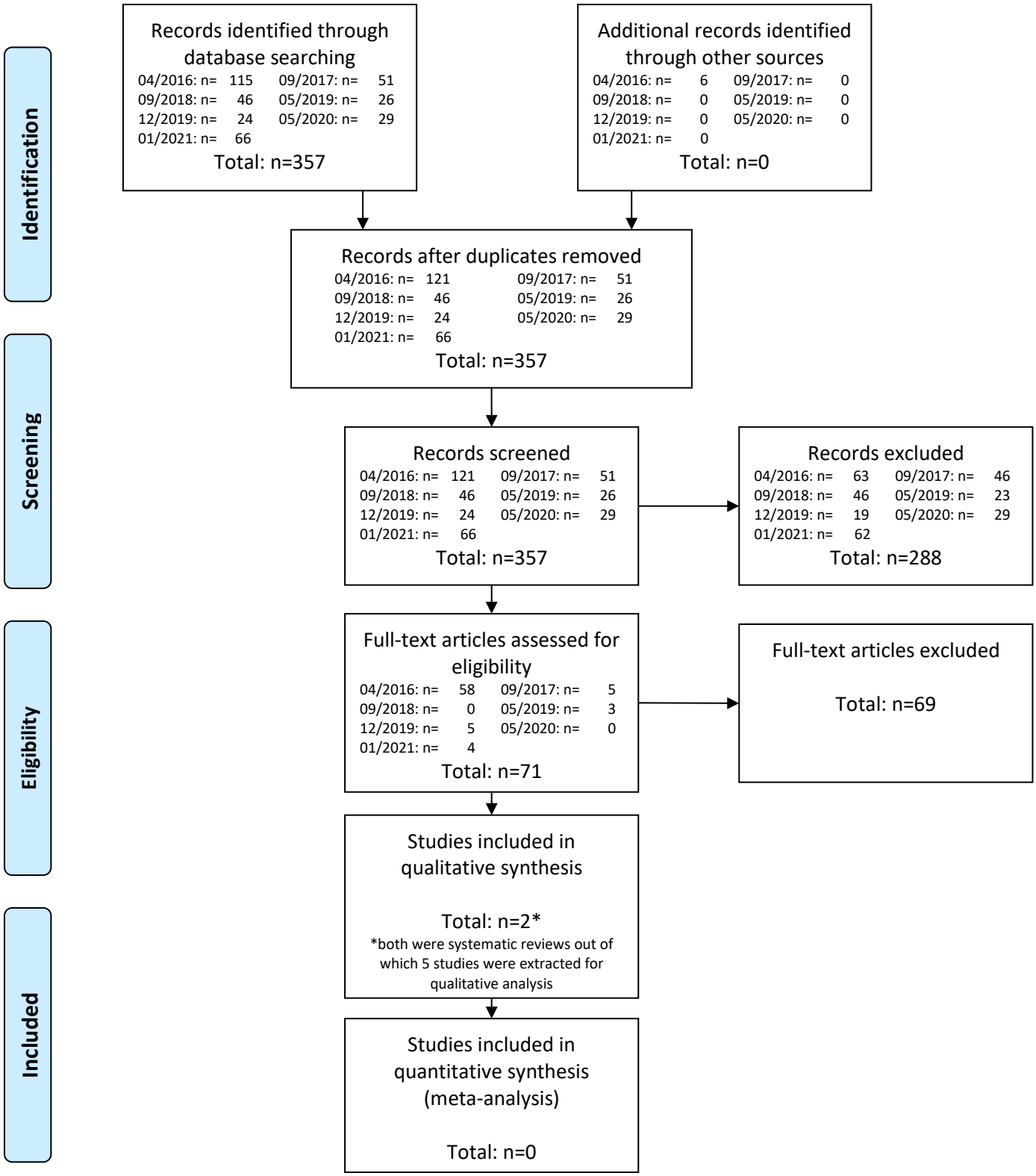
From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). *P*referred *R*eporting *I*tems for *S*ystematic Reviews and *M*eta-*A*nalyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

PRISMA 2009 Flow Diagram for PICO 7a-d (cancer registries and quality indicators [7a], specialized lung cancer services [7b], individual quality improvement measures [7c], audits/quality indicator systems [7d])



From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). *P*referred *R*eporting *I*tems for *S*ystematic Reviews and *M*eta-*A*nalyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

PRISMA 2009 Flow Diagram for PICO 8 (patient decision tools)



From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). *P*referred *R*eporting *I*tems for *S*ystematic Reviews and *M*eta-*A*nalyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit www.prisma-statement.org.

References

1. Asking Answerable Clinical Questions. *In:* Sackett D L SSE, Richardson W S, Rosenberg W, Haynes R B, ed. Evidence-Based Medicine How to Practice and Teach EBM, 2nd ed. Churchill Livingstone, Edinburgh, 2000; pp. 13–27.
2. Oxford ORHIUo. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA). 2015 [cited 27.11.2017]; Available from: http://www.prisma-statement.org/
3. Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, Savovic J, Schulz KF, Weeks L, Sterne JA, Cochrane Bias Methods G, Cochrane Statistical Methods G. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011: 343: d5928.
4. Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, Montori V, Akl EA, Djulbegovic B, Falck-Ytter Y, Norris SL, Williams JW, Jr., Atkins D, Meerpohl J, Schunemann HJ. GRADE guidelines: 4. Rating the quality of evidence--study limitations (risk of bias). *J Clin Epidemiol* 2011: 64(4): 407-415.
5. Blum TG, Rich A, Baldwin D, Beckett P, De Ruysscher D, Faivre-Finn C, Gaga M, Gamarra F, Grigoriu B, Hansen NC, Hubbard R, Huber RM, Jakobsen E, Jovanovic D, Konsoulova A, Kollmeier J, Massard G, McPhelim J, Meert AP, Milroy R, Paesmans M, Peake M, Putora PM, Scherpereel A, Schonfeld N, Sitter H, Skaug K, Spiro S, Strand TE, Taright S, Thomas M, van Schil PE, Vansteenkiste JF, Wiewrodt R, Sculier JP. The European initiative for quality management in lung cancer care. *Eur Respir J* 2014: 43(5): 1254-1277.
6. Santesso N, Glenton C, Dahm P, Garner P, Akl EA, Alper B, Brignardello-Petersen R, Carrasco-Labra A, De Beer H, Hultcrantz M, Kuijpers T, Meerpohl J, Morgan R, Mustafa R, Skoetz N, Sultan S, Wiysonge C, Guyatt G, Schünemann HJ. GRADE guidelines 26: informative statements to communicate the findings of systematic reviews of interventions. *J Clin Epidemiol* 2020: 119: 126-135.
7. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, Norris S, Falck-Ytter Y, Glasziou P, Debeer H. GRADE guidelines: 1. Introduction—GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 2011: 64(4): 383-394.
8. Alonso-Coello P, Oxman AD, Moberg J, Brignardello-Petersen R, Akl EA, Davoli M, Treweek S, Mustafa RA, Vandvik PO, Meerpohl J, Guyatt GH, Schunemann HJ, Group GW. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 2: Clinical practice guidelines. *BMJ* 2016: 353: i2089.
9. Alonso-Coello P, Schunemann HJ, Moberg J, Brignardello-Petersen R, Akl EA, Davoli M, Treweek S, Mustafa RA, Rada G, Rosenbaum S, Morelli A, Guyatt GH, Oxman AD, Group GW. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. *BMJ* 2016: 353: i2016.
10. Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, Devereaux PJ, Montori VM, Freyschuss B, Vist G, Jaeschke R, Williams JW, Jr., Murad MH, Sinclair D, Falck-Ytter Y, Meerpohl J, Whittington C, Thorlund K, Andrews J, Schunemann HJ. GRADE guidelines 6. Rating the quality of evidence--imprecision. *J Clin Epidemiol* 2011: 64(12): 1283-1293.
11. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, Alonso-Coello P, Glasziou P, Jaeschke R, Akl EA, Norris S, Vist G, Dahm P, Shukla VK, Higgins J, Falck-Ytter Y, Schunemann HJ, Group GW. GRADE guidelines: 7. Rating the quality of evidence--inconsistency. *J Clin Epidemiol* 2011: 64(12): 1294-1302.

12. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, Alonso-Coello P, Falck-Ytter Y, Jaeschke R, Vist G, Akl EA, Post PN, Norris S, Meerpohl J, Shukla VK, Nasser M, Schunemann HJ, Group GW. GRADE guidelines: 8. Rating the quality of evidence--indirectness. *J Clin Epidemiol* 2011: 64(12): 1303-1310.
13. Guyatt GH, Oxman AD, Montori V, Vist G, Kunz R, Brozek J, Alonso-Coello P, Djulbegovic B, Atkins D, Falck-Ytter Y, Williams JW, Jr., Meerpohl J, Norris SL, Akl EA, Schunemann HJ. GRADE guidelines: 5. Rating the quality of evidence--publication bias. *J Clin Epidemiol* 2011: 64(12): 1277-1282.
14. Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, Guyatt GH, Harbour RT, Haugh MC, Henry D, Hill S, Jaeschke R, Leng G, Liberati A, Magrini N, Mason J, Middleton P, Mrukowicz J, O'Connell D, Oxman AD, Phillips B, Schunemann HJ, Edejer T, Varonen H, Vist GE, Williams JW, Jr., Zaza S, Group GW. Grading quality of evidence and strength of recommendations. *BMJ* 2004: 328(7454): 1490.
15. Andrews J, Guyatt G, Oxman AD, Alderson P, Dahm P, Falck-Ytter Y, Nasser M, Meerpohl J, Post PN, Kunz R, Brozek J, Vist G, Rind D, Akl EA, Schünemann HJ. GRADE guidelines: 14. Going from evidence to recommendations: the significance and presentation of recommendations. *J Clin Epidemiol* 2013: 66(7): 719-725.
16. Andrews JC, Schunemann HJ, Oxman AD, Pottie K, Meerpohl JJ, Coello PA, Rind D, Montori VM, Brito JP, Norris S, Elbarbary M, Post P, Nasser M, Shukla V, Jaeschke R, Brozek J, Djulbegovic B, Guyatt G. GRADE guidelines: 15. Going from evidence to recommendation-determinants of a recommendation's direction and strength. *J Clin Epidemiol* 2013: 66(7): 726-735.
17. Guyatt GH, Alonso-Coello P, Schunemann HJ, Djulbegovic B, Nothacker M, Lange S, Murad MH, Akl EA. Guideline panels should seldom make good practice statements: guidance from the GRADE Working Group. *J Clin Epidemiol* 2016: 80: 3-7.