NIHR | National Institute for Health and Care Research

PROSPERO
International prospective register of systematic reviews

Artificial Intelligence of predictive models for recurrence thyroid cancer following surgery: A systematic review and Meta- Analysis

Andrea Paola Solis Pazmino, Lenara Golbert, Vanessa Mattevi, Erika Laurini de Souza Meyer, Debora Nienow, Freddy Molina, Kamille Guidolin, Laura Berton Eidt, Oscar Ponce, Luis Figueroa

Citation

Andrea Paola Solis Pazmino, Lenara Golbert, Vanessa Mattevi, Erika Laurini de Souza Meyer, Debora Nienow, Freddy Molina, Kamille Guidolin, Laura Berton Eidt, Oscar Ponce, Luis Figueroa. Artificial Intelligence of predictive models for recurrence thyroid cancer following surgery: A systematic review and Meta- Analysis. PROSPERO 2025 CRD420251004644. Available from https://www.crd.york.ac.uk/PROSPERO/view/CRD420251004644.

REVIEW TITLE AND BASIC DETAILS

Review title

Artificial Intelligence of predictive models for recurrence thyroid cancer following surgery: A systematic review and Meta- Analysis

Condition or domain being studied

Machine learning; Artificial intelligence; Thyroid Cancer; Recurrent Disease

This study addresses the global increase in the incidence of thyroid cancer, with a particular focus on papillary thyroid carcinoma (PTC), which now accounts for 90% of differentiated thyroid malignancies. In the United States, the incidence of PTC has tripled over the past three decades, and in Brazil, it increased from 1.1% in 2000 to 11.0% in 2014. Although most cases are low risk, the increasing number of survivors requires ongoing monitoring. Predictive artificial intelligence (AI) models have been developed to predict cancer recurrence and improve clinical management.

Rationale for the review

The application of AI in the medical field has grown in recent years, included PTC. Predictive models for recurrence in papillary thyroid carcinoma (PTC) are designed to identify patients at higher risk of recurrence, thereby enabling more personalized follow-up and treatment plans. These models integrate a variety of clinical, pathological, and molecular factors. Clinical factors

include patient age at diagnosis, sex, family history of thyroid cancer, and prior radiation exposure. Pathological factors consider tumor size, extrathyroidal extension, multifocality, lymph node involvement, histological subtype, and surgical margin status.

Although AI models have demonstrated high accuracy in internal validation, they lack external validation to improve their generalizability. This study aims to externally validate existing predictive models using data from DTC patients in Latin America, including diverse geographic and ethnic/racial populations, molecular markers, and follow-up variables. This approach aims to address current limitations and improve the generalizability and accuracy of models in predicting outcomes for a broader range of patients.

Review objectives
What are the best artificial intelligence models in Differentiated Thyroid Cancer Recurrence?

Keywords
Artificial intelligence; Thyroid cancer; Recurrence; Differentiated thyroid cancer; Machine learning

Country
Brazil

ELIGIBILITY CRITERIA

Population

Included
Adults with a confirmed diagnosis of differentiated thyroid cancer

Excluded
I. Non differentiated thyroid cancer or medullary thyroid cancer

II. Conference/meeting abstracts

III. Non-English language publications

Intervention(s) or exposure(s)

Included
Artificial intelligence; Thyroid; Machine learning

Artificial Intelligence of predictive models for recurrence thyroid cancer

(machine learning, linear, ridge, svm, support vector machine, knn, K-Nearest Neighbors, lasso, decision tree regressor, random forest regressor, xgboost regressor, deep machine learning, neural network, Transfer Learning OR Computational Intelligence OR Machine Intelligence OR SVM OR Gradient Boosting Machine OR GBM OR Adaboost OR Radiomics OR radiomic OR radiogenomic OR radiomics-based)

Comparator(s) or control(s)
This review does not have any comparators

Study design
Only nonrandomized study types will be included.

Included
longitudinal studies (either observational or experimental), cohort studies, clinical trials, case control studies, case series.

Excluded
systematic reviews or meta- analysis

Context
Although AI models have demonstrated high accuracy in internal validation, they lack external validation to improve their generalizability. This study aims to externally validate existing predictive models using data from DTC patients in Latin America, including diverse geographic and ethnic/racial populations, molecular markers, and follow-up variables. This approach aims to address current limitations and improve the generalizability and accuracy of models in predicting outcomes for a broader range of patients.

TIMELINE OF THE REVIEW

Date of first submission to PROSPERO
04 March 2025

Review timeline
Start date: 4 March 2025. End date: 4 March 2026.

Date of registration in PROSPERO
15 April 2025

AVAILABILITY OF FULL PROTOCOL

Availability of full protocol
A full protocol has been written and uploaded to PROSPERO. The protocol will be made available after the review is completed.

SEARCHING AND SCREENING

Search for unpublished studies
Only published studies will be sought.

Main bibliographic databases that will be searched
The main databases to be searched are *CENTRAL - Cochrane Central Register of Controlled Trials*, *CLIB - The Cochrane Library*, *Embase - Embase via Ovid*, *Embase.com*, *LILACS - Latin American and Caribbean Health Sciences Literature*, *MEDLINE*, *PubMed*, *SCI - Science Citation Index*, *SSCI - Social Science Citation Index* and *Scopus*.

Search language restrictions
The review will only include studies published in English.

Search date restrictions
There are no search date restrictions.

Other methods of identifying studies

Other studies will be identified by: *contacting authors or experts*, *searching conference proceedings* and *searching trial or study registers*.

Link to search strategy
A full search strategy is available in the full protocol as described in the *Availability of full protocol* section

Selection process
Studies will be screened independently by at least two people (or person/machine combination) with a process to resolve differences.

Other relevant information about searching and screening
A comprehensive search of several databases from each database's inception to January 17th, 2025, any language, was conducted. The databases included Ovid MEDLINE(R) and Epub Ahead of Print, In-Process & Other Non-Indexed Citations, and Daily, Ovid EMBASE, Ovid Cochrane Central Register of Controlled Trials, Ovid Cochrane Database of Systematic Reviews, Scopus, and Web of Science. The search strategy was designed and conducted by an experienced librarian with input from the study's principal investigator. Controlled vocabulary supplemented with keywords was used to search for the use of artificial Intelligence for prediction of differentiated thyroid cancer recurrence in adults. The actual strategy listing all search terms used and how they are combined is available in the appendix.

DATA COLLECTION PROCESS

Data extraction from published articles and reports
Data will be extracted independently by at least two people (or person/machine combination) with a process to resolve differences.

Authors will be asked to provide any required data not available in published reports.

Study risk of bias or quality assessment
Risk of bias will be assessed using: *Cochrane RoB-2*

Data will be assessed independently by at least two people (or person/machine combination) with a process to resolve differences.

Additional information will be sought from study investigators if required information is unclear or unavailable in the study publications/reports.

Reporting bias assessment
Study quality will be assessed by two independent reviewers based on CLARITY tool for cohort-studies and case-control studies. The domains for cohort studies are: 1) selection of the cohorts that are exposed and non-exposed take in from the same population; 2) assessment of exposure; 3) outcome of interest was not present at start of study; 4) Match cohorts for all variables related to the outcome of interest or statistical adjustment for these prognostic variables; 5) evaluation of the presence or absence of prognostic factors; 6) assessment of the outcome; 7) adequate follow up of cohorts; and 8) similarity of co-interventions between groups. The domains for case-control assessment include: 1) evaluation of exposure; 2) Security that the cases had developed the result of interest and the controls had not; 3) proper selection of cases; 4) proper selection of controls; and 5) coincidence of cases and controls according to

important prognostic variables or statistical adjustment for these variables. The risk of bias will for randomized clinical trials be assessed using the revised Cochrane risk of bias tool for randomized clinical trials (RoB 2). This tool takes into consideration five domains: (1) randomization process, (2) intended interventions (effect of assignment to intervention), (3) missing outcome data, (4) measurement of the outcome, and (5) reported result. Each domain risk of bias will be judged as low risk of bias, intermediate risk of bias and high risk of bias, by using a predetermined algorithm.

Overall risk of bias will be calculated for each study. Articles will be deemed to be at low risk of bias when all domains are at low risk of bias. They will be graded as intermediate risk of bias when at least one domain is assessed as intermediate risk of bias and all the other domains are at low risk of bias, and articles will be considered at high risk of bias when at least one domain is at high risk of bias or when more than one domain is judged to have intermediate risk of bias.

Lastly, the quality of evidence will be appraised for each outcome and using the Grading of Recommendations Assessment, Development and Evaluation approach (GRADE). Two reviewers will perform this assessment, and disagreements will be resolved by consensus with a third reviewer.

Certainty assessment

The quality of evidence will be appraised for each outcome and using the Grading of Recommendations Assessment, Development and Evaluation approach (GRADE). Two reviewers will perform this assessment, and disagreements will be resolved by consensus with a third reviewer.

OUTCOMES TO BE ANALYSED

Main outcomes

The outcomes of interest for this study are:

I. Recurrence: it occurs when a new thyroid arises after completely removing the original disease.

II. Persistent: refers to thyroid cancer that remains detectable or recurs after initial treatment

Additional outcomes

I. Disease-specific survival (DSS): percentage of people in a study or treatment group who have not died from a specific disease in a defined period of time (usually from diagnosis or treatment initiation to death)

II. Progression-free survival (PFS): length of time during and after treatment that a patient lives with the disease but it does not get worse

III. Locoregional recurrence (LRR): recurrence of cancer cells at the same site as the original (primary) tumor or the regional lymph nodes after a disease free period

PLANNED DATA SYNTHESIS

Strategy for data synthesis

For dichotomous outcomes, we will calculate relative risks (RR) and their corresponding 95% confidence intervals (CI) using random-effects models fitted with restricted maximum-likelihood estimator (REML). These effect sizes will be calculated for multiple follow-up times when possible. Additionally, if we find time-to-event outcomes (e.g. survival), we will perform random effect analysis using the log form of hazard ratios (HR) and their variances.

For continuous outcomes, we will pool outcomes with identical definitions by using mean differences (MD) and their CI. If outcomes have different constructs, we will calculate standardized mean difference (SMD) and their CI with the Hedges method. All of these analyses will be performed with random-effects models and the REML.

Heterogeneity will be assess visually and by using the I^2 statistic will. Broadly, an I^2 of 25% might reflect low inconsistency and an I^2 of 75% could reflect high inconsistency.

CURRENT REVIEW STAGE

Stage of the review at this submission

Review stage	Started	Completed
Pilot work	✓	
Formal searching/study identification	✓	
Screening search results against inclusion criteria		
Data extraction or receipt of IPD		
Risk of bias/quality assessment		
Data synthesis		

Review status

The review is currently planned or ongoing.

Publication of review results

Results of the review will be published in English.

REVIEW AFFILIATION, FUNDING AND PEER REVIEW

Review team members

Dr Andrea Paola Solis Pazmino (review guarantor and contact) Santa Casa de Misericordia, Porto Alegre. Brazil.

No conflict of interest declared.

Dr Lenara Golbert. Santa Casa de Misericordia, Porto Alegre. Brazil.

No conflict of interest declared.

Dr Vanessa Mattevi. Santa Casa de Misericordia, Porto Alegre. Brazil.

No conflict of interest declared.

Dr Erika Laurini de Souza Meyer. Santa Casa de Misericordia, Porto Alegre. Brazil.

No conflict of interest declared.

Dr Debora Nienow. Santa Casa de Misericordia, Porto Alegre. Brazil.

No conflict of interest declared.

Dr Freddy Molina. Santa Casa de Misericordia, Porto Alegre. Brazil.

No conflict of interest declared.

Dr Kamille Guidolin. Santa Casa de Misericordia, Porto Alegre. Brazil.

No conflict of interest declared.

Dr Laura Berton Eidt. Santa Casa de Misericordia, Porto Alegre. Brazil.

No conflict of interest declared.

Dr Oscar Ponce. Catalina Research Initiative. England.

No conflict of interest declared.

Dr Luis Figueroa. Catalina Research Initiative. Ecuador.

No conflict of interest declared.

Named contact

Dr Andrea Paola Solis Pazmino (andrea.pazmino@ufcspa.edu.br). Santa Casa de Misericordia, Porto Alegre. Brazil.

Review affiliation

Santa Casa de Misericordia, Porto Alegre, Brazil

Funding source

Review has no specific/external funding but is supported by guarantor/review team (non-commercial) institutions.

Peer review

Search results will be uploaded into an online platform called Rayyan. The abstract screening will be performed by 6 reviewers and it consists of reviewing titles and abstracts. The articles included by at least one reviewer will follow to full-text screening. After the first phase, full-text material from included articles will be obtained and screened. At this phase, articles included by the pair of reviewers will be deemed included for this systematic review. Reviewers will work independently and in duplicates in both phases. In case of disagreements at full-text screening, they will be resolved by an external expert doctor. At the beginning of each phase, a pilot with 5 selected articles will be carried out to assess the understanding and clarity of the eligibility criteria.

ADDITIONAL INFORMATION

Review conflict of interest

Declared individual interests are recorded under team member details.. No additional interests are recorded for this review.

Medical Subject Headings

Artificial Intelligence; Brazil; Humans; Incidence; Machine Learning; Neoplasm Recurrence, Local; Neural Networks, Computer; Radiomics; Random Forest; Support Vector Machine; Survivors; Thyroid Cancer, Papillary; Thyroid Neoplasms; United States

SIMILAR REVIEWS

Check for similar records already in PROSPERO

PROSPERO identified a number of existing PROSPERO records that were similar to this one (last check made on 4 March 2025). These are shown below along with the reasons given by that the review team for the reviews being different and/or proceeding.

- Current Status of Artificial Intelligence Technology in Predicting Delayed Cerebral Ischemia after Spontaneous Subarachnoid Hemorrhage: A Systematic Review [published 28 October 2024] [CRD42024600866]. The review was judged **not to be similar**
- Accuracy of Machine Learning Models in Detecting Pancreatic Cancer in Computed Tomography Images: A Systematic Review and Meta-Analysis [published 29 March 2024] [CRD42024525983]. **The authors did not check this review**
- The Diagnostic Accuracy of Artificial Intelligence Models in Detecting Lymph Node Metastases in Lung Cancer Using Endobronchial Ultrasound (EBUS) Images: A Systematic Review and Diagnostic Meta-analysis [published 17 January 2025] [CRD42025635581]. The review was judged **not to be similar**
- Artificial intelligence in the prediction of lymph node metastasis in thyroid cancer patients: a systematic review and meta-analysis [published 12 March 2022] [CRD42022310118]. The review was judged **not to be similar**

PROSPERO version history

- Version 1.0, published 15 Apr 2025

Disclaimer

The content of this record displays the information provided by the review team. PROSPERO does not peer review registration records or endorse their content.

PROSPERO accepts and posts the information provided in good faith; responsibility for record content rests with the review team. The guarantor for this record has affirmed that the information provided is truthful and that they understand that deliberate provision of inaccurate information may be construed as scientific misconduct.

PROSPERO does not accept any liability for the content provided in this record or for its use. Readers use the information provided in this record at their own risk.

Any enquiries about the record should be referred to the named review contact