



PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Application of artificial intelligence models in eosinophilic esophagitis: a systematic review.	Title
ABSTRACT			
Abstract	2	Background: Artificial intelligence (AI) tools are increasingly being integrated into computer-aided diagnosis systems that can be applied to improve the recognition and clinical and molecular characterization of allergic diseases, including eosinophilic esophagitis (EoE). This review aims to systematically evaluate current applications of AI, machine learning (ML), and deep learning (DL) methods in EoE characterization and management. Methods: A search strategy was designed to retrieve all articles via the online database PubMed, Embase, and Web of Science. All studies that met the following criteria were included in the review. The risk of bias and applicability for eligible studies was assessed according to the prediction model study risk of bias assessment tool (PROBAST). Results: 120 articles were found. After removing 68 duplicates, 52 articles were reviewed based on the title and abstract, and 34 articles were excluded. Eleven full texts were assessed for eligibility, met the inclusion criteria, and were analyzed for the systematic review. All articles were published since 2018. AI technologies have been applied to create and validate models to improve the endoscopic and histologic diagnosis the clinical and molecular characterization of EoE heterogeneity. Discussion: AI technologies could promote more accurate evidence-based management of EoE by integrating the results of molecular signature, clinical, histology and endoscopic features. However, the era of AI application in medicine is just beginning; therefore, further studies with model validation in the real-world environment are required.	Abstract
INTRODUCTION			
Rationale	3	AI tools are increasingly being integrated into computer-aided diagnosis systems that can be applied to improve the recognition and clinical and molecular characterization of EoE.	Introduction
Objectives	4	The aim of this review is the systematic evaluation of current applications of AI, ML, and DL methods in EoE characterization and management.	Introduction
METHODS			
Eligibility criteria	5	Inclusion criteria: 1) original research articles (retrospective or prospective studies) published in English in peer-reviewed journals, 2) participants were children and adult patients with a diagnosis of EoE histologically confirmed (≥ 15 eosinophils/HPF) according to guidelines.	Methods
Information sources	6	PubMed, Embase and Web of Science.	Methods
Search strategy	7	An extensive search strategy was designed to retrieve all articles via the online databases PubMed, Embase and Web of Science combining the terms "artificial intelligence" AND "eosinophilic esophagitis", "machine learning" AND "eosinophilic esophagitis" and "deep learning" AND "eosinophilic esophagitis" and following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.	Methods
Selection process	8	The review of literature was performed in May 2023, including all publication years. Search results were compiled using the software Refworks®. Two independent researchers screened retrieved articles. The same investigators independently assessed full texts of records deemed eligible for inclusion. Any discrepancies were resolved by discussion and consensus. Authors of publications reporting unclear data were contacted by email for clarification.	Methods
Data collection process	9	Two independent reviewers extracted data from each eligible study using a standardized data extraction sheet and then proceeded to cross-check the results. We extracted the following information: first author name; date of publication; AI methodology; study outcome; accuracy, sensitivity, and specificity of the AI models.	Methods
Data items	10a	We extracted the following information: first author name; date of publication; AI methodology; study outcome; accuracy, sensitivity, and specificity of the AI models	Methods
	10b	Not applicable.	
Study risk of bias assessment	11	The risk of bias and applicability for eligible studies was assessed according to the Prediction model study Risk Of Bias Assessment Tool (PROBAST). The risk of bias and applicability are classified as low, unclear, or high. The evaluation tool contains 20 signaling questions from four domains: participants, predictors, outcomes, and analyses.	Methods



PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
Effect measures	12	Not applicable.	
Synthesis methods	13a	Not applicable.	
	13b	Not applicable.	
	13c	Not applicable.	
	13d	Not applicable.	
	13e	Not applicable.	
	13f	Not applicable.	
Reporting bias assessment	14	Not applicable.	
Certainty assessment	153	Not applicable.	
RESULTS			
Study selection	16a	We found 120 articles. After removing 68 duplicates, 52 articles were reviewed based on the title and abstract, and 34 articles were excluded. Eleven full texts were assessed for eligibility, met the inclusion criteria, and were analyzed for the systematic review. All articles were published since 2018.	Results
	16b	Six congress abstracts were excluded because of data were too limited to be assessed for bias risk	Results
Study characteristics	17	<p>Okimoto et al., developed a CNN system algorithm to analyze multiple esophageal endoscopic images demonstrating high sensitivity (90.8%), specificity (96.6%), and accuracy (94.7%) for detection of EoE.</p> <p>Guimarães et al., established and trained a CNN-based approach to distinguish the endoscopic appearance of EoE from normal findings and candida esophagitis. The CNN algorithm showed a global accuracy of 91.5%, sensitivity of 87.1%, specificity of 93.6%, and AUC 0.966, that were higher than the endoscopists.</p> <p>Römmele et al., developed an AI algorithm for detecting and quantifying the endoscopic features of EoE in white light images, integrated by the EREFS. The overall sensitivity, specificity, and accuracy of the algorithm were 0.96, 0.94, 0.95 respectively, while the AUC was 0.992.</p> <p>Adorno et al., created and validated an automated eosinophil detection model. The authors realized a CNN model that predicted the location of eosinophils on histological images with an overall accuracy of 99.0%, sensitivity of 100%, and specificity of 98.2% and linked biopsy features with treatments and clinical phenotypes.</p> <p>Czyzewski et al. developed a platform based on a DCNN that analyzed esophageal biopsies with an accuracy of 85%, sensitivity of 82.5%, and specificity of 87%.</p> <p>Daniel et al. recently developed a ML pipeline to identify and quantify esophageal eosinophils at the whole slide image level, detecting intact and not-intact eosinophils with a global accuracy of 94.75%, sensitivity of 94.13%, and specificity of 95.25%.</p> <p>Larey et al. developed a platform, using ML, that provided a complete quantification of the eosinophils and basal cells fraction over the entire slide, quantifying the peak count, basal cell fraction, the percent of HPFs that have more than 15 eosinophils and the percent of HPFs that have more than 25% basal cells within them. This algorithm predicted the histological severity better than the gold-standard method.</p> <p>Archila et al. developed an AI-based digital pathology model for the evaluation of histologic features in the spectrum of EoE. The model showed an excellent performance for the recognition of various EoE histologic features, including the lamina propria remodeling, representing an accurate and reproducible method for semi-automated quantitative analysis to be used in the evaluation of esophageal biopsies.</p> <p>Sallis et al. developed an AI based automated algorithm to generate the diagnostic probability score for EoE [p(EoE)], based on esophageal mRNA transcripts from patients with EoE, gastroesophageal reflux disease (GERD), and healthy controls. The p(EoE) score ≥ 25 detected active EoE with</p>	Results



PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
		high accuracy (sensitivity 91%, specificity 93%, and AUC 0.98) and improved diagnosis of doubtful cases by 85%, distinguishing EoE from GERD. Sallis et al. authors analyzed the esophageal transcript of EoE patients presenting with (EoE + FI) and without FI using ML techniques. The algorithm identified EoE+FI patients with a sensitivity of 93% and a specificity of 100%. Shoda et al., identified relationships between the endoscopic, histologic, and molecular (EoE diagnostic panel) features and determined EoE endotypes, using a ML approach. Notably, the authors found 3 distinct EoE endotypes (EoEe).	
Risk of bias in studies	18	Using PROBAST, 18% and 64% of the studies were classified as having a low risk of bias and applicability. In the domains of participants, predictors, and outcomes, most studies were classified as low risk. However, in the domain of analyses, most studies were classified as unclear risk. The results were analyzed and classified according to the field of AI technologies application, including the diagnosis (endoscopically, histologically, molecularly) and the assessment of EoE heterogeneity.	Results
Results of individual studies	19	Not applicable.	
Results of syntheses	20a	Not applicable.	
	20b	Not applicable.	
	20c	Not applicable.	
	20d	Not applicable.	
Reporting biases	21	Not applicable.	
Certainty of evidence	22	Not applicable.	
DISCUSSION			
Discussion	23a	The use of AI could promote more accurate evidence-based management of EoE by integrating the results of molecular signature, clinical, histology and endoscopic features.	Conclusion
	23b	The included studies were highly heterogeneous in nature with a variety of clinical themes, and evaluation methods. Therefore, we could not extract any universal conclusion from the systematic review. Although we conducted a broad search strategy encompassing three large databases, relevant publications might have been missing, such as conference or congress abstracts, and preprint articles published in other online databases.	Discussion
	23c	Moreover, biases are present. Most studies showed a low risk of bias in the domains of participants, predictors, and outcomes. However, in the domain of analyses, most studies had an unclear risk, owing to the appropriate handling of missing data, the appropriate evaluation of model performance, as well as accounting for model overfitting, underfitting, and optimism. In particular, the lack of external validation was the main limitation of several studies.	Discussion
	23d	The era of AI application in medicine is just beginning. This field is certainly high on promise, but relatively low on data and proof. AI models in medicine, including EoE, requires rigorous studies and extensive validation in a real-world environment	Discussion
OTHER INFORMATION			
Registration and protocol	24a	The systematic review is under review to be published in PROSPERO (protocol number will be provided when available)	Methods
	24b	PROSPERO CRD42023451048	Methods
	24c	Not applicable.	
Support	25	No financial support.	
Competing interests	26	No competing interests	



PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
Availability of data, code and other materials	27	Not applicable.	