

Supplementary material

Appendix 1 Literature search strategy

Number of results before and after de-duplication

Database	Number of initial hits	After de-duplication
EBM Reviews	112	38
Embase	2260	1508
Medline	940	874
Scopus	2805	1512
Web of Science	1430	313
Totals	7547	4245

EBM Reviews

((digestive or gastr* or GI or alimentary or esophag* or oesophag* or stomach or intestin* or bowel* or colon* or colorectal or rectal or rectum or sigmoid or duoden* or ileum or ileal or jejun* or anal or anus) adj3 (polyp* or mass* or lesion* or tumor* or tumour* or carcin* or adeno* or neoplas* or cancer* or malignan* or sarcoma* or lymphoma* or leiomyosarcoma*)).ab,hw,ti.) AND ((endoscop* or enteroscop* or gastroscop* or colonoscop* or duodenoscop* or rectoscop* or sigmoidoscop* or ileocolonoscop* or chromoendoscop* or esophagogastroduodenoscop* or esophagoscop* or oesophagogastroduodenoscop* or proctoscop* or ERCP or anoscop* or endomicroscop* or oesophagoscop* or gastroduodenoscop* or sigmoidoscop* or diagnos* or patholog*).ab,hw,ti.) AND (“artificial intelligence” or “machine learning” or “machine intelligen*” or computer-aided or “computational intelligen*” or “deep learning” or “deep unified network*” or “data mining” or datamining or “supervised learning” or “semi-supervised learning” or “unsupervised learning” or “automated pattern recognition” or “Bayesian learning” or “computer heuristics” or “hidden Markov model*” or “k-nearest neighbor*” or “kernel method*” or “learning algorithm*” or “natural language processing” or “support vector” or “vector machine” or Gaussian or Bootstrap or “regression tree*” or “linear discriminant analysis” or “naive Bayes” or “learning vector” or “random forest*” or “chi-square automatic interaction detection” or “iterative dichotom*” or fuzzy or “neural network*” or perceptron* or (computer adj1 heuristic*).ab,hw,ti.)

Embase (1974+)

(digestive system cancer/ or exp esophagus cancer/ or exp intestine cancer/ or exp stomach cancer/ or digestive system tumor/ or exp esophagus tumor/ or exp gastrointestinal tumor/ or exp intestine tumor/ or exp stomach tumor/ or ((digestive or gastr* or GI or alimentary or esophag* or oesophag* or stomach or intestin* or bowel* or colon* or colorectal or rectal or rectum or sigmoid or duoden* or ileum or ileal or jejun* or anal or anus) adj3 (polyp* or mass* or lesion* or tumor* or tumour* or carcin* or adeno* or neoplas* or cancer* or malignan* or sarcoma* or lymphoma* or leiomyosarcoma*)).ab,kw,ti.)

AND (digestive tract endoscopy/ or exp chromoendoscopy/ or exp endoscopic retrograde cholangiopancreatography/ or exp esophagogastroduodenoscopy/ or exp esophagoscopy/ or exp gastrointestinal endoscopy/ or digestive endoscope/ or exp anoscope/ or exp balloon enteroscope/ or exp capsule endoscopy/ or exp colonoscope/ or exp digestive endomicroscope/ or exp duodenoscope/ or exp esophagoscope/ or exp gastroduodenoscope/ or exp gastroscop* or exp proctoscope/ or exp sigmoidoscope/ or

(endoscop* or enteroscop* or gastroscop* or colonoscop* or duodenoscop* or rectoscop* or sigmoidoscop* or ileocolonoscop* or chromoendoscop* or esophagogastroduodenoscop* or esophagoscop* or oesophagogastroduodenoscop* or proctoscop* or ERCP or anoscop* or endomicroscop* or oesophagoscop* or gastroduodenoscop* or sigmoidoscop* or diagnos* or patholog*).ab,kw,ti.) AND (exp artificial intelligence/ or exp machine learning/ or (“artificial intelligence” or “machine learning” or “machine intelligen*” or computer-aided or “computational intelligen*” or “deep learning” or “deep unified network*” or “data mining” or datamining or “supervised learning” or “semi-supervised learning” or “unsupervised learning” or “automated pattern recognition” or “Bayesian learning” or “computer heuristics” or “hidden Markov model*” or “k-nearest neighbor*” or “kernel method*” or “learning algorithm*” or “natural language processing” or “support vector” or “vector machine” or Gaussian or Bootstrap or “regression tree*” or “linear discriminant analysis” or “naive Bayes” or “learning vector” or “random forest*” or “chi-square automatic interaction detection” or “iterative dichotom*” or fuzzy or “neural network*” or perceptron* or (computer adj1 heuristic*).ab,kw,ti.) NOT (exp animal/ not exp human/, exp child/ not exp adult/, “case report”.kw,pt,ti.) Limit to English

Ovid MEDLINE(R) 1946 to Present and Epub Ahead of Print, In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) Daily

(exp Gastrointestinal Neoplasms/ or ((digestive or gastr* or GI or alimentary or esophag* or oesophag* or stomach

or intestin* or bowel* or colon* or colorectal or rectal or rectum or sigmoid or duoden* or ileum or ileal or jejun* or anal or anus) adj3 (polyp* or mass* or lesion* or tumor* or tumour* or carcin* or adeno* or neoplas* or cancer* or malignan* or sarcoma* or lymphoma* or leiomyosarcoma*).ab,kf,ti.) AND (exp Endoscopy, Digestive System/ or exp Endoscopes, Gastrointestinal/ or (endoscop* or enteroscop* or gastroscop* or colonoscop* or duodenoscop* or rectoscop* or sigmoidoscop* or ileocolonoscop* or chromoendoscop* or esophagogastroduodenoscop* or esophagoscop* or oesophagogastroduodenoscop* or proctoscop* or ERCP or anoscop* or endomicroscop* or oesophagoscop* or gastroduodenoscop* or sigmoidoscop* or diagnos* or patholog*).ab,kf,ti.) AND (exp Artificial Intelligence/ or (“artificial intelligence” or “machine learning” or “machine intelligen*” or computer-aided or “computational intelligen*” or “deep learning” or “deep unified network*” or “data mining” or datamining or “supervised learning” or “semi-supervised learning” or “unsupervised learning” or “automated pattern recognition” or “Bayesian learning” or “computer heuristics” or “hidden Markov model*” or “k-nearest neighbor*” or “kernel method*” or “learning algorithm*” or “natural language processing” or “support vector” or “vector machine” or Gaussian or Bootstrap or “regression tree*” or “linear discriminant analysis” or “naive Bayes” or “learning vector” or “random forest*” or “chi-square automatic interaction detection” or “iterative dichotom*” or fuzzy or “neural network*” or perceptron* or (computer adj1 heuristic*).ab,kf,ti.) NOT (exp Animals/ not Humans/, exp CHILD/ not exp ADULT/, “case report”.kf,pt,ti.) Limit to English

Scopus

(TITLE-ABS-KEY ((digestive OR gastr* OR gi OR alimentary OR esophag* OR oesophag* OR stomach OR intestin* OR bowel* OR colon* OR colorectal OR rectal OR rectum OR sigmoid OR duoden* OR ileum OR ileal OR jejun* OR anal OR anus) W/3 (polyp* OR mass* OR lesion* OR tumor* OR tumour* OR carcin* OR adeno* OR neoplas* OR cancer* OR malignan* OR sarcoma* OR lymphoma* OR leiomyosarcoma*))) AND (TITLE-ABS-KEY (endoscop* OR enteroscop* OR gastroscop* OR colonoscop* OR duodenoscop* OR rectoscop* OR sigmoidoscop* OR ileocolonoscop* OR chromoendoscop* OR esophagogastroduodenoscop* OR esophagoscop* OR oesophagogastroduodenoscop* OR proctoscop* OR ercp OR anoscop* OR endomicroscop* OR oesophagoscop* OR gastroduodenoscop* OR sigmoidoscop* OR diagnos* OR

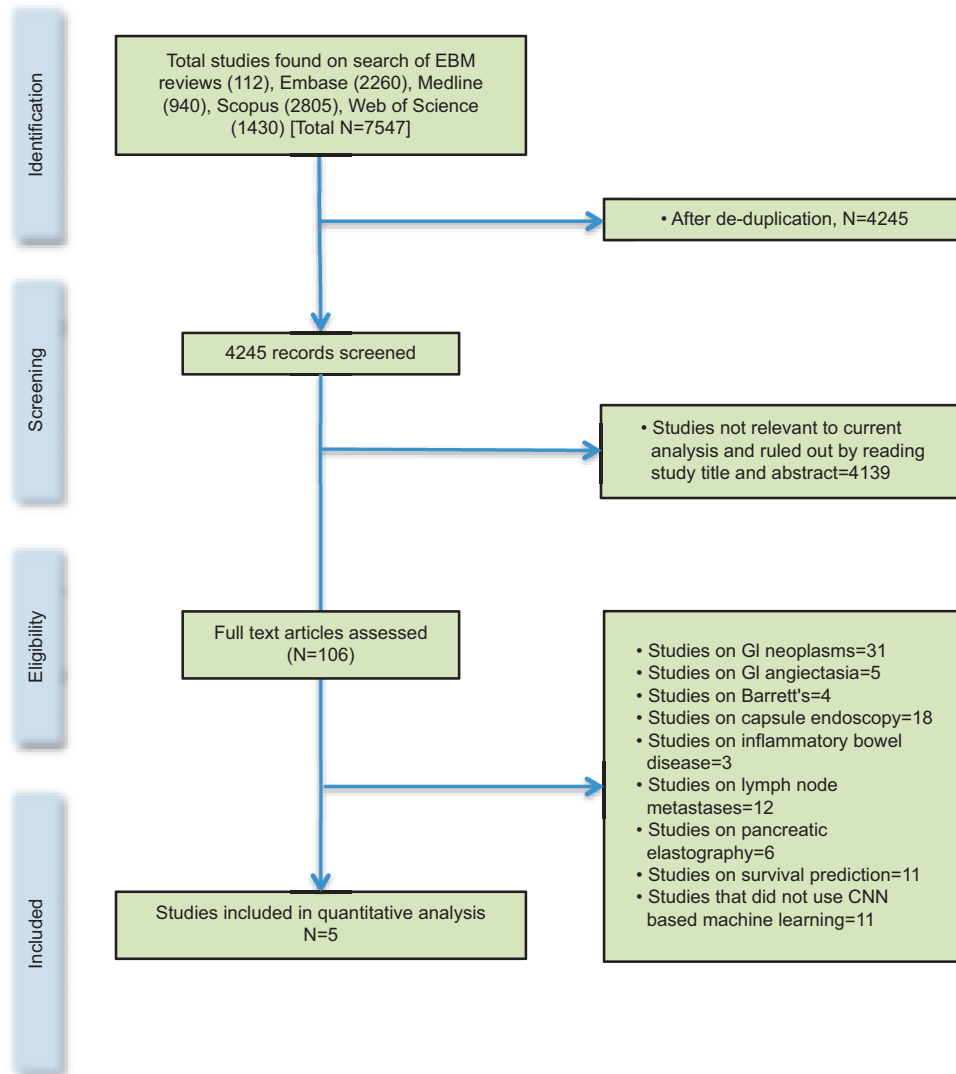
patholog*)) AND (TITLE-ABS-KEY (“artificial intelligence” or “machine learning” OR “machine intelligen*” OR computer-aided OR “computational intelligen*” OR “deep learning” OR “deep unified network*” OR “data mining” OR datamining OR “supervised learning” OR “semi-supervised learning” OR “unsupervised learning” OR “automated pattern recognition” OR “Bayesian learning” OR “computer heuristics” OR “hidden Markov model*” OR “k-nearest neighbor*” OR “kernel method*” OR “learning algorithm*” OR “natural language processing” OR “support vector” OR “vector machine” OR gaussian OR bootstrap OR “regression tree*” OR “linear discriminant analysis” OR “naive Bayes” OR “learning vector” OR “random forest*” OR “chi-square automatic interaction detection” OR “iterative dichotom*” OR fuzzy OR “neural network*” OR perceptron* OR (computer AND W/1 AND heuristic*)) AND (LIMIT-TO (LANGUAGE , “English”))

Web of Science

TS=((digestive or gastr* or GI or alimentary or esophag* or oesophag* or stomach or intestin* or bowel* or colon* or colorectal or rectal or rectum or sigmoid or duoden* or ileum or ileal or jejun* or anal or anus) NEAR/3 (polyp* or mass* or lesion* or tumor* or tumour* or carcin* or adeno* or neoplas* or cancer* or malignan* or sarcoma* or lymphoma* or leiomyosarcoma*)) AND TS=(“artificial intelligence” or “machine learning” or “machine intelligen*” or computer-aided or “computational intelligen*” or “deep learning” or “deep unified network*” or “data mining” or datamining or “supervised learning” or “semi-supervised learning” or “unsupervised learning” or “automated pattern recognition” or “Bayesian learning” or “computer heuristics” or “hidden Markov model*” or “k-nearest neighbor*” or “kernel method*” or “learning algorithm*” or “natural language processing” or “support vector” or “vector machine” or Gaussian or Bootstrap or “regression tree*” or “linear discriminant analysis” or “naive Bayes” or “learning vector” or “random forest*” or “chi-square automatic interaction detection” or “iterative dichotom*” or fuzzy or “neural network*” or perceptron* or (computer NEAR/1 heuristic*)) Limit to English

Appendix 2 MOOSE checklist

Item No	Recommendation	Reported on Page No
Reporting of background should include		
1	Problem definition	5
2	Hypothesis statement	-
3	Description of study outcome(s)	5
4	Type of exposure or intervention used	5
5	Type of study designs used	5
6	Study population	5
Reporting of search strategy should include		
7	Qualifications of searchers (e.g., librarians and investigators)	7, Appendix 1
8	Search strategy, including time period included in the synthesis and key words	7, Appendix 1
9	Effort to include all available studies, including contact with authors	7
10	Databases and registries searched	7, Appendix 1
11	Search software used, name and version, including special features used (e.g., explosion)	Appendix 1
12	Use of hand searching (e.g., reference lists of obtained articles)	-na-
13	List of citations located and those excluded, including justification	Appendix 1
14	Method of addressing articles published in languages other than English	-na-
15	Method of handling abstracts and unpublished studies	7
16	Description of any contact with authors	7
Reporting of methods should include		
17	Description of relevance or appropriateness of studies assembled for assessing the hypothesis to be tested	7-8
18	Rationale for the selection and coding of data (e.g., sound clinical principles or convenience)	7-8
19	Documentation of how data were classified and coded (e.g., multiple raters, blinding and inter-rater reliability)	-na-
20	Assessment of confounding (e.g., comparability of cases and controls in studies where appropriate)	-na-
21	Assessment of study quality, including blinding of quality assessors, stratification or regression on possible predictors of study results	7-8
22	Assessment of heterogeneity	8
23	Description of statistical methods (e.g., complete description of fixed or random effects models, justification of whether the chosen models account for predictors of study results, dose-response models, or cumulative meta-analysis) in sufficient detail to be replicated	8
24	Provision of appropriate tables and graphics	Table 1 supplementary materials
Reporting of results should include		
25	Graphic summarizing individual study estimates and overall estimate	Supplementary materials
26	Table giving descriptive information for each study included	Table 1
27	Results of sensitivity testing (e.g., subgroup analysis)	11
28	Indication of statistical uncertainty of findings	12
Item No	Recommendation	Reported on Page No
Reporting of discussion should include		
29	Quantitative assessment of bias (e.g., publication bias)	12
30	Justification for exclusion (e.g., exclusion of non-English language citations)	-na-
31	Assessment of quality of included studies	-na-
Reporting of conclusions should include		
32	Consideration of alternative explanations for observed results	13-15
33	Generalization of the conclusions (i.e., appropriate for the data presented and within the domain of the literature review)	13-15
34	Guidelines for future research	15



Supplementary Figure 1 Literature search flow chart

EBM, evidence-based medicine; GI, gastrointestinal; *H. pylori*, *Helicobacter pylori*; CNN, convolutional neural network