Applications of Artificial Intelligence to Eosinophilic Esophagitis

Eric Reuben Smith 1,* and Jay Shah 2

1 Internal Medicine Program, Baylor Scott and White-Round Rock, Round Rock, TX 78665, USA
2 Department of Pediatrics, Division of Gastroenterology, Hepatology and Nutrition, University of Texas Health Science Center at San Antonio, San Antonio, TX 78229, USA; shahj3@uthscsa.edu
* Correspondence: smithe4@uthscsa.edu

Abstract: Eosinophilic Esophagitis (EoE) is a chronic immune-related inflammation, and challenges to its diagnosis and treatment evaluation persist. This literature review evaluates all AI applications to EOE, including 15 studies using AI algorithms for counting eosinophils in biopsies, as well as newer diagnostics using mRNA transcripts in biopsies, endoscopic photos, blood and urine biomarkers, and an improved scoring system for disease classification. We also discuss the clinical impact of these models, challenges faced in applying AI to EoE, and future applications. In conclusion, AI has the potential to improve diagnostics and clinical evaluation in EoE, improving patient outcomes.

Keywords: eosinophilic esophagitis; artificial intelligence; machine learning

1. Introduction

Eosinophilic esophagitis (EoE) is a chronic Th2 antigen mediated inflammation defined by increased eosinophils in the esophageal mucosa [1]. EoE is associated with dysphagia, GERD-like symptoms, emesis, abdominal pain, and food impaction. Although EoE can present with esophageal trachealization, longitudinal furrows, and whitish exudates on endoscopy, diagnosis is confirmed by esophageal biopsy showing more than 15 eosinophils per high-power field after ruling out other causes of esophageal eosinophilia. Pharmacologic treatment involves swallowed steroids, proton-pump inhibitors, and immunosuppressants in severe cases. Food elimination diets can also help prevent chronic esophageal remodeling and strictures, which can be managed by balloon dilation. There is evidence suggesting that EoE is increasing in incidence and prevalence in both pediatric and adult populations, and prevalence is estimated between 25–50 per 100,000 worldwide [2,3]. EoE has been recognized as a leading cause of food impaction and dysphagia in adults, and is a considerable cause of failure to thrive in children [4].

Many diagnostic and therapeutic challenges exist in the management of EoE. Although progress has been made since EoE was first described in 1993, EoE is a relatively new diagnosis and there are questions about proper diagnostic criteria and therapeutic evaluation [5]. While PPIs are now recognized as a first line treatment for EoE, controversy remains over the best initial food elimination diet, optimal time to wait for treatment evaluation, and definition of treatment response. As a lesser-known disease that presents similarly to more common diseases, diagnosis is delayed on average by 2 years in children and 6 years in adults [6]. Diagnosis can be challenging as symptoms of EoE can easily be confused with GERD. EoE symptoms are not uniform across age ranges, further complicating diagnosis (dysphagia common in adults, vomiting and abdominal pain in children). That said, late diagnosis can increase risk of developing strictures, and stricture rates increase by 9% per year of undiagnosed EoE [7]. Diagnostic biopsies are both invasive and costly, including anesthesia and procedural risk, as well as time away from work and school. Histologic analysis can be time consuming and there are no standardized cut offs for response to therapy in EoE. Treatment outcomes can be difficult to predict as there are multiple EoE phenotypes with seemingly unique pathophysiology, including
involvement of aerosol-antigen sensitization in some cases [8]. Treatment can take years of follow up and multiple endoscopic biopsies, which is concerning as exposure to anesthesia in children has been linked to impaired brain development [9]. To date, no endoscopic scoring system or biomarker has been found to safely replace esophageal biopsy. Many speculate that the future of EoE management will have to focus on prevention, precision, and personalization [10].

Artificial intelligence (AI) has been applied to many other diseases in gastroenterology and is well suited to revolutionize the management of EoE [11]. AI is a branch of computer science which tries to imitate the way humans learn and solve problems using large datasets. Supervised machine learning is an approach within AI where an algorithm is fed thousands of training examples, which includes any number of independent variables tied to a dependent variable [12]. Common algorithms used in supervised machine learning include neural networks, random forest, and supervised vector machines, and each can predict in terms of probabilities (percent chance of disease state) or numerical value (number of weeks to wait before adequate treatment response). After the algorithm is trained on a sample dataset, it can be further refined using prospective feedback such as correct diagnosis and treatment outcome. Unsupervised machine learning is another branch of AI which finds patterns in datasets without labeled outcomes, categorizing data into groups. For example, with multifactorial disease, unsupervised machine learning can categorize phenotype or endotype of each EoE patient. AI algorithms provide a promising future in EoE management by improving preventative care, innovating new non-invasive diagnostics, and providing personalized treatment plans.

2. Materials and Methods

We searched PubMed, MedLine, EMbase, and Google scholar databases up to 16 April 2022 using the search string (“eosinophilic esophagitis” OR “eosinophilic oesophagitis”) AND (“artificial intelligence” OR “deep learning” OR “machine learning”). Our inclusion criteria were defined as original studies that documented AI applications to eosinophilic esophagitis and did not include literature reviews or opinion articles. Results were limited to publications in English. We also conducted a separate query for AI applications to similar allergic gastrointestinal diseases such as celiac and IBD, in order to find correlates that could inform future use of AI in EoE.

3. Results

Our search string yielded 851 results, and all were reviewed by a single researcher (see Figure 1). All studies included are cohort studies. Only 15 publications fit the inclusion criteria of AI applications to EoE, 10 full manuscripts, one poster presentation, and four abstracts. We emailed authors of manuscripts we could not gain access to, as well as abstracts to obtain full manuscripts and poster presentations. Of the included publications, AI applications included: analyzing biopsy histology (6), diagnostics using biopsy mRNA transcripts (3), diagnosing EOE via endoscopic photos (5), and automating measurement of layers of esophageal tissue thickened in EoE (1). All studies except two abstracts met criteria for good study design using the QUADAS quality assessment tool. See Table 1 for full description of included studies.
Table 1. Description of included studies.

<table>
<thead>
<tr>
<th>Study</th>
<th>Type</th>
<th>Outcome</th>
<th>AI Model</th>
<th>Data and Sample Size</th>
<th>Study Results/Validation Cohort</th>
<th>QUADAS Quality Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Catalano et al., 2020</td>
<td>Poster</td>
<td>Eosinophil quantification (PEC and average eosinophil count), diagnosis, prediction of treatment response</td>
<td>U-net segmentation convolutional neural network (CNN)</td>
<td>91 biopsies (36 from patients with EoE)</td>
<td>Diagnostic accuracy of 96% with an average error of + 0.16 eosinophils per HPF, standard deviation of 1.20. They also found high average eosinophil counts were associated with response to four to six food elimination diets and higher numbers of eosinophils in the mid vs. distal esophagus.</td>
<td>7</td>
</tr>
<tr>
<td>Adorno III et al., 2021</td>
<td>Manuscript</td>
<td>Eosinophil quantification (PEC, average eosinophil count, percent patches with 0, &gt;5, 10, 15 eosinophils, average eosinophil size), diagnosis, prediction of treatment response</td>
<td>Compared 11 CNNs: U-Net vs. Res. U-Net vs. R2U-Net vs. Attn. U-Net with various test set dice coefficients</td>
<td>Biopsies from 101 patients (44 with EoE)</td>
<td>Diagnostic accuracy of 99.0%, 100% sensitivity, and 98% specificity. Higher maximum eosinophils in 4/6 FED responders over PPI and steroid responders.</td>
<td>9</td>
</tr>
<tr>
<td>Javaid et al., 2021</td>
<td>Manuscript</td>
<td>Eosinophil quantification (PEC), diagnosis, prediction of treatment response</td>
<td>U-net and VGG16 CNNs</td>
<td>Biopsies from 77 patients (36 with EoE)</td>
<td>Diagnostic accuracy of 99.9%, eosinophil quantification SD of −0.3 Eos/HPF; 4/6 FED responders had higher PEC.</td>
<td>10</td>
</tr>
<tr>
<td>Czyzewski et al., 2021</td>
<td>Manuscript</td>
<td>Diagnosis from biopsy images using deep CNN, determine which image size is ideal</td>
<td>ResNet50 (Deep CNN)</td>
<td>420 biopsy images (210 with EoE)</td>
<td>Diagnostic accuracy of 85%, 448 × 448 pixels downscaled to 224 × 224 performed better than 224 × 224 pixel images, suggesting that global features contribute to model.</td>
<td>10</td>
</tr>
<tr>
<td>Shi et al., 2022</td>
<td>Manuscript</td>
<td>Diagnosis from biopsy through data augmentation using small dataset</td>
<td>ResNet50 and Bit-M CNNs</td>
<td>202 biopsy images from 15 EoE patients compared to 404 normal biopsies</td>
<td>Diagnostic accuracy of 62%, ResNet50 outperformed Bit-M, limited data successfully augmented by random flipping, increasing contrast, and weight to training loss function.</td>
<td>8</td>
</tr>
<tr>
<td>Daniel et al., 2021</td>
<td>Manuscript</td>
<td>Eosinophil quantification (PEC), diagnosis, intact vs. not intact eosinophils</td>
<td>U-net</td>
<td>Biopsies from 23 patients with EoE</td>
<td>Diagnostic accuracy of 95%, distinguishes intact vs. not intact eosinophils with 98.8% accuracy.</td>
<td>10</td>
</tr>
<tr>
<td>Sallis et al., 2018 #1</td>
<td>Manuscript</td>
<td>Diagnosis of EoE vs. GERD vs. controls using RNA transcripts</td>
<td>Random Forest</td>
<td>Biopsies from 113 patients, (38 with EoE)</td>
<td>Diagnostic accuracy of 85% in patients with equivocal histology, created p(EoE) score that predicts diagnosis with AUC 0.985.</td>
<td>10</td>
</tr>
<tr>
<td>Sallis et al., 2018 #2</td>
<td>Manuscript</td>
<td>Identify patients with history of food impaction using RNA transcripts</td>
<td>Random Forest</td>
<td>Biopsies from 215 EoE patients, (26 with food impaction)</td>
<td>Predicts food impaction with 93% sensitivity and 100% specificity</td>
<td>10</td>
</tr>
<tr>
<td>Lin et al., 2018</td>
<td>Abstract</td>
<td>Diagnosis using RNA transcripts and buccal biopsies</td>
<td>Not clear</td>
<td>Not clear</td>
<td>“EoE status can be predicted using buccal epithelial tissue biopsies . . . [enabling] more accurate diagnosis of EoE using less invasive and lower-cost biopsy protocols”</td>
<td>3</td>
</tr>
</tbody>
</table>
Table 1. Cont.

<table>
<thead>
<tr>
<th>Study</th>
<th>Type</th>
<th>Outcome</th>
<th>AI Model</th>
<th>Data and Sample Size</th>
<th>Study Results/Validation Cohort</th>
<th>QUADAS Quality Assessment (Strong Study Is &gt; 6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strbkova et al., 2020</td>
<td>Manuscript</td>
<td>Using time lapse images to classify cells as entering endothelial-mesenchymal transition (EMT) in real time, which correlates with strictures in EoE.</td>
<td>Compared 19 AI models, including various decision trees, k-nearest neighbor neural networks, and supervised vector machines</td>
<td>180 cells monitored every 5 min over 48 h period</td>
<td>Models averaged 98% accuracy at predicting cells going through EMT, improving future studies on stricturing EoE.</td>
<td>10</td>
</tr>
<tr>
<td>Rommele et al., 2021</td>
<td>Abstract</td>
<td>Endoscopic diagnosis</td>
<td>ResNet CNN, comparing image only vs. image with EREFS score augmented</td>
<td>1272 endoscopic white light images (410 EoE)</td>
<td>EREFS-augmented model was strongest, with sensitivity, specificity, and F1-score of 0.85, 0.95, and 0.86, respectively.</td>
<td>8</td>
</tr>
<tr>
<td>Guimaraes et al., 2022</td>
<td>Abstract</td>
<td>Endoscopic diagnosis (EoE vs. Candida vs. control)</td>
<td>CNN with deep Taylor decomposition</td>
<td>484 endoscopic images from 134 patients</td>
<td>Diagnostic accuracy of 92%, sensitivity of 87%, specificity of 94%, better than endoscopists.</td>
<td>8</td>
</tr>
<tr>
<td>Okimoto et al., 2019</td>
<td>Manuscript</td>
<td>Endoscopic diagnosis</td>
<td>ResNet50</td>
<td>2384 endoscopic images (1192 EoE)</td>
<td>95% accuracy, 91% sensitivity, and 97% specificity.</td>
<td>10</td>
</tr>
<tr>
<td>Wang et al., 2020</td>
<td>Manuscript</td>
<td>Automatic segmentation of esophageal OCT images from guinea pigs with EoE, diagnosis by layer width</td>
<td>Several CNNs including Segnet, U-net, pix2pix, and adversarial convoluted network (ACN)</td>
<td>1100 OCT images from five healthy and two animals with EoE</td>
<td>ACN outperformed other models, with 97% accuracy at segmenting esophageal tissue layers, basal layer significantly larger in EoE guinea pigs.</td>
<td>10</td>
</tr>
<tr>
<td>Ryu et al., 2019</td>
<td>Abstract</td>
<td>EoE diagnosis using images from tethered capsule using reflectance endomicroscopy</td>
<td>CNN</td>
<td>200 images with labeled regions of hypereosinophilia</td>
<td>86% accurate at identifying HPF-sized images positive or negative.</td>
<td>6</td>
</tr>
</tbody>
</table>
4. AI for Biopsy Analysis

The widely accepted diagnostic criteria of peak eosinophil count (PEC) greater than 15 per high power field (HPF) is essential for a diagnosis of EoE. Pathologists have also developed an EoE Histology Scoring System (HSS) to quantify the presence of other classic EoE tissue findings such as fibroblasts in the lamina propria [28]. Documenting PEC and HSS can be time consuming, and studies have found significant inter-observer differences in PEC [29]. Multiple AI models have successfully predicted PEC [13-15].

Convolutional neural networks (CNNs) are supervised machine learning algorithms that can be used to classify images, including calculating PEC. Another group of scientists used a smaller dataset including 4345 images from 23 patients with EoE, where a trained pathologist labeled each pixel as including an intact eosinophil, non-intact eosinophil, or non-eosinophil. Using a U-Net model, they could predict EoE diagnosis with 95% accuracy [18]. Differentiating intact vs. non-intact eosinophils (which are not included in PEC) can be challenging for pathologists, and this model correctly identified 98.8% of intact eosinophils. Similarly, Catalano et al. found a 96% accuracy in diagnosing EoE using 274 patches (500 × 500 pixels) from 91 esophageal biopsies [13]. Their accuracy in eosinophil quantification had an impressive average error of + 0.16 and standard deviation of 1.3, showing that AI can accurately predict cellular quantification. Improvements were made by Adorno et al. who compared eleven CNNs built using a data set of over 47,000 500 × 500 pixel patches [14]. Their most powerful U-Net model predicted EoE diagnosis with 99.0% accuracy, 100% sensitivity, and 98.2% specificity. Javaid et al. used a similarly sized dataset and replicated a CNN model with 99% accuracy, this time with a SD for eosinophil quantification of 0.3 [15]. Authors conclude that CNNs can responsibly take the place of manual eosinophil counts.

While CNNs show promise in replacing manually calculated PEC, they may also measure other variables that are tedious to collect manually, such as average eosinophil count per WSI, total number of eosinophilic patches, average eosinophil size, and quantification of PEC over 60. One CNN used these variables to differentiate symptomatic vs. asymptomatic EoE and successfully predicted treatments leading to remission [14]. Catalano et al. also found that a high average eosinophil count predicted successful treatment with four or six food elimination diets, with higher eosinophil counts in the middle compared to the distal esophagus [13]. Histological differences between treatment responders were again found by a model constructed by Adorno et al., who found that higher PEC, higher percent of HPFs with >60 eosinophils, and larger eosinophils correlated with four or six food elimination diet success over PPI and steroids [14]. A similar CNN was able to differentiate stricturing from PPI-responsive EoE, which may aid in prognostication [15]. By relying
solely on the cut off of 15 eosinophils per HPF, valuable information for making clinical predictions may be lost.

Histologic characteristics besides eosinophils have been used in creating AI models. One publication compared multiple CNNs to predict EoE in both cropped and downscaled biopsy whole slide images (WSI), with the best algorithm finishing with an accuracy of 85% [16]. Cropped images of the 147 EoE positive slides that did not include eosinophils and zoomed out images where eosinophils were not clearly visible seemed to contribute to the model, suggesting that global tissue structure contributed to the prediction. Researchers have also constructed CNNs to identify characteristics used in the HSS, with less accurate results [17]. That said, this study may have been limited by a small dataset, including biopsies from only 15 patients with EoE. Similar algorithms may be eventually used to spotlight HSS features and expedite analysis for pathologists.

Many suspected EoE patients have equivocal PEC and HSS findings on initial biopsy [30], and AI algorithms have explored additional approaches to tissue analysis. One group used RNA transcript analysis to predict active EoE vs. GERD vs. normal tissue. Their model calculated a p (EoE) score or probability of EoE for each sample and resulted in 85% accurate diagnosis of patients with equivocal initial biopsy results [19]. Their algorithm also quantified the extent of IgE contribution to inflammation, which could potentially predict treatment response to newer IgE-blocking monoclonal therapies [31]. In a separate publication, they created a model using RNA transcripts from 215 patients with EoE to predict whether the patient had a history of food impaction. Their model resulted in 93% sensitivity and 100% specificity, showing that AI may help flag patients prospectively for food impaction [20]. Another team replicated this model, finding that RNA sequencing could be used to accurately predict EoE from less invasive buccal mucosal biopsies [21].

Machine learning also has applications to understanding pathophysiology of EoE and development of new therapeutics. The worst complications of untreated EoE are related to esophageal remodeling, which is correlated with the transition of epithelial cells to assume mesenchymal cell properties. Epithelial–Mesenchymal transition (EMT) can be difficult to classify, and machine learning algorithms have been used to understand and quantify this process in real time [22]. These researchers followed 180 cells with digital holographic microscopy images taken every 5 minutes over 48 hours and were able to predict cells which have entered EMT with 98% accuracy. Interleukin-13 inhibitors have been shown to prevent EMT in EoE, using AI models to quantify histologic changes [32].

5. Endoscopic Imaging

In order to avoid invasive biopsies, many researchers have hoped for diagnostic classification using endoscopic imaging, such as the EoE Reference Score (EREFS) [33]. This classification system includes features such as rings, exudates, furrows, and edema and can diagnose EoE with 89% accuracy. EREFS scores can also identify inactive vs. active disease posttreatment. Unfortunately, some estimate that 10–20% of children with EoE present with no gross abnormalities on endoscopy, and biopsy is still required for diagnosis [34]. AI using endoscopic imaging has proven effective for diagnosis of Barrett’s esophagus and squamous cell carcinoma of the esophagus [35].

Researchers trained a CNN using endoscopic white light images of EoE patients and controls, in addition to manually calculated EREFS scores. They found that when the model was augmented with EREFS scores it predicted EoE with a sensitivity of 0.85 and specificity of 0.95, compared to a non-augmented model which had a sensitivity of 0.72 and specificity of 0.96 [23]. They suggest that their tool may be improved through application to endoscopic videos in the future. Modest improvements in analyzing endoscopic white light images were replicated by another group of German researchers, which then used deep-Taylor decomposition to understand that the CNN was using similar areas of images to make predictions as endoscopists [24]. Calculating EREFS scores can be tedious for endoscopists, and CNNs may improve efficiency while preserving accuracy [25]. They hypothesized that their model might improve using narrow band imaging, instead of
a standard white light camera, although no publication has used this approach to date. Another new approach on the horizon for EoE endoscopy Optical Coherence Tomography (OCT). Similar to ultrasound but using light waves instead of sound, OCT provides detailed images of nearby tissue. CNNs have also been used to analyze OCT images shot from endoscopic probes. Artificial intelligence has been used to identify and measure the five layers of esophageal tissue as basal layers are thickened in EoE. To date, automated tissue segmentation has only been performed in animal models of EoE, and CNNs were better able to identify the increased basal lamina thickness associated with EoE than manual scoring [26,36]. Their most predictive model used an adversarial convoluted network (ACN) in which the computer purposely tries to trick the algorithm by feeding the model challenging test cases. This model was able to identify esophageal tissue layers with 97% accuracy.

Even if endoscopic imaging could diagnose EoE, the costs and risks of sedation would remain. Researchers have developed a swallowable tethered capsule equipped with spectrally-encoded reflectance microscopy that could identify the reflective properties of individual eosinophils [37]. The capsule is swallowed without sedation and pulled out with an attached string. The process generates many images which can be time consuming to analyze manually, and researchers designed a CNN to look for >15 eosinophils per approximated HPF [27]. The diagnosis of EoE was 86% accurate, comparable to other endoscopic diagnostics. Further refinement of the capsule camera and reflectance microscopy could mean tethered capsules hopefully improve predictive value. Endocytoscopy is a process where images are taken from an ultra-magnified endoscope system, allowing surface epithelial visualization in real time. This approach has been coupled with CNNs in diagnosing precancerous esophageal lesions and may show promise in describing EoE lesions as well [38].

6. Non-Invasive Diagnosis

Many studies have explored non-invasive EoE diagnostics, but very few biomarkers have correlated with disease activity. Swallowed and retrievable strings and sponges capturing luminal eosinophil-derived proteins, as well as absolute eosinophil count and eosinophil progenitor cells in peripheral blood samples, show promise, but sensitivity and specificity to EoE diagnoses are still inferior to EREFS scores [39–42]. While current biomarkers on their own lack accuracy, panels of blood tests and reported symptoms have been combined with multivariate analysis to accurately separate small groups of treatment responders from non-responders [43].

Artificial intelligence combined with large datasets has the potential to improve the effectiveness of biomarker panels beyond linear regression. Wechler et al. used a random forest model to combine six eosinophil associated proteins in the blood and two urine biomarkers to accurately predict EoE 84% of the time [44]. Their model also found a new biomarker with significant association with positive EoE. Unsupervised cluster models have also found distinctions between food impactors and other phenotypes of EoE using peripheral blood transcriptome analysis and microarrays [45]. Cluster analysis has already contributed to the understanding of EoE phenotypes and unsupervised machine learning will play a role in discoveries of new EoE variants [8].

7. Future Application

The future of EoE management will require personalized care which can be delivered by AI. Preventative measures are possible with an AI that flags patients at risk using data such as birth and family history, biomarker profiles, as well as genome and microbiome data, and such models have proven useful in application to other rare disorders [10,46]. Regression models have been used to predict features missing from shallow biopsies, and AI will likely decrease the percent of equivocal biopsies [47]. Studies have shown that EoE patients exhibit altered esophageal microbiomes, and AI may help manage future applications of microbiome data [48]. After diagnosis, it can be difficult to predict offending
food allergens and the best type of initial food elimination diet. AI has already been used to predict polysensitization patterns in allergies, and combining demographic, histologic, and endoscopic data will likely yield useful predictions in EoE sensitization [49]. Nanobiosensors have been developed to screen food for allergens and diet compliance, and AI can synthesize large data outputs [50]. As EoE is a chronic disease, precise follow up schedules can be suggested using machine learning algorithms in order to prevent remodeling and strictures.

While the future is promising, AI applications to EoE share concerns with any AI application to medicine. While AI will assuredly work with physicians to solve clinical problems, many leaders in the field are concerned about losing their jobs to a computer and may resist change. AI is considered a “black box”, and it can be difficult to understand factors that lead to mistakes when they occur. Databases used to train AI may lack diversity and not be as beneficial for minority groups [51]. Although most data are anonymized, there is a risk of personal data being leaked as well as adversarial attacks which distort AI feedback to confuse a model [52].

In conclusion, CNNs have proven successful in predicting PEC and diagnosing EoE using other characteristics on histology, while the accuracy of CNNs using endoscopic photos must be improved. As databases grow and more predictive variables are identified, there is potential for diagnostic prediction using blood and urine panels. AI will continue to play a larger role in EoE management through new diagnostic and surveillance modalities that decreases risks of repeated endoscopies and anestheisa exposure.

Author Contributions: Conceptualization, E.R.S. and J.S.; Methodology, E.R.S. and J.S.; Writing-Original Draft, E.R.S.; Reviewing and Editing, J.S.; Project Supervision, J.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References


