Development and Validation of Web-Based Tool to Predict Lamina Propria Fibrosis in Eosinophilic Esophagitis

Girish Hiremath, MD, MPH[1], Lili Sun, PhD[2], Hernan Correa, MD[3], Sari Acra, MD, MPH[1], Margaret H. Collins, MD[4], Peter Bonis, MD[5], Nicoleta C. Arva, MD, PhD[6], Kelley E. Capocelli, MD[7], Gary W. Falk, MD, MS[8], Eileen King, PhD[9], Nirmala Gonsalves, MD[10], Sandeep K. Gupta, MD[11], Ikuo Hirano, MD[10], Vincent A. Mukkada, MD[12], Lisa J. Martin, PhD[12], Philip E. Putnam, MD[12], Jonathan M. Spergel, MD, PhD[13], Joshua B. Wechsler, MD, MS[1], Guang-Yu Yang, MD, PhD[14], Seema S. Aceves, MD, PhD[15], Glenn T. Furuta, MD[1], Marc. E. Rothenberg, MD, PhD[16], Tatsuki Koyama, PhD[2] and Evan S. Dellon, MD, MPH[17]

INTRODUCTION: **Approximately half of esophageal biopsies from patients with eosinophilic esophagitis (EoE) contain inadequate lamina propria, making it impossible to determine the lamina propria fibrosis (LPF). This study aimed to develop and validate a web-based tool to predict LPF in esophageal biopsies with inadequate lamina propria.**

METHODS: **Prospectively collected demographic and clinical data and scores for 7 relevant EoE histology scoring system epithelial features from patients with EoE participating in the Consortium of Eosinophilic Gastrointestinal Disease Researchers observational study were used to build the models. Using the least absolute shrinkage and selection operator method, variables strongly associated with LPF were identified. Logistic regression was used to develop models to predict grade and stage of LPF. The grade model was validated using an independent data set.**

RESULTS: **Of 284 patients in the discovery data set, median age (quartiles) was 16 (8–31) years, 68.7% were male patients, and 93.4% were White. Age of the patient, basal zone hyperplasia, dyskeratotic epithelial cells, and surface epithelial alteration were associated with presence of LPF. The area under the receiver operating characteristic curve for the grade model was 0.84 (95% confidence interval: 0.80–0.89) and for stage model was 0.79 (95% confidence interval: 0.74–0.84). Our grade model had 82% accuracy in predicting the presence of LPF in an external validation data set.**

DISCUSSION: **We developed parsimonious models (grade and stage) to predict presence of LPF in esophageal biopsies with inadequate lamina propria and validated our grade model. Our predictive models can be easily used in the clinical setting to include LPF in clinical decisions and determine its effect on treatment outcomes.**

SUPPLEMENTARY MATERIAL accompanies this paper at http://links.lww.com/AJG/C336, http://links.lww.com/AJG/C337, http://links.lww.com/AJG/C338, http://links.lww.com/AJG/C339, http://links.lww.com/AJG/C340, http://links.lww.com/AJG/C341, http://links.lww.com/AJG/C342

Am J Gastroenterol 2022;117:272–279. https://doi.org/10.14309/ajg.0000000000001587

INTRODUCTION

Eosinophilic esophagitis (EoE) is an allergen-mediated chronic inflammatory condition affecting the esophagus (1). It is estimated to affect 1 in 2,000 individuals in the United States (2). Children affected by EoE typically present with feeding difficulties, vomiting, and abdominal pain due to the inflammatory phenotype (3,4). A

[1]Division of Pediatric Gastroenterology, Hepatology and Nutrition, Monroe Carell Jr. Children's Hospital at Vanderbilt, Vanderbilt University Medical Center, Nashville, Tennessee, USA; [2]Department of Biostatistics, Vanderbilt University Medical Center, Nashville, Tennessee, USA; [3]Division of Pediatric Pathology, Monroe Carell Jr. Children's Hospital at Vanderbilt, Vanderbilt University Medical Center, Nashville, Tennessee, USA; [4]Department of Pathology and Laboratory Medicine, Cincinnati Children's Hospital, Cincinnati, Ohio, USA; [5]Wolters Kluwer Health, Waltham, Massachusetts, USA; [6]Department of Pathology and Laboratory Medicine, Ann & Robert H. Laurie Children's Hospital at Chicago, Illinois, USA; [7]Department of Pathology, Children's Hospital Colorado, Aurora, Colorado, USA; [8]Department of Medicine, Perelman Center for Advanced Medicine, Philadelphia, Pennsylvania, USA; [9]Division of Biostatistics and Epidemiology, Department of Pediatrics, Cincinnati Children's Hospital, Cincinnati, Ohio, USA; [10]Division of Gastroenterology and Hepatology, Northwestern Medicine, Chicago, Illinois, USA; [11]Division of Pediatric Gastroenterology, Hepatology, and Nutrition, Indiana University School of Medicine, Indianapolis, Indiana, USA; [12]Department of Pediatrics, Cincinnati Children's Hospital, Cincinnati, Ohio, USA; [13]Division of Allergy-Immunology, Children's Hospital of Philadelphia, Perelman School of Medicine, University of Pennsylvania, Pennsylvania, USA; [14]Division of Anatomical Pathology, Northwestern Feinberg School of Medicine, Chicago, Illinois, USA; [15]Division of Rheumatology, Allergy and Immunology, Rady Children's Hospital, San Diego, San Diego, California, USA; [16]Division of Allergy and Immunology, Cincinnati Children's Hospital, Cincinnati, Ohio, USA; [17]Division of Gastroenterology and Hepatology, University of North Carolina at Chapel Hill School of Medicine, Chapel Hill, North Carolina, USA. **Correspondence:** Girish Hiremath, MD, MPH. E-mail: Girish.Hiremath@vumc.org.
Received July 1, 2021; accepted November 19, 2021; published online December 21, 2021

delay in diagnosis or suboptimal treatment may lead to persistent eosinophilic inflammation (EI) and involvement of the subepithelium and lamina propria fibrosis (LPF), which in turn can result in esophageal remodeling or the fibrostenotic phenotype (5,6). As such, adolescents and adults with EoE typically present with dysphagia and esophageal food impaction requiring endoscopic interventions (7,8).

The EoE diagnostic guidelines recommend multiple esophageal mucosal biopsies from 2 or more levels to optimize the diagnostic yield (9,10). Subsequently, the esophageal biopsies are assessed for the intensity of EI by the peak eosinophil count per high-power field (eos/hpf). With the recognition that the disease severity does not strongly depend on the intensity of the eosinophilic infiltration alone, the EoE histology scoring system (EoEHSS) was developed to quantify the grade (degree) and stage (extent) of EoE-relevant histologic changes in epithelium and subepithelium (11,12). However, nearly half of the esophageal mucosal biopsies obtained in routine clinical practice by using standard forceps had inadequately sampled subepithelium (13). This makes it impossible to assess the extent of LPF—a key histologic feature of the esophageal remodeling process (12). As such, developing approaches to predict LPF in esophageal biopsies with inadequate lamina propria sampling can facilitate clinical decision making and inform treatment choices to more accurately assess treatment response, prevent future complications, and improve the clinical outcomes.

We previously reported a high concordance between the presence of surface epithelial alteration (SEA) and dyskeratotic epithelial cells (DEC) and the presence of LPF in children with EoE (14). Based on these observations, we hypothesized that in EoE, the level of involvement of certain esophageal epithelial features can predict LPF in situations where it is impossible to determine the status of lamina propria. Therefore, the aims of this study were to construct and validate computational models to accurately predict LPF in esophageal biopsies with inadequate lamina propria. We additionally sought to share our prediction model with the healthcare community as a web-based tool to facilitate management of their patients with EOE.

METHODS

Ethical considerations

All participants provided consent to partake in the Outcome Measures for Eosinophilic Gastrointestinal Diseases across Ages (OMEGA) study and for future use of their samples and data, as per both central institutional review board (IRB) and local IRB requirements. This study is a secondary analysis of these data and was approved by the IRB at Cincinnati Children's Hospital Medical Center (CCHMC) and at the Vanderbilt University Medical Center.

Data source

We analyzed the demographic, clinical, and histologic data collected as part of the OMEGA study—a multicenter, observational study, aimed at understanding the natural history of EoE and other non-EoE eosinophilic gastrointestinal diseases such as eosinophilic gastritis and colitis. This study was conducted from 2015 through 2019 under the auspices of the Consortium of Eosinophilic Gastrointestinal Disease Researchers (CEGIR)—a national collaborative network of 13 academic centers caring for adults and children with eosinophilic gastrointestinal diseases (15,16).

Model development

Patients with a diagnosis of EoE who were aged 3 years or older were included in the discovery data set. Diagnosis of EoE was based on the published criteria of having symptoms of esophageal dysfunction and presence of a peak of ≥15 eos/hpf in at least one of their multiple esophageal biopsies (9). Patients could be at any point in their EoE course (newly diagnosed, active off treatment, successfully treated, or treatment-refractory) and with any level of disease activity. Patients with a history of intestinal surgery other than G tube placement, planned or recent enrollment in blinded investigational studies, esophageal stricture < 3 mm, other identifiable potential causes for esophageal eosinophilia. Individuals with any physical, mental, or social condition or a history that might have interfered with study procedures or the ability of the subject to adhere to and complete the study were excluded. However, this did not directly affect this study because we performed secondary analysis of the data already collected in the original study.

The esophageal biopsies were collected during a clinically indicated esophagogastroduodenoscopy (EGD) and were taken at the discretion of the gastroenterologist performing the endoscopy. The location of acquisition of the esophageal biopsies was noted. The biopsies were processed locally, and the slides were then scanned using the 2-dimensional Aperio Digital Pathology Slide Scanner (Leica Biosystems, Buffalo Grove, IL). The scanned images were sent to the CCHMC. The CEGIR Central Review Pathology committee accessed the images stored on a CCHMC server.

The pathology committee was comprised of 3 pathologists with expertise in EoE. They assessed the scanned images for the grade and stage of tissue pathology per the EoEHSS. The EoEHSS assesses 8 EoE relevant histologic features: EI, basal zone hyperplasia (BZH), eosinophilic abscess (EA), eosinophilic surface layering (ESL), dilated intercellular spaces (DIS), SEA, DEC, and LPF when adequate lamina propria was available for evaluation. Each feature is scored on a 4-point scale (0–3) for severity (grade) or extent (stage) of the abnormality, with 0 representing normal features and 3 denoting most severe or extensive pathology (11). EoEHSS has been shown to have excellent interobserver and intraobserver reliability (17,18).

To develop our prediction model, we used the initial biopsy with adequate lamina propria from patients with EoE after being included in the OMEGA study. This allowed us to ascertain the reliability of our data and eliminate the confounding effect of intraindividual relationship in patients who underwent EGD with biopsies at multiple time points. We used the peak score for each of the 8 epithelial features irrespective of the location of the sampling in the esophagus. This allowed us to develop prediction models based on the epithelial features and optimized for real-world clinical application. We excluded patients in whom lamina propria could not be assessed due to insufficient sampling because their information would not have contributed toward model development.

External validation

The external validation data set comprised demographic, clinical, and histologic information of children (aged 3–18 years) with EoE undergoing EGD with biopsies at Monroe Carrell Jr. Children's Hospital at Vanderbilt between 2017 and 2018. Details of this study have been previously published (14). In brief, in this study, multiple biopsies were collected from the proximal and distal esophagus from children with EoE undergoing EGD for clinical care at the discretion of their pediatric gastroenterologist. Each of the biopsies were examined. The fragment with most

prominent and abnormal histologic and architectural changes and the highest amount of eosinophilic infiltration in a given subject was scored for the EoEHSS grade score. This was the same protocol as used by the CEGIR pathologists. The EoEHSS stage score was not collected in this study.

Statistical analysis

Descriptive statistics including counts and percentages for categorical variables and medians and quartiles for continuous variables were used to summarize the characteristics of patients included in the discovery and validation data sets.

We first investigated whether the grade and stage scores could be used interchangeably by examining the agreement between peak grade and peak stage scores. Differences between grade and stage scores were calculated for each pair, and we *a priori* determined that to be used interchangeably, at least 80% of the grade–stage pairs for each of the features had to be in agreement. Next, we computed 2 Spearman correlation matrix, one each for grade and stage scores, to investigate the relationship between each of the features in all patients and in patients stratified by their LPF status (LPF = 0 or absent, and LPF = 1, 2, or 3 or present).

Given the high-dimensional nature of the data, we used the least absolute shrinkage and selection operator (LASSO) method to simultaneously select the variables and estimate their regression coefficients. LASSO is a type of analysis that uses shrinkage to force some regression coefficients to be zero. It is particularly well-suited for automating certain parts of the model selection such as the variable selection or parameter elimination (19,20). As such, LASSO allowed us to develop a simple and sparse prediction model (i.e., a model with fewer predictor variables) while retaining the highest ability to predict LPF in biopsies with inadequate lamina propria.

The predictor variables comprised of patient characteristics (including age at biopsy collected used in model development, sex, and race), clinical factors (environmental allergies, duration of EoE monitored defined as age at biopsy collected used in model development – age at diagnosis of EoE, and ongoing EoE treatment), and the 7 epithelial features were fit into LASSO models to predict absence or presence of LPF (absence = 0, presence = 1). The area under the receiver operating characteristic curve (AUC) was calculated to assess how well the parsimonious models classified the presence or absence of LPF.

We performed sensitivity analysis to determine the robustness of our prediction model. In this analysis, we assessed the strength of agreement between the epithelial features, particularly the features selected in our prediction models, in the esophageal biopsies collected from proximal, middle, and distal esophagus. Furthermore, because treatment can interfere with LPF, we examined the effect of treatment status (treatment naive vs on treatment) and individual treatment approaches (topical steroids, elimination diet, empiric elimination diet, and elemental diet) on the predictors. Finally, the prediction model for grade of LPF was externally validated using our single-center EoEHSS grade scores in children with EoE. All analyses were performed in R Statistical Software (version 4.0, R foundation for Statistical Computing, Vienna, Austria).

RESULTS

CEGIR data

In all, 1,253 esophageal biopsies (proximal: 511, middle: 156, and distal: 586) collected from 419 patients were included in the CEGIR data set. Of these, the lamina propria was adequately sampled and assessed in 614 biopsies (proximal: 255, middle: 77, and distal: 282) collected from 284 patients. The peak value for each epithelial feature per EoEHSS per subject was included in the discovery data set to construct the prediction models (Figure 1).

Cohort characteristics

In the discovery data set, 93.4% were White, 68.7% were male patients, and the median (quartiles) age was 16 (8–31) years. More than half (58%) of the patients in the discovery data set were in the pediatric age group (younger than 18 years). The external validation data set comprised 87 children. Most of them were White (75.9%) and male patients (77.0%), and the median (quartiles) age was 10 (7–13) years. Similar proportion of patients in discovery (54%) and validation (47%) data set had EoE for ≥24 months. A significantly higher proportion of patients in the discovery data set were on swallowed topical steroids when compared with that of the validation data set (60.9% vs 12.6%) (Table 1). The peak grade and stage scores among proximal, middle, and distal esophageal biopsies included in the analyses are summarized in Supplementary Table 1 (see Supplementary Digital Content 3, http://links.lww.com/AJG/C338).

Agreement and correlation between grade and stage scores

The probabilities of agreement between grade and stage scores for each of the features did not meet our *a priori* threshold of ≥80% to be used interchangeably. The peak grade and stage score agreement was less than 80% for EI (53%), BZH (74%), DIS (52%), and LPF (68%). For DIS and EI, the grade score was higher than the stage score (26% and 33%, respectively) (see Supplementary Table 2, Supplementary Digital Content 4, http://links.lww.com/AJG/C339).

The correlation between the peak grade and peak stage scores was high (≥0.88) for EI, BZH, EA, ESL, SEA, DEC, and LPF and was low (0.58) for DIS. On stratified analysis, in patients with LPF = 0, a moderate correlation was noted between BZH and EI for grade scores (0.66) and a strong correlation was noted between BZH and EI (0.71) for stage scores. A moderate correlation was also noted between grade and stage scores for ESL and EI (0.51 and 0.53), BZH (0.54 and 0.52), and EA (0.53 and 0.51), respectively. By contrast, the only modest correlation was noted for grade scores between BZH and EI (0.52) (see Supplementary Figure 1, Supplementary Digital Content 1, http://links.lww.com/AJG/C336). Based on these findings, we determined that the grade and stage scores could not be used interchangeably to develop a prediction model.

Parsimonious model to predict LPF

Using the LASSO approach, age of the patient, BZH, DEC, and SEA were identified as the variables that were associated with LPF, for both grade and stage scores (see Supplementary Table 3, Supplementary Digital Content 5, http://links.lww.com/AJG/C340). These variables were fit into separate prediction models. The AUC for grade model was 0.84 (95% confidence interval [CI]: 0.80–0.89) and for stage model was 0.79 (95% CI: 0.74–0.84) (Figure 2). The link to the web-based prediction tool is https://ls2021.shinyapps.io/pre_lpf/.

Sensitivity analysis

Agreement between epithelial features. In all, 47 patients had esophageal biopsies collected from proximal, middle, and distal sites. The epithelial features identified by LASSO as predictors of LPF were in strong agreement across all levels for the grade scores

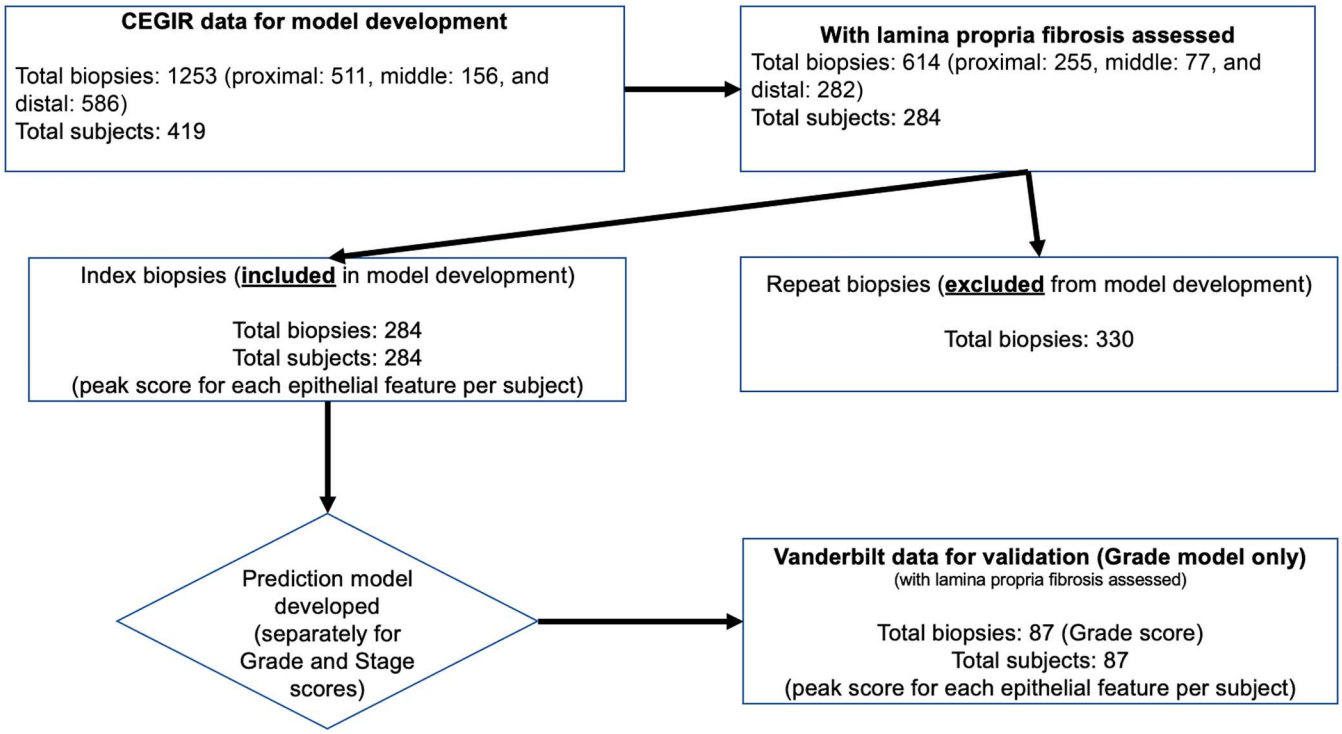

Figure 1. Schematic illustration of the study. CEGIR, Consortium of Eosinophilic Gastrointestinal Disease Researchers.

(BZH: 57%–66%; DEC: 91%; and SEA: 81%–85%) and the stage scores (BZH: 60%–70%; DEC: 89%–94%; and SEA: 81%–85%) (see Supplementary Table 4, Supplementary Digital Content 6, http://links.lww.com/AJG/C341).

Effect of treatment on model performance. In all, 58 patients (20%) were treatment naive and 226 patients (80%) were on treatment. Of those on treatment, 72 (32%) were on topical steroids alone, 21 (9%) were on elimination diet alone, 19 (8%) were on empiric elimination diet alone, and 2 (<1%) were on elemental diet alone. Approximately 50% of individuals were on a combination treatment.

Neither the treatment status nor exposure to swallowed topical steroids confounded the ability of the predictors (i.e., age, BZH, DEC, and SEA) to predict LPF. Given the small number of patients, we were unable to conduct meaningful analysis to examine the effect of dietary therapy on our prediction model.

External validation. We used grade scores from 87 patients in the Vanderbilt data set to validate our LPF grade prediction model. Our model correctly predicted absence of LPF (grade LPF = 0) in 60 patients (80%) and presence of LPF (grade LPF = 1) in 27 patients (85%), with a cumulative accuracy of 82%. The AUC was 0.78 (95% CI: 0.60–0.95) (see Supplementary Figure 2, Supplementary Digital Content 2, http://links.lww.com/AJG/C337). Because the stage scores were unavailable, we were unable to validate the prediction model for the stage of LPF in this study.

DISCUSSION

In EoE, almost half of the esophageal mucosal biopsies obtained using standard forceps inadequately sample the subepithelial space. This makes it impossible to assess the subepithelial involvement including the LPF. Using a large and diverse data set, we developed highly accurate computational models to predict the presence or absence of LPF, individually for grade and stage models, in esophageal biopsies with inadequate lamina propria. Our models included patients' characteristics and epithelial features as assessed per the EoEHSS. The predictor epithelial features were in strong agreement across the biopsy sites, and the performance of our model was not confounded by treatment status and exposure to swallowed topical steroids. The grade model was externally validated using our single-center data set that comprised children with EoE, and its total accuracy was 82%.

A drawback of assessing esophageal biopsies from patients with EoE per the EoEHSS for both grading and staging of tissue pathology is that it is time consuming and thus impractical in clinical practice setting in contrast to the research setting. This holds true even though the grade and stage scores have been previously believed to track together (21). In this study, we found that the correlation between grade and stage scores was high but the agreement between the 2 scores did not meet our predetermined threshold. This suggests that there can be incongruence between the grading and staging of tissue pathology in EoE, and these metrics may need to be assessed separately.

Previously, in a single-center study involving only pediatric patients with EoE, we reported that peak grade scores of DEC (r = 0.75), SEA (r = 0.70), ESL (r = 0.60), and EI (r = 0.59) and EA (r = 0.52) were associated with LPF in biopsies with adequate lamina propria sampling (14). We also reported that the presence of SEA and DEC strongly correlated with presence of LPF. In this study involving multicenter data comprised both pediatric and adult patients with EoE, we found that BZH, DEC, and SEA were highly associated with LPF in biopsies with adequate lamina propria. We were unable to confirm our previous finding related to the association between ESL, EI, and EA and the presence of LPF.

In our analysis, the duration of EoE monitored was not identified as one of the optimal variables to predict the LPF. Based

Table 1. Characteristics of the patients included in the discovery data set and validation data set

	CEGIR data (discovery data set), N = 284	Vanderbilt data (validation data set), N = 87
Age (yr), median [quartile]	16 [8, 31]	10 [7, 13]
Male sex, n (%)	195 (68.7)	67 (77.0)
Race, n (%)		
White	256 (90.1)	66 (75.9)
African American	11 (3.9)	14 (16.1)
Asian	1 (0.4)	0 (0)
Others/unknown	16 (5.7)	7 (8.0)
Environmental allergy, n (%)		
Yes	120 (42.3)	
No	54 (19.0)	
Unknown	110 (38.7)	
Duration of EoE monitored (mo), n (%)		
≤6	31 (10.9)	27 (31.0)
6–12	31 (10.9)	12 (13.8)
12–24	29 (10.2)	7 (8.0)
>24	154 (54.2)	41 (47.1)
Unknown	39 (13.7)	
Ongoing treatment, n (%)		
Swallowed topical steroids	173 (60.9)	11 (12.6)
Elimination diet	93 (32.7)	
Empiric elimination diet	83 (29.2)	
Elemental diet	27 (9.5)	
Other	41 (14.4)	
Patients from each of the centers, n (%)		
Cincinnati Children's Hospital	61 (31.5)	
Children's Hospital Colorado	54 (19.0)	
Children's Hospital of Philadelphia	6 (2.1)	
Laurie Children's Hospital of Chicago	11 (3.9)	
Northwestern University	16 (5.6)	
Riley Children's Hospital	2 (0.7)	
Rady Children's Hospital	28 (9.9)	
Tuft's Medical Center	19 (6.7)	
University of California San Diego	6 (2.1)	
University of Colorado Denver	15(5.3)	
University of Illinois at Peoria	5 (1.8)	
University of North Carolina	31 (10.9)	
University of Pennsylvania	30 (10.6)	

CEGIR, Consortium of Eosinophilic Gastrointestinal Disease Researchers; EoE, eosinophilic esophagitis.

on the cross-sectional data, the current disease paradigm suggests that the fibrostenotic complications can occur in EoE in a time-dependent manner (3,5). Perhaps, future prospective studies will be able provide more data on the natural history of EoE and the factors associated with fibrostenotic complications. Similarly, the markers of esophageal eosinophilia (EI and EA) were also not selected as optimal variables predictive of LPF. This highlights the ongoing dilemma about the role of eosinophils as a histologic marker of EoE activity (22), and the unmet need to identify more reliable histologic markers of tissue involvement in EoE. Our findings suggest that BZH, DEC, and SEA may serve as efficient histologic markers of EoE activity.

Because LPF can be an early feature even in the absence of overt endoscopic findings such as esophageal narrowing or stricture (23), a variety of approaches are being used to indirectly or directly predict the presence of LPF or fibrosis deeper in the esophageal layers in patients with EoE. For instance, application of EndoFLIP—a novel approach using high-resolution impedance planimetry to determine regional variations in pressure in a cross-sectional area of a hollow organ such as esophagus has revealed that esophagus is less distensible in patients with EoE compared with controls and decreased esophageal distensibility was associated with future food impactions (24–26). Similarly, using specialized forceps to obtain deep esophageal biopsies allowed sampling of subepithelial space in more than 90% of adult patients with EoE (27). However, these approaches are invasive and can be unsafe particularly in children. They may also require specialized equipment that are not widely available and incur considerable expense. Similarly, molecular markers of epithelial-stromal crosstalk and fibrosis in EoE such as upregulation of periostin and transforming growth factor β1–induced plasminogen activator inhibitor-1 in active EoE and its correlation with LPF have shown promise in research setting, but their utility in clinical practice remains to be studied (28,29). In addition, our parsimonious predictive models can be easily used to reliably predict presence or absence of LPF (grade and stage) by inputting the patient's age and maximum score (grade and stage scores separately) for specific epithelial features, which are routinely assessed by pathologists in both clinical and research settings. Our model is available at: https://ls2021.shinyapps.io/pre_lpf/.

Our study has limitations. Although we used data from a relatively large and diverse group of patients with EoE for analysis, we do not know whether the presence of impenetrable LPF affected procurement of LP in esophageal biopsies in the original data set. Given that patients with EoE often have delayed diagnosis or symptoms dating back years before diagnosis, we were unable to use the exact duration of EoE in our models; instead, we used duration of EoE monitored. Next, we did not have sufficient data to develop highly accurate models to predict presence of LPF with more granularity. So, at this point, our models can be used to predict presence or absence of LPF (dichotomous outcome) as opposed to providing a breakdown grade or stage of LPF by subscores (0–3) per the EoEHSS. Similarly, we also had limited data to assess the effect of dietary therapy on the performance of our prediction model. Our models were developed on the data that were already collected as part of an observational study. The models need to be tested longitudinally to assess their performance in a real-world clinical setting. The validation data set was only able to focus on grade scores in children with EoE because the stage was not available. Additional studies are needed to validate our model using the stage scores in children and grade and stage scores in adults.

Despite these limitations, our study has several strengths. We used a large, diverse, prospectively collected, and multiinstitutional data set to develop our prediction models, highlighting the

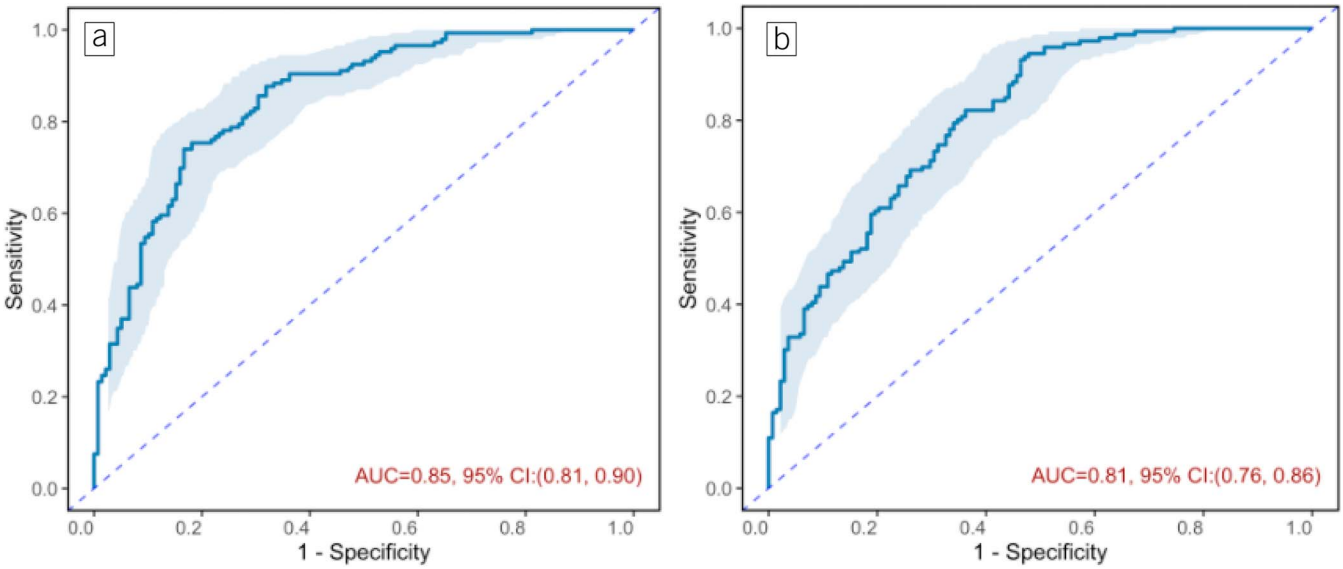

Figure 2. Area under the curve of prediction models: (**a**) grade and (**b**) stage of lamina propria fibrosis. AUC, area under the receiver operating characteristic curve; CI, confidence interval.

value of CEGIR's collaborative infrastructure to develop novel approaches to facilitate a better understanding of the disease natural history and improve clinical outcomes in patients with EoE (30–34). By using LASSO, we were able to optimally use our high-dimensional data for variable selection and model building. Furthermore, this approach has been shown to be superior to usual methods of automatic variable selection such as forward, backward, and stepwise selection for such tasks (35). Because the model is based on the grading and the staging of selected epithelial features at a given time point, predictive ability of our model will be independent of the ongoing EoE therapy. Mirroring a real clinical practice scenario, we corroborated the high accuracy of our prediction model (grade) in an independent single-center data set that included only pediatric patients with EoE, wherein the EoEHSS scoring was performed by an independent pathologist. Taken together, these illustrate the generalizability and clinical applicability of our prediction model. Finally, we have made the prediction tool freely available for the clinical community. We envision that either pathologists could use this to expand on their report or clinicians can use the pertinent patient information and data from their pathology report to inform their management of patients with EoE.

The current disease paradigm suggests that EoE is a chronic, progressive condition and can lead to esophageal remodeling due to LPF. As such, recognizing signs of fibrosis can provide information on the severity and progression of the disease. Finding LPF even when the eosinophilia is controlled could lead to escalation of treatment, a careful search for strictures (e.g., using barium swallow, EndoFLIP, and balloon sizing of the esophagus), or more close monitoring. We envision that our prediction tool would be used in both clinical and research settings. In clinical setting, we anticipate that this will help the clinicians to better understand their patients symptoms, chart the course of their disease, prepare for complications associated with LPF such as persistent dysphagia, future food impactions, and the need for esophageal dilation(s), and inform themselves and patients about the treatment options if there is LPF (e.g., escalating care by considering topical steroids if a patient is on a PPI alone and improving compliance in a noncompliant patient). It will also alleviate the need to collect deeper esophageal biopsies or biopsies with larger forceps to obtain adequate lamina propria and assess its health. In the research setting, we foresee that our prediction tool will aid the researchers to correlate the esophageal distensibility (as measured by EndoFLIP) with LPF and become a part of the therapeutic trials so that the effect of the new drug can be estimated on the health of lamina propria in the setting where adequate lamina propria is unavailable.

In conclusion, we developed prediction models based on the grade or stage of alterations in epithelial features to predict grade or stage scores of LPF in esophageal samples with inadequate lamina propria. Prospective use of models in routine clinical practice, patient-oriented research, and therapeutic drug trials will allows us to assess its performance and further document the effect of LPF on disease progression and clinical outcomes in EoE.

CONFLICTS OF INTEREST

Guarantor of the article: Girish Hiremath, MD, MPH.
Specific author contributions: Study concept and design: G.H., S.A., T.K., E.S.D. Acquisition, analysis, or interpretation of data: G.H., L.S., H.C., S.A., M.H.C., P.B., N.C.A., K.E.C., G.W.F., E.K., N.G., S.K.G., I.H., V.A.M., L.J.M., P.E.P., J.M.S., J.B.W., G.-Y.Y., S.S.A., G.T.F., M.E.R., T.K., and E.S.D. Drafting the manuscript: G.H. and L.S. Critical revision of the manuscript for important intellectual content: G.H., L.S., H.C., S.A., M.H.C., P.B., N.C.A., D.A., K.E.C., G.W.F., E.K., N.G., S.K.G., I.H., V.A.M., L.J.M., P.E.P., J.M.S., J.B.W., G.-Y.Y., S.S.A., G.T.F., M.E.R., T.K., and E.S.D. All authors approved the final draft submitted.
Financial support: G.H. is supported by the American College of Gastroenterology Junior Faculty Development Award and K12 award supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under Award Number K12HD087023. L.S. and T.K. are supported by the Vanderbilt Digestive Disease Research Center NIH Grant DK058404. G.H., M.H.C., P.B., N.C.A., K.E.C., G.W.F., E.K.,

N.G., S.K.G., I.H., V.A.M., L.J.M., P.E.P., J.M.S., J.B.W., G.-Y.Y., S.S.A., G.T.F., M.E.R., and E.S.D. are supported by CEGIR. CEGIR is part of the Rare Diseases Clinical Research Network (RDCRN), an initiative of the Office of Rare Diseases Research (ORDR), National Center for Advancing Translational Sciences (NCATS), and is funded through collaboration between the National Institute of Allergy and Infectious Diseases (NIAID), National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), and NCATS. CEGIR is also supported by patient advocacy groups, including the American Partnership for Eosinophilic Disorders (APFED), Campaign Urging Research for Eosinophilic Diseases (CURED), and Eosinophilic Family Coalition (EFC). The work was independent of these funding sources. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.
Potential competing interests: M.H.C. is a consultant for Allakos, AstraZeneca, Calypso, Esocap, GlaxoSmithKline, Receptos/Celgene, Regeneron, and Shire a Takeda company and has received research funding from AstraZeneca, Shire a Takeda company, Receptos/Celgene, and Regeneron. S.K.G is a consultant for Abbott, Adare, Allakos, Celgene, Gossamer Bio, QOL, UpToDate, Medscape, and Viaskin and has received research support from Shire, Allakos, Adare, and NIH U54 grant to CEGIR. Other authors declare no competing interests related to this study.

Study Highlights

WHAT IS KNOWN

- ✓ In eosinophilic esophagitis (EoE), lamina propria fibrosis (LPF) is central to esophageal remodeling and fibrostenotic complications.
- ✓ However, almost half of esophageal mucosal biopsies do not contain adequate lamina propria, thereby making it impossible to ascertain LPF.
- ✓ Developing an easy and widely applicable approach to predict LPF in esophageal biopsies with inadequate lamina propria sampling can contribute toward improving clinical outcomes in EoE.

WHAT IS NEW HERE

- ✓ Using patient characteristics and the peak grade and stage score for each of the features of the EoE histology scoring system, we developed parsimonious models to predict the presence of LPF (grade and stage) in esophageal biopsies with inadequate lamina propria.
- ✓ The area under the receiver operating characteristic curve of our model to predict of LPF (grade) was 0.84 (95% confidence interval [CI]: 0.80–0.89) and that for the LPF (stage) was 0.79 (95% CI: 0.74–0.84).
- ✓ Our grade model predicted presence of LPF with 82% accuracy in an independent data set (external validation).
- ✓ The prediction model is made available as a web-based tool: https://ls2021.shinyapps.io/pre_lpf/.

REFERENCES

1. Furuta GT, Katzka DA. Eosinophilic esophagitis definition and differential diagnosis. N Engl J Med 2016;373:1640–8.
2. Jensen ET, Kappelman MD, Martin CF, et al. Health-care utilization, costs, and the burden of disease related to eosinophilic esophagitis in the United States. Am J Gastroenterol 2015;110:626–32.
3. Dellon ES, Hirano I. Epidemiology and natural history of eosinophilic esophagitis. Gastroenterology 2018;154:319–32.e3.
4. Lucendo AJ, Sánchez-Cazalilla M. Adult versus pediatric eosinophilic esophagitis: Important differences and similarities for the clinician to understand. Expert Rev Clin Immunol 2012;8:733–45.
5. Schoepfer AM, Safroneeva E, Bussmann C, et al. Delay in diagnosis of eosinophilic esophagitis increases risk for stricture formation in a time-dependent manner. Gastroenterology 2013;145:1230–6.e2.
6. Warners MJ, Oude Nijhuis RAB, de Wijkerslooth LRH, et al. The natural course of eosinophilic esophagitis and long-term consequences of undiagnosed disease in a large cohort. Am J Gastroenterol 2018;113:836–44.
7. Hirano I. Clinical relevance of esophageal subepithelial activity in eosinophilic esophagitis. J Gastroenterol 2020;55:249–60.
8. Menard-Katcher C, Benitez AJ, Pan Z, et al. Influence of age and eosinophilic esophagitis on esophageal distensibility in a pediatric cohort. Am J Gastroenterol 2017;112:1466–73.
9. Liacouras CA, Furuta GT, Hirano I, et al. Eosinophilic esophagitis: Updated consensus recommendations for children and adults. J Allergy Clin Immunol 2011;128:3–20.e6.
10. Dellon ES, Liacouras CA, Molina-Infante J, et al. Updated international consensus diagnostic criteria for eosinophilic esophagitis: Proceedings of the AGREE conference. Gastroenterology 2018;155:1022–33.e10.
11. Collins MH, Martin LJ, Alexander ES, et al. Newly developed and validated eosinophilic esophagitis histology scoring system and evidence that it outperforms peak eosinophil count for disease diagnosis and monitoring. Dis Esophagus 2017;30:1–8.
12. Dellon ES, Katzka DA, Collins MH, et al. Safety and efficacy of budesonide oral suspension maintenance therapy in patients with eosinophilic esophagitis. Clin Gastroenterol Hepatol 2019;17:666–73.e8.
13. Wang J, Park JY, Huang R, et al. Obtaining adequate lamina propria for subepithelial fibrosis evaluation in pediatric eosinophilic esophagitis. Gastrointest Endosc 2018;87:1207–14.e3.
14. Hiremath G, Choksi YA, Acra S, et al. Factors associated with adequate lamina propria sampling and presence of lamina propria fibrosis in children with eosinophilic esophagitis. Clin Gastroenterol Hepatol 2020;19(9):1814–23.e1.
15. Gupta SK, Falk GW, Aceves SS, et al. Consortium of eosinophilic gastrointestinal disease researchers: Advancing the field of eosinophilic GI disorders through collaboration. Gastroenterology 2019;156:838–42.
16. Aceves S, Collins MH, Rothenberg ME, et al. Advancing patient care through the Consortium of Eosinophilic Gastrointestinal Disease Researchers (CEGIR). J Allergy Clin Immunol 2020;145:28–37.
17. Warners MJ, Ambarus CA, Bredenoord AJ, et al. Reliability of histologic assessment in patients with eosinophilic oesophagitis. Aliment Pharmacol Ther 2018;47:940–50.
18. Arias A, Pérez-Martínez I, Tenías JM, et al. Systematic review with meta-analysis: The incidence and prevalence of eosinophilic oesophagitis in children and adults in population-based studies. Aliment Pharmacol Ther 2016;43:3–15.
19. Sirimongkolkasem T, Drikvandi R. On regularisation methods for analysis of high dimensional data. Ann Data Sci 2019;6:737–63.
20. Tibshirani R. Regression shrinkage and selection via the Lasso. J R Stat Soc Ser B 1996;58:267–88.
21. Collins MH, Dellon ES, Katzka DA, et al. Budesonide oral suspension significantly improves eosinophilic esophagitis histology scoring system results: Analyses from a 12-week, phase 2, randomized, placebo-controlled trial. Am J Surg Pathol 2019;43:1501–9.
22. Blanchard C, Simon D, Schoepfer A, et al. Eosinophilic esophagitis: Unclear roles of IgE and eosinophils. J Intern Med 2017;281:448–57.
23. Hirano I, Aceves SS. Clinical implications and pathogenesis of esophageal remodeling in eosinophilic esophagitis. Gastroenterol Clin North Am 2014;43:297–316.
24. Lin Z, Kahrilas PJ, Xiao Y, et al. Functional luminal imaging probe topography: An improved method for characterizing esophageal distensibility in eosinophilic esophagitis. Therap Adv Gastroenterol 2013;6:97–107.
25. Muir AB, Wang JX, Nakagawa H. Epithelial-stromal crosstalk and fibrosis in eosinophilic esophagitis. J Gastroenterol 2019;54:10–8.
26. Hassan M, Aceves SS, Newbury R, et al. Esophageal distensibility as a predictor of clinical phenotype in pediatric patients with eosinophilic esophagitis. Gastroenterology 2017;152:S432.
27. Bussmann C, Schoepfer AM, Safroneeva E, et al. Comparison of different biopsy forceps models for tissue sampling in eosinophilic esophagitis. Endoscopy 2016;48:1069–75.
28. Rawson R, Yang T, Newbury RO, et al. TGF-β1–induced PAI-1 contributes to a profibrotic network in patients with eosinophilic esophagitis. J Allergy Clin Immunol 2016;138:791–800.e4.

29. Williamson P, Proudfoot J, Gharibans A, et al. Plasminogen activator inhibitor-1 as a marker of esophageal functional changes in pediatric eosinophilic esophagitis. Clin Gastroenterol Hepatol 2020. [Epub ahead of print September 30, 2020.] doi: 10.1016/j.cgh.2020.09.040.
30. Shoda T, Wen T, Caldwell JM, et al. Correlation of the eosinophilic histopathological scoring system with esophageal gene expression in patients with eosinophilic esophagitis. J Allergy Clin Immunol 2018;141: AB138.
31. Pesek RD, Reed CC, Muir AB, et al. Increasing rates of diagnosis, substantial co-occurrence, and variable treatment patterns of eosinophilic gastritis, gastroenteritis and colitis based on 10 year data across a multi-center consortium. bioRxiv. 2019;114(6):984–94.
32. Hirano I, Collins MH, King E, et al. 357—Prospective evaluation of a novel, endoscopic activity assessment system for eosinophilic gastritis. Gastroenterology 2019;156:S-71–S-72.
33. Aceves SS, King E, Collins MH, et al. Alignment of parent- and child-reported outcomes and histology in eosinophilic esophagitis across multiple CEGIR sites. J Allergy Clin Immunol 2018;142:130–8.e1.
34. Shoda T, Wen T, Aceves SS, et al. Eosinophilic oesophagitis endotype classification by molecular, clinical, and histopathological analyses: A cross-sectional study. Lancet Gastroenterol Hepatol 2018;3:477–88.
35. Lima E, Davies P, Kaler J, et al. Variable selection for inferential models with relatively high-dimensional data: Between method heterogeneity and covariate stability as adjuncts to robust selection. Sci Rep 2020;10:8002.