Online resource 1. Search strategies

Original search strategies used for MEDLINE, PREMEDLINE and Embase databases in September 2016 (Yu et al., 2019)

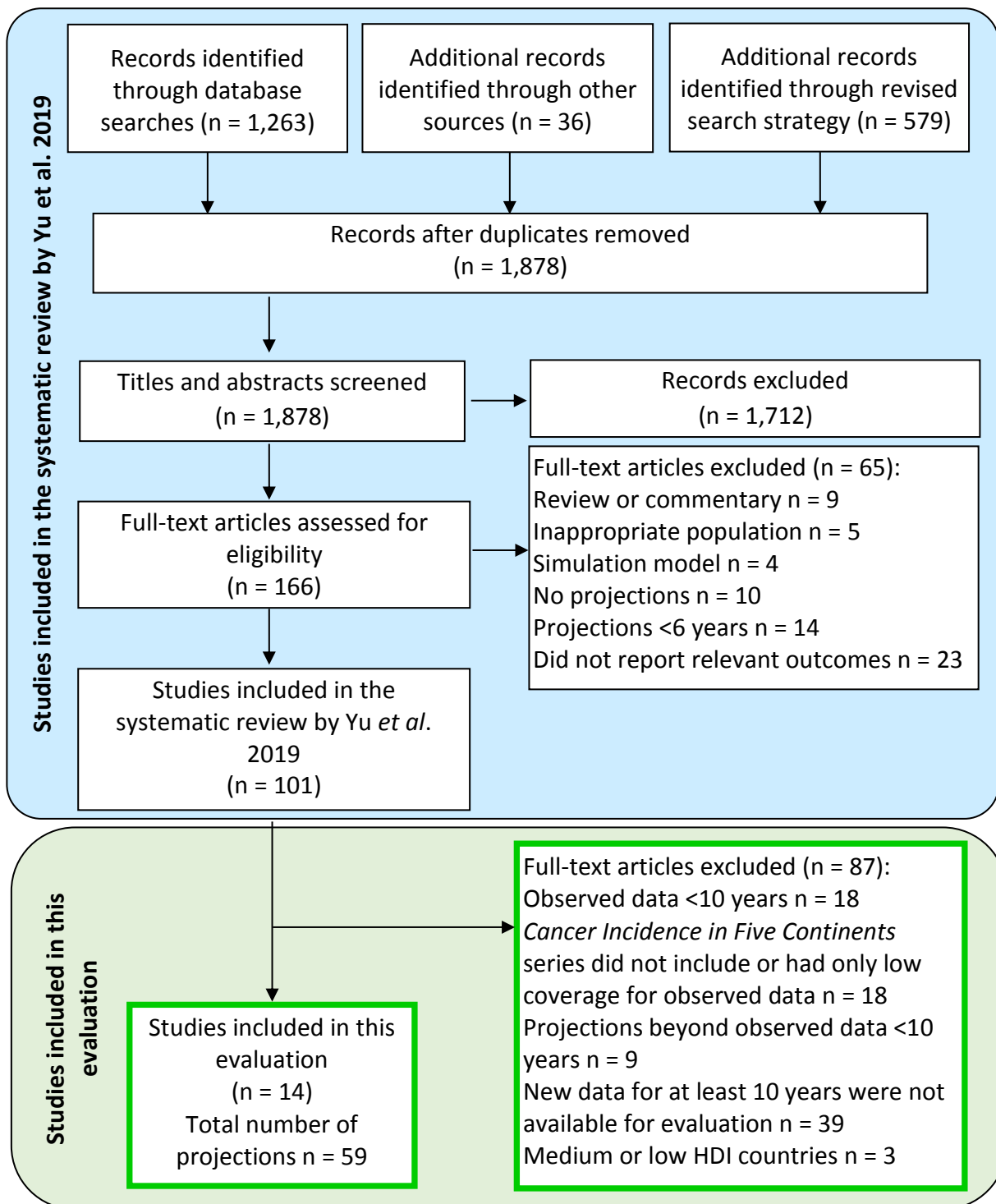| #	Searches
1	forecast$.mp.
2	extrapolat$.mp.
3	(projecte$ or projection$).mp.
4	*models,statistical/ or *likelihood functions/ or *linear models/ or *logistic models/
5	*population model/ or *prediction model/
6	exp forecasting/
7	1 or 2 or 3 or 4 or 5 or 6
8	exp lung Neoplasms/
9	(lung adj6 (cancer$ or carcinoma$ or malig$ or tumo?r$ or neoplas$ or adeno$)).mp.
10	(pulmon$ adj6 (cancer$ or carcinoma$ or malig$ or tumo?r$ or neoplas$ or adeno$)).mp.
11	8 or 9 or 10
12	7 and 11
13	tomograph$.ti. or tomograph$.ab.
14	(radiotherap$ or radiograph$ or radiation).mp.
15	(beam$ or radon or neuroendocrine or kinase or chemotherapy or genetic or genom$ or nodule$).mp.
16	13 or 14 or 15
17	12 not 16
18	limit 17 to english language
19	limit 18 to yr="1988 -Current"
20	limit 19 to (conference abstract or conference paper or "conference review" or editorial or letter or note or case reports or clinical conference or comment or congresses or consensus development conference or consensus development conference, nih or festschrift or guideline or practice guideline) [Limit not valid in Embase,Ovid MEDLINE(R),Ovid MEDLINE(R) In-Process; records were retained]
21	19 not 20
22	remove duplicates from 21

Revised search strategies used for Embase Ovid MEDLINE(R) and Epub Ahead of Print, In-Process & Other Non-Indexed Citations and Daily in August 2018 (Yu et al., 2019)

#	Searches
1	forecast$.mp.
2	extrapolat$.mp.
3	(projecte$ or projection$).mp.
4	(lung cancer adj4 predict*).tw.
5	(lung cancer adj4 to project).tw.
6	models,statistical/ or likelihood functions/ or linear models/ or logistic models/
7	population model/ or prediction model/
8	exp forecasting/
9	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8
10	exp lung Neoplasms/
11	(lung adj9 (cancer$ or carcinoma$ or malig$ or tumo?r$ or neoplas$ or adeno$)).mp.
12	(pulmon$ adj9 (cancer$ or carcinoma$ or malig$ or tumo?r$ or neoplas$ or adeno$)).mp.
13	10 or 11 or 12
14	9 and 13
15	tomograph$.ti. or tomograph$.ab.
16	(radiotherap$ or radiograph$ or radiation).mp.
17	(beam$ or radon or neuroendocrine or kinase or chemotherapy or genetic or genom$ or nodule$).mp.
18	15 or 16 or 17
19	14 not 18
20	limit 19 to english language
21	limit 20 to yr="2016 -Current"
22	limit 21 to (conference abstract or conference paper or "conference review" or editorial or letter or note or case reports or clinical conference or comment or congresses or consensus development conference or consensus development conference, nih or festschrift or guideline or practice guideline) [Limit not valid in Embase,Ovid MEDLINE(R),Ovid MEDLINE(R) Daily Update,Ovid MEDLINE(R) In-Process,Ovid MEDLINE(R) Publisher; records were retained]
23	21 not 22
24	remove duplicates from 23

Online resource 2: Criteria for inclusion and exclusion of studies

Domain	Inclusion criteria	Exclusion criteria
Study type	Population-based statistical projections beyond the observed data.	Studies did not provide projected rates beyond the original observed data. For example, some published methodological studies evaluated the performance of projection methods by withholding the most recent observed data from the projection model and so didn't include projected rates beyond the observation period.
Study population	General population in any country.	Restricted to selected groups, i.e. selected cancer patients or high-risk populations.
Outcomes	Reports projections of lung cancer incidence and/or mortality rates.	No relevant outcomes are reported, i.e. no lung cancer specific outcomes.
Statistical method	Uses a statistical method for the projection, including studies which used simulation methods to estimate confidence intervals, i.e. Bayesian technique.	Uses mathematical or micro-simulation models which generate outcomes based on a proposed theoretical model of the disease's natural history.
Publication type	Full text published.	Conference proceedings, abstracts, posters.
Time of publication	Published from 1st January 1988 to 14th August 2018.	Published before 1988.
Language	English	Language other than English
Data quality	Used at least 10 years of good quality observed data (e.g. incidence data included in the Cancer Incidence in Five Continents series (2014), or mortality data from a source considered by the World Health Organization's (WHO) Mortality Database to have a high population coverage (World Health Organization, 2017)) for countries ranked as high or very high on the Human Development Index (HDI) in 2016 (United Nations Development Programme (UNDP), 2018).	Used less than 10 years of observed data or the data source was reported but the majority of the observed data used are not included in Five Continents series (2014), or have low population coverage as stated in the WHO database, or there is insufficient information to make an assessment, or countries ranked as medium or low on the HDI in 2016 (United Nations Development Programme (UNDP), 2018).
Number of years projected	Projected future lung cancer rates for at least 10 years beyond the original observation period.	Does not report projections of lung cancer rates, i.e. only explains past trends, or reports projections for less than 10 years.
Number of years of projections evaluated	More recent observed rates for comparison were available for a minimum of 10 years from the beginning of the projection period.	Less than 10 years of available data for the evaluation period

Online resource 3: Inclusion and exclusion of studies for this evaluation study

Studies included in the systematic review by Yu et al. 2019

Records identified through database searches (n = 1,263)

Additional records identified through other sources (n = 36)

Additional records identified through revised search strategy (n = 579)

Records after duplicates removed (n = 1,878)

Titles and abstracts screened (n = 1,878)

Records excluded (n = 1,712)

Full-text articles assessed for eligibility (n = 166)

Full-text articles excluded (n = 65):
Review or commentary n = 9
Inappropriate population n = 5
Simulation model n = 4
No projections n = 10
Projections <6 years n = 14
Did not report relevant outcomes n = 23

Studies included in the systematic review by Yu *et al*. 2019 (n = 101)

Studies included in this evaluation

Studies included in this evaluation (n = 14)
Total number of projections n = 59

Full-text articles excluded (n = 87):
Observed data <10 years n = 18
Cancer Incidence in Five Continents series did not include or had only low coverage for observed data n = 18
Projections beyond observed data <10 years n = 9
New data for at least 10 years were not available for evaluation n = 39
Medium or low HDI countries n = 3

Reference:
Yu XQ, Luo Q, Hughes S, et al. Statistical projection methods for lung cancer incidence and mortality: a systematic review. *BMJ Open* 2019; **9**, e028497. doi:10.1136/bmjopen-2018-028497

Online resource 4: Studies included in the evaluation

First author and publication year	Country	Outcome	Observed data	Number of years for projection	Number of years for evaluation	Statistical projection method used	Software	Standard population
1. Brown 1988	USA	Mortality	1958-1982	45	35	GLM with smoking variable	No information	USA 1970
2. Negri 1990a	Italy	Mortality	1955-1984	15	15*	A and B: APC models. C: APC model including smoking effect.	GLIM	European
3. Negri 1990b	Switzerland	Mortality	1950-1984	15	15*	A and B: APC models. C: APC model including smoking effect.	GLIM	European
4. Kuroishi 1992	Japan	Mortality	1969-1989	21	21	Other GLMs	No information	Japanese
5. Engeland 1995	5 countries	Mortality	1958-1987	25	25*	Other GLMs	No information	Segi World
6. Hristova 1997	Bulgaria	Incidence	1968-1992	25	15*	APC model	GLIM	Segi World
7. Kubik 1998	2 countries	Mortality	1960-1989	20	20*	APC model	GLIM	Segi World
8. Moller 2002	5 countries	Incidence	1958-1997	25	15*	APC model	Nordpred	Segi World
9. Kaneko 2003	Japan	Mortality	1952-2001	30	15	APC model	WinBUGs	Japanese
10. O'Lorcain 2004	Ireland	Mortality	1954-2000	15	15	Other GLMs	STATA, Joinpoint	European
11. Shibuya 2005	4 countries	Mortality	1950-1999	35	15*	GLM with smoking variable	No information	WHO World
12. Byers 2006	USA	Mortality	1990-2002	13	12	Other GLMs	No information	USA 2000
13. Moller 2007	England	Incidence	1974-2003	20	10*	APC model	Nordpred	European
14. Eilstein 2008	France	Mortality	1978-2002	10	10	APC model	WinBUGs	Segi World

* For studies that only provided age-standardized rates (ASR) graphically, data were independently extracted by two authors (QL and JS) using the open source software program WebPlotDigitizer (Rohatgi 2018), which is considered to be a reliable tool for extracting data from figures (Drevon et al., 2017). The mean values of the two extractions were used in the analyses. The agreement between the two authors' data extractions was high: the absolute difference between the extracted ASRs ranged from 0.0002 to 1.9 per 100000.

USA: United States of America. GLM: generalised linear model. APC: age-period-cohort.

References:

Rohatgi A (2018). WebPlotDigitizer. Version: 4.1 [Online]. Austin, Texas, USA. Available: Website: https://automeris.io/WebPlotDigitizer [Accessed 5 Apr 2018].

Drevon D, Fursa SR, Malcolm AL (2017). Intercoder Reliability and Validity of WebPlotDigitizer in Extracting Graphed Data. Behav Modif, **41**, 323-39.