Published in final edited form as:

Clin Exp Allergy. 2015 August ; 45(8): 1317–1327. doi:10.1111/cea.12523.

Accuracy of digital mRNA profiling of oesophageal biopsies as a novel diagnostic approach to eosinophilic oesophagitis

Willem S. Lexmond, MD¹, Lan Hu, PhD², Michael Pardo¹, Nicole Heinz¹, Katharine Rooney¹, Jessica LaRosa¹, Eleonora Dehlink, MD PhD^{1,3}, Edda Fiebiger, PhD^{#1}, and Samuel Nurko, MD MPH^{#1}

¹ Division of Gastroenterology and Nutrition, Center for Motility and Functional Gastrointestinal Disorders, and Eosinophilic Gastrointestinal Disease Center Boston Children's Hospital and Harvard Medical School; Boston, Massachusetts

² Center for Cancer Computational Biology, Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute; Boston, Massachusetts

[#] These authors contributed equally to this work.

Abstract

Background—Quantification of tissue eosinophils remains the golden standard in diagnosing eosinophilic oesophagitis (EoE), but this approach suffers from poor specificity. It has been recognized that histopathological changes that occur in patients with EoE are associated with a disease-specific tissue transcriptome.

Objective—We hypothesized that digital mRNA profiling targeted at a set of EoE-specific and Th2 inflammatory genes in oesophageal biopsies could help differentiate patients with EoE from those with reflux oesophagitis (RE) or normal tissue histology (NH).

Methods—The mRNA expression levels of 79 target genes were defined in both proximal and distal biopsies of 196 patients with nCounter® (Nanostring) technology. According to clinicopathological diagnosis, these patients were grouped in a training set (35 EoE, 30 RE, 30 NH) for building of a three-class prediction model using the random forest method, and a blinded predictive set (n=47) for model validation.

Results—A diagnostic model built on ten differentially expressed genes was able to differentiate with 100% sensitivity and specificity between conditions in the training set. In a blinded predictive set, this model was able to correctly predict EoE in 14 out of 18 patients in distal (sensitivity 78%, 95% CI 52%-93%) and 16 out of 18 patients in proximal biopsies (sensitivity 89%, 95% CI 64%-98%), without false positive diagnosis of EoE in RE or NH patients (specificity 100%, 95% CI 85%-100%). Sensitivity was increased to 94% (95% CI 71%-100%) when either the best predictive distal or proximal biopsy was used.

To whom correspondence should be addressed: Samuel Nurko, MD MPH, 300 Longwood Avenue, Hunnewell Ground, Boston, MA 02115, Tel: +1 617 355 6055; fax: 6173557043; samuel.nurko@childrens.harvard.edu..

³Current address: Department of Paediatrics and Adolescent Medicine, Division of Paediatric Pulmonology, Allergy and Endocrinology, Medical University of Vienna, Vienna, Austria.

CONFLICT OF INTEREST

None of the authors have any competing financial interests to disclose

Conclusion & Clinical Relevance—We conclude that mRNA profiling of oesophageal tissue is an accurate diagnostic strategy in detecting EoE.

INTRODUCTION

Quantification of tissue eosinophils under proton-pump inhibitor (PPI) therapy in combination with appraisal of clinical symptomatology remains the gold standard in diagnosing patients with eosinophilic oesophagitis (EoE) (1). Unfortunately, this approach suffers from poor specificity since a wide variety of gastrointestinal disorders can give rise to oesophageal eosinophilia, and EoE-like symptoms such as failure-to-thrive or feeding difficulties are generally non-specific, especially in young children (1,2). In daily practice, the distinction between reflux-associated eosinophil infiltration and true EoE remains particularly difficult, because these disease entities have considerable symptomatic and histologic overlap and are typically patchy in nature.(3) Differentiation can furthermore be complicated by poor patient compliance with PPI therapy prior to diagnostic endoscopy, as well as the recent recognition of PPI-responsive EoE variants (3,4). These diagnostic difficulties often cause considerable delay between onset of disease and diagnosis, hence delaying installment of optimal therapy (5,6). As such, a need for more refined diagnostic methods to reliably differentiate EoE from reflux oesophagitis (RE) and other causes of oesophageal eosinophilia clearly persists (7).

It has been recognized that EoE is associated with a tissue-specific transcriptome and that increased oesophageal mRNA levels of genetic markers such as eotaxin-3 and periostin display considerable diagnostic sensitivity and specificity (8,9). Additionally, we and others have shown that the EoE-transcriptome can be used to identify involvement of specific inflammatory pathways that are operative in subsets of EoE patients (10-12). Following the identification of an EoE-specific transcriptome, it has been hypothesized that an oesophageal mRNA expression pattern across multiple disease-specific and Th2-inflammatory genes can provide additional diagnostic benefit over the current golden standard of histopathological eosinophil count (13). However, analysis of mRNA expression levels of EoE-markers in oesophageal biopsies has not yet been introduced into routine clinical practice. This may in part be due to the relatively challenging requirements for large-scale mRNA tissue diagnostics, which in the case of quantitative RT-PCR requires multi-step processing that is both time and manipulation-sensitive. In the current study, we have taken advantage of a digital mRNA expression profiling method (Nanostring technology, nCounter® system) that has been described to have sensitivity comparable to quantitative RT-PCR, yet does not require tissue extraction procedures or generation of cDNA as an intermediate step (14). Using this direct, state-of-the-art mRNA profiling methodology, we here aimed to define an oesophageal mRNA pattern stamp that can distinguish patients with EoE from those with RE or normal tissue histology (NH) and lead to improved tissue diagnosis.

METHODS

In the reporting of this study, we have adhered to the guidelines put forth by the STARD initiative (15).

Study population

The patients described in this study were included at Boston Children's Hospital as part of a prospective cohort study that focuses on EoE pathophysiology and diagnostics. Details pertaining this study have been previously described (11). Study inclusion was based on clinical presentation. Children (1-18 years of age) who presented with EoE-like symptoms (such as failure-to-thrive, food aversion, regurgitation or dysphagia) and who were scheduled for diagnostic upper oesophagogastrroduodenoscopy were invited to participate. In all patients, endoscopy was performed after a minimum of four weeks of PPI therapy (mean duration 10.8 months, with a range of 1-120 months). Additional patient information was collected through a standardized questionnaire at time of enrollment. For each patient, an additional study biopsy was obtained from both the proximal and distal oesophagus, defined as lying either 10 cm or 1-2 cm from the gastro-oesophageal junction respectively. Biopsies were directly stored in RNAlater (Qiagen) for a minimum of 24h and frozen at -80°C. From our cohort, we randomly selected patients to undergo digital transcriptional profiling of both proximal and distal oesophageal biopsies. Approval for this study was obtained from the Investigational Review Board of Boston Children's Hospital (Harvard Medical School, Boston, MA, approval number: 07-11-0460). All patients or their legal guardians provided written informed consent prior to enrollment.

Clinicopathological diagnosis (reference standard)

A board-certified pediatric gastroenterologist, who was blinded to the results of pattern profiling, reviewed the clinicopathological diagnosis of every patient by taking into account all available clinical information, which included symptomatology at presentation, laboratory testing, the pathology report of oesophageal histology and tissue eosinophil count, as well as the clinical course after initial diagnosis, and the results of follow-up endoscopies if performed. This review took place after a minimum follow-up of two years after the enrollment endoscopy, and was used in our study as the reference standard against which we aimed to validate our novel diagnostic approach (1). *Following consensus guidelines* (1), we considered patients to have *EoE* when they met the following criteria: 1. treatment with PPI for 4 weeks prior to diagnostic endoscopy; 2. tissue eosinophil count >15/hpf in at least one biopsy; 3. exclusion of other origins of oesophageal eosinophilia. Use of corticosteroids was considered an exclusion criteria. Conversely, patients were classified as *RE* when they showed: 1. histological evidence of oesophageal tissue inflammation such as basal zone hyperplasia and an inflammatory cell infiltrate; 2. eosinophil count 1-15/hpf; 3. a clinical history suggestive of reflux-associated symptoms 4. Evidence of pathologic GERD either by abnormal pH/impedance studies, or by erosive oesophagitis that healed after antacid therapy, and 5 no evidence of development of EoE after long term follow up. Lastly, *NH* patients were defined as having: 1. normal tissue histology in all routine biopsies, and 2. no evidence of underlying gastrointestinal disease for at least 3 months after endoscopy in the absence of antacid therapy. Patients that did not meet any of these three diagnostic categories were excluded from training and predictive patient set.

Sample processing and mRNA profiling with the nCounter® system

Biopsies were homogenized in RLT buffer (Qiagen) and further processed with the nCounter® Prep Station and Digital Analyzer, following the manufacturer's instructions (nCounter® system, www.nanostring.com). Samples were analyzed using a customized panel that consisted of five housekeeping genes and 79 genes of interest based on previously published microarray data (8). This code set is summarized in supplemental Table S1. Expression data from separate nCounter® runs were normalized through quantile normalization, and then log2 transformed prior to downstream analysis. Outlying samples with low readout in the internal positive controls were excluded from further analysis

Definition of a training and predictive patient set

A total of 95 unambiguously diagnosed patients were otherwise randomly selected into a training set, which was used to identify differentially expressed genes to include in a diagnostic prediction model. The remaining unambiguous patients were used for the predictive patient test set. For training set patients, both the clinicopathological diagnosis from the reference standard and the mRNA pattern profile were provided to the statistician, who performed differential gene expression analysis and diagnostic model building. For the predictive set, the statistician was blinded to the histopathological diagnosis and only the mRNA profile was provided.

Differential gene expression analysis

Three individual linear statistical models were built (R Bioconductor limma package) to compare training set patients (EoE vs. NH, RE vs. NH, and EoE vs. RE, respectively) to identify genes that were differentially expressed between all three disease conditions (p-value<0.05).

Diagnostic model

A three-class (EoE, RE, and NH) diagnostic model was built with 10-fold cross validation using the random forest method. In each round of the cross validation process, the ratio of EoE, RE, and NH samples was set to be the same as in the complete training set. Once the model for a given biomarker gene set was trained in the training samples, the expression profile of the same biomarker gene set from the predictive set samples was fitted on the trained model and the EoE/RE/NH classification diagnostic probability, i.e. the probability of having each diagnosis, was calculated for the predictive samples. A predicted probability >50% was considered a positive mathematical diagnosis for that particular condition.

Statistical analysis

Comparison of clinical characteristics and probability scores between diagnostic groups was performed with ANOVA or Kruskal-Wallis test for continuous variables or Fisher's exact test for dichotomous predictors. Correlation analysis was performed using Pearson correlation coefficient. Values are expressed as mean $\pm$ SD unless otherwise indicated. Analyses were performed using Stata 12 (StataCorp, TX, USA).

RESULTS

Patient inclusion

The 196 patients analyzed in this study were randomly selected from a previously described (11,16), longitudinal cohort of 429 children who were included at our center between April 2008 and May 2012. For every patient, one proximal and one distal research biopsy was analyzed with the nCounter® system to obtain a digital mRNA expression profile of the oesophageal tissue. The collection of two additional research biopsies was not associated with adverse events.

Patient classification

Histological evidence of oesophageal inflammation was present in 110 out of 196 patients (56%). In 96 (87%) of those, an unambiguous clinicopathological diagnosis of either EoE or RE could be made based on comprehensive case review after a follow-up time of a minimum of two years. These 96 cases were used to form the basis of a training and a predictive patient set. From those unequivocally diagnosed patients, we randomly selected 35 EoE and 30 RE patients, together with 30 randomly selected NH cases to give rise to the training set. The remaining 31/96 unambiguous patients with oesophageal inflammation were combined with another 16 NH patients to generate the predictive test set that we used for validation of the diagnostic model. In 14/110 patients with oesophageal inflammation (13%), no unambiguous diagnosis of either RE or EoE could be established with the available clinical and *histologic* information and were excluded from further analysis. These patients either suffered from non-RE, non-EoE oesophageal inflammation (e.g. candida oesophagitis), or had undocumented or unclear use of a PPI prior to diagnostic endoscopy. We also excluded patients when a diagnosis of EoE was hampered by prior installment of some form of EoE-specific treatment before study enrollment if biopsies did not have the established number of eosinophils to diagnose them as EoE and 2-year follow-up had not revealed a clear underlying condition.

The flow diagram of patient inclusion and distribution over the training and predictive sets is shown in Figure 1. The distal biopsy of one RE patient in the predictive set was excluded from analysis due to overall low mRNA counts. Clinical, endoscopic and histological features of patients in both sets are summarized in Table 1. In line with previous reports (1,11,17), statistically significant differences between diagnostic groups were found in distribution of gender, presenting symptoms, allergic comorbidity and findings during diagnostic endoscopy.

Selection of predictor genes in the training set of patients

To identify which of the 79 genes in our panel were differentially expressed between disease conditions, we first performed comparative gene expression analysis in patients with EoE, RE or NH that were included in the training set. As expected from previous studies (8,11,18), this analysis revealed striking differences between conditions in both proximally and distally-derived biopsies (Figure 2 A, B). Differential gene expression analysis for EoE vs. NH, RE vs. NH and EoE vs. RE patients identified 38 genes in proximal biopsies and 43 genes in distal biopsies that were differentially expressed in at least one comparison at

$\alpha=0.05$ (Figure 2 A, B). We hypothesized *a priori* that most discriminative power would be attained by a model that only included genes that were differentially expressed between all three pathologies. For proximal biopsies, seven genes met this criterion, whereas 10 genes were found to be differentially expressed in distal biopsies at a p -value <0.05 in all three comparisons (Table 2). Because distal biopsies gave rise to a greater abundance of significantly regulated genes, we selected those biopsies to build our diagnostic model. As such, our strategy resulted in inclusion of the major EoE-specific gene eotaxin-3 (CCL26) as well as the mast cell-marker carboxypeptidase A3 (CPA3) and β -chain of the high-affinity IgE receptor (Fc ϵ R1), but not periostin, which although highly upregulated in both biopsies in EoE patients compared to RE or NH patients (t -statistic >5.5 , p -value $<10^{-6}$) was not significantly different between NH controls and RE patients in proximal or distal tissue biopsies (p -value >0.05 , Figure 2). Additional upregulated genes were the eotaxin-3 receptor CCR3, which is mainly expressed on eosinophils and basophils, and the Th2-cytokine IL-13. EoE-specific downregulation of genes was noted for CD4, CCL17, CD207, Fc ϵ R1 α and LGALS3 (galectin-3). In proximal, but not distal, biopsies, statistical significance was reached for CCL22 and CD40 ligand (Figure 2 and Table 2). These results suggested that disease-specific regulation of these ten target genes could potentially be used to discriminate between patients with EoE, RE, and NH.

Definition of a predictive diagnostic model

A three-class statistical model was built to predict the diagnosis of patients from the expression levels of the ten genes that were identified to be significantly regulated in distal biopsies. The primary output of our model was a probability score per patient for each of the three diagnostic categories (NH, EoE or RE), which allowed for quantification of the confidence of the model in predicting the underlying diagnosis. After 10-fold cross validation on distal training set samples, the disease-specific probability scores discriminated perfectly between all three diagnostic categories: mean predicted EoE-probability in EoE patients 0.92 ± 0.09 (range 0.64-1.0); RE-probability in RE patients 0.81 ± 0.08 (range 0.64-0.94) and NH-probability in NH patients 0.87 ± 0.07 (range 0.67-0.96). In patients with EoE, a positive correlation was observed between distal tissue eosinophil count and the predicted EoE probability, which suggests that a more pronounced eosinophilic infiltrate results in better performance of the diagnostic model (Figure 3A). However, even in patients with a modest eosinophilic infiltrate in the distal oesophagus (20-25 eosinophils/hpf), a clear prediction of EoE over RE or NH etiology could be made with probabilities ranging from 0.74 to 0.97. These results indicated that distal mRNA levels of ten differentially expressed genes are sufficiently dissimilar between disease conditions to allow for clear mathematical differentiation.

This finding, however, does not necessarily guarantee a similar degree of performance on biopsies that do themselves not further shape the algorithmic parameters of the model. In assessing how our pre-defined model performed on an independent subset of biopsies, we first further validated our prediction model on the paired proximal biopsies from the patients in the training set. This analysis revealed that application of the distally-defined prediction model on proximal biopsies from the same patients resulted in an equally clear differentiation between patients from all three conditions (Figure 3 B-C), and that a

diagnostic cut-off of >50% EoE probability identified EoE patients from NH or RE patients with 100% sensitivity and specificity in our training set. Importantly, these results are also indicative of a high degree of test reproducibility as proximal and distal biopsies had been independently collected, stored, and processed prior to digital mRNA profiling.

Performance of diagnostic model in the blinded predictive patient set

To further assess the power of our model in differentiating EoE from non-EoE patients, we next applied it to the patients in the predictive subset (Figure 4). In 14 out of 18 patients with EoE, the model correctly predicted the presence of EoE with >50% probability in distal biopsies (sensitivity 78%, 95% CI 52%-93%). Conversely, in none of the NH or RE control patients a diagnosis of EoE was made using the model (specificity 100%, 95% CI 85%-100%). For proximal biopsies, similar results were obtained, with 16 out of 18 patients with EoE correctly identified (sensitivity 89%, 95% CI 64%-98%), again with no false positives in the 29 non-EoE control patients. It is well recognized that EoE is a patchy disease and histopathological analysis of multiple biopsies is required for a reliable diagnosis.^(1,19) When a cut-off of >50% EoE probability in either proximal or distal biopsy was used to identify patients with EoE, sensitivity was increased to 94% (95% CI 71%-100%) without compromising specificity, which remained at 100%. These results indicate that gene expression analysis of multiple biopsies enhances the diagnostic power of this predictive model for EoE.

In contrast to the performance of the model on the proximal biopsies in the training set patients, differentiation between RE and NH patients in this blinded predictive patient set was poor. No significant differences were found in either the NH or RE probabilities between patients with a clinicopathological diagnosis of these conditions (Figure 4). Our diagnostic model therefore failed to identify patients suffering from RE from the controls within the pool of non-EoE patients. It is, however, conceivable that a gene set for RE diagnostic could be defined with our approach.

DISCUSSION

In the current study, we have tested the hypothesis that a diagnosis of EoE can reliably be made based on changes in the mRNA expression levels of a limited number of EoE-specific target genes. Using a diagnostic panel of merely ten targets, the analysis of a single research biopsy could detect EoE with a sensitivity of 78% for distal and 89% for proximal biopsies, which was increased further to 94% when results of two investigated biopsies were combined. Importantly, no false-positive EoE predictions were made amongst 29 normal or inflammatory controls.

The increased sensitivity upon analysis of multiple biopsies is not unexpected given the patchy nature of the disease, and is also true in histological assessment of oesophageal tissue, where it has been shown that sensitivity increases from 73% in a single biopsy, to 84%, 97%, and 100% with obtaining two, three, and six biopsies respectively (20). Our results thus compare favorably to eosinophil count alone, which suggests that the transcriptome signature may be a more sensitive readout of disease involvement than histological evaluation. Although we did observe a positive correlation between eosinophil

count and predicted EoE probability, the model still reliably identified EoE at moderate eosinophil counts of 20-25 eosinophils/hpf. One of the main findings of our current study is therefore that an objective, observer-independent diagnosis of EoE could be made using the oesophageal transcriptome within our study population. We think that this technique will be a useful adjunct to the clinical evaluation of patients with suspected EoE, and may allow establishment of an accurate diagnosis from fewer biopsies and help in equivocal cases.

A rather surprising observation that can be made from this work is the seemingly paradoxical condition of EoE-specific tissue downregulation of the α -chain of the high affinity IgE receptor Fc ϵ RI in combination with upregulation of its β -chain. Increased β -chain expression can be readily explained by esophageal infiltration of mast cells, as has been well documented to occur in EoE (10). Less is known, however, about reduced tissue expression of Fc ϵ RI α . In addition to its expression on mast cells as part of the tetrameric Fc ϵ RI complex, the α -chain is also expressed in the GI-tract by Langerhans cells to form part of a β -chain-independent, trimeric Fc ϵ RI (21). As such, it is conceivable that reduced mRNA levels of Fc ϵ RI α signify lower numbers of Langerhans cells in the esophagus of patients with EoE. In support of this notion, we found significantly lower transcript numbers of CD207 (Langerin, Table 2) and CD1a (results not shown) in EoE patients compared to non-EoE patients. Since CD207⁺ dendritic cells have been associated with the establishment of oral tolerance to ingested antigens (22), it could be hypothesized that altered numbers of these cells contribute to disease pathogenesis and future experimental studies addressing the role of these dendritic cells in EoE are therefore warranted.

One strength of our study is that we were able to follow the included patients for a minimum follow-up of two years, ensuring that the initial clinicopathological diagnosis that was given initially was accurate. Our strategy allowed us to incorporate additional clinical information such as response to treatment, progression of symptoms and follow-up endoscopies. Because of this comprehensive follow-up we were able to validate our novel diagnostic approach against an optimized gold standard reference. Mutual blinding of the reviewing clinician and analyzing statistician eliminated the potential of bias during this process.

Our results indicate that the EoE signature is reliably identified in almost all investigated patients. Unfortunately, even though the model performed well in separating EoE patients from those with RE, the sensitivity and specificity of our model was too low to be considered of diagnostic value in differentiating RE patients from NH controls. This distinction is, however, not a difficult problem in daily clinical practice. To more reliably make a transcriptome-based diagnosis of RE-mediated inflammation, it is likely that a broader gene array panel needs to be screened, since our combination of 76 genes only yielded two genes that were differentially regulated between NH controls and RE but not EoE patients.

Our findings are in line with a recently published study that also tested the diagnostic potential of oesophageal transcriptome changes in adult and pediatric patients with EoE (13). Using a real-time quantitative PCR approach on a panel of 94 EoE-specific genes, of which four (CCL26, CPA3, IL13 and CCR3) are also included in our set of 10 genes, this group attained a 96% sensitivity and 98% specificity rate for a diagnosis of EoE over non-

EoE control patients. The results presented here, showing equivalently high degrees of diagnostic accuracy, thus confirm the great diagnostic potential of the oesophageal transcriptome in detecting EoE, and further validate this approach using a different technical platform in another patient cohort. In addition, our data demonstrate that a diagnosis of EoE can reliably be made with a greatly reduced number of target genes and that analysis with a novel high-throughput mRNA expression system that eliminates the need for intermediary sample processing steps is a feasible approach to obtaining a tissue diagnosis of EoE. Although restricted availability and costs of this technology limit its current applicability, it is possible that with technological advances these costs will decrease in the future. Furthermore, transcriptome diagnostics with digital mRNA profiling may prove applicable to other disorders as well, which will help further promote distribution of this novel diagnostic technique.

CONCLUSION

In summary, using a high-throughput digital mRNA expression platform, we show that a statistical model built on changes in expression of 10 EoE-specific genes detects EoE patients from non-EoE patients with high diagnostic sensitivity and specificity. This novel strategy could help stratify patients at time of their first diagnostic endoscopy and lead to rapid installment of appropriate EoE-specific therapy.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

This work was supported by the Gerber Foundation (to E.F. and S.N.). Further support came from National Institutes of Health (NIH) grants R01AI075037 (to E.F.), K24DK82792-1 (to S.N.) and the Research Council, Boston Children's Hospital (pilot study, to S.N. and E.F.). This study was further supported by an unrestricted gift from Mead Johnson Nutrition Co to the Division of Gastroenterology and Nutrition (Wayne I. Lencer), E.F. and W.L. (unrestricted postdoctoral research fellowship). W.L. is further supported by grants from the Ter Meulen Fund of the Royal Netherlands Academy of Arts and Sciences and the Banning de Jong Fund in The Netherlands. E.D. was funded by the APART Program of the Austrian Academy of Sciences. This work was also supported by the Harvard Digestive Diseases Center Grant P30 DK034854.

REFERENCES

1. 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. quiz 21–2. [PubMed: 21477849]
2. Brown-Whitehorn T, Liacouras CA. Eosinophilic esophagitis. *Curr. Opin. Pediatr.* 2007; 19:575–580. [PubMed: 17885478]
3. Papadopoulou A, Koletzko S, Heuschkel R, et al. Management guidelines of eosinophilic esophagitis in childhood. *Journal of Pediatric Gastroenterology and Nutrition.* 2014; 58:107–118. [PubMed: 24378521]
4. Dellon ES, Speck O, Woodward K, et al. Clinical and endoscopic characteristics do not reliably differentiate PPI-responsive esophageal eosinophilia and eosinophilic esophagitis in patients undergoing upper endoscopy: a prospective cohort study. *Am J Gastroenterology.* 2013; 108:1854–1860.

5. Schoepfer AM, Simon D, STRAUMANN A. Eosinophilic oesophagitis: latest intelligence. *Clin. Exp. Allergy*. 2011; 41:630–639. [PubMed: 21429051]
6. Hruz P, Straumann A, Bussmann C, et al. Escalating incidence of eosinophilic esophagitis: a 20-year prospective, population-based study in Olten County, Switzerland. *J. Allergy Clin. Immunol.* 2011; 128:1349–1350. e5. [PubMed: 22019091]
7. Katzka DA. Unmet diagnostic needs in eosinophilic esophagitis. *Dig Dis*. 2014; 32:139–142. [PubMed: 24603398]
8. Blanchard C, Wang N, Stringer KF, et al. Eotaxin-3 and a uniquely conserved gene-expression profile in eosinophilic esophagitis. *J. Clin. Invest.* 2006; 116:536–547. [PubMed: 16453027]
9. Blanchard C, Mingler MK, McBride M, et al. Periostin facilitates eosinophil tissue infiltration in allergic lung and esophageal responses. *Mucosal Immunol.* 2008; 1:289–296. [PubMed: 19079190]
10. Abonia JP, Blanchard C, Butz BB, et al. Involvement of mast cells in eosinophilic esophagitis. *J. Allergy Clin. Immunol.* 2010; 126:140–149. [PubMed: 20538331]
11. Lexmond WS, Pardo M, Rooney K, et al. Elevated levels of leukotriene C4 synthase mRNA distinguish a subpopulation of eosinophilic oesophagitis patients. *Clin. Exp. Allergy*. 2013; 43:902–913. [PubMed: 23889244]
12. Lexmond WS, Neves JF, Nurko S, et al. Involvement of the iNKT Cell Pathway Is Associated With Early-Onset Eosinophilic Esophagitis and Response to Allergen Avoidance Therapy. *Am J Gastroenterology*. 2014
13. Wen T, Stucke EM, Grotjan TM, et al. Molecular diagnosis of eosinophilic esophagitis by gene expression profiling. *Gastroenterology*. 2013; 145:1289–1299. [PubMed: 23978633]
14. Geiss GK, Bumgarner RE, Birditt B, et al. Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat. Biotechnol.* 2008; 26:317–325. [PubMed: 18278033]
15. Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. 2003:40–44.
16. Dehlink E, Platzer B, Baker AH, et al. A soluble form of the high affinity IgE receptor, Fc-epsilon-RI, circulates in human serum. *PLoS ONE*. 2011; 6:e19098. [PubMed: 21544204]
17. Rothenberg ME. Biology and treatment of eosinophilic esophagitis. *Gastroenterology*. 2009; 137:1238–1249. [PubMed: 19596009]
18. Blanchard C, Stucke EM, Rodriguez-Jimenez B, et al. A striking local esophageal cytokine expression profile in eosinophilic esophagitis. *J. Allergy Clin. Immunol.* 2011; 127:208–17. 217, e1–7. [PubMed: 21211656]
19. Nielsen JA, Lager DJ, Lewin M, et al. The optimal number of biopsy fragments to establish a morphologic diagnosis of eosinophilic esophagitis. *Am J Gastroenterology*. 2014; 109:515–520.
20. Shah A, Kagalwalla AF, Gonsalves N, et al. Histopathologic variability in children with eosinophilic esophagitis. *Am J Gastroenterology*. 2009; 104:716–721.
21. Bannert C, Bidmon-Fliegenschnee B, Stary G, et al. Fc-Epsilon-RI, the High Affinity IgE-Receptor, Is Robustly Expressed in the Upper Gastrointestinal Tract and Modulated by Mucosal Inflammation. *PLoS ONE*. 2012; 7:e42066. [PubMed: 22848703]
22. Chang S-Y, Cha H-R, Chang J-H, et al. Lack of retinoic acid leads to increased langerin-expressing dendritic cells in gut-associated lymphoid tissues. *Gastroenterology*. 2010; 138:1468–1478. e1–6. [PubMed: 19914251]

Figure 1.
Flow diagram of patient inclusion and composition of training and predictive sets.

Figure 2. Heatmap display of differentially regulated genes in either distal (A) or proximal (B) oesophageal biopsies by clinicopathological diagnosis. Bubble plots display levels of statistical significance and the direction of regulation (color: blue for down and red for up-regulation), and the statistical significance of the effect (size: larger bubbles signify smaller p-values).

Figure 3.

(A) EoE probability in EoE training set patients correlates positively with the severity of tissue eosinophil infiltration. (B) Graphical representation of cumulative disease-specific probability scores in proximal biopsies of training set patients, with summary statistics per diagnosis depicted under (C).

Figure 4.

Predicted probability of underlying diagnosis from predictive set patients in distal or proximal biopsies, with corresponding summary statistics of diagnostic probability per clinicopathological category. Graphs in the combined panel display the highest EoE probability in either proximal or distal biopsy.

Table 1
Clinical details and characteristics of patients used in training and predictive sets.

PARAMETER	TRAINING SET				PREDICTIVE SET				* p-value
	NH	EOE	RE	NH	EOE	RE	RE		
n	30	35	30	16	18	13	13		
Age (years, median, range)	10.3 (1.25-18.0)	10.6 (2-18.5)	10.0 (1.4-18.6)	9.5 (1.6-17.2)	6.8 (2.8-16)	13.5 (4.2-18.5)	13.5 (4.2-18.5)	0.55	
Male gender	10/30 (33%)	20/35 (57%)	20/30 (67%)	5/16 (31%)	11/18 (61%)	9/13 (69%)	9/13 (69%)	0.003	
Symptomatology in last year									
Dysphagia	10/30 (33%)	16/35 (46%)	6/30 (20%)	4/16 (25%)	9/18 (50%)	4/13 (31%)	4/13 (31%)	0.04	
Food impaction	0/30 (0%)	4/35 (11%)	0/30 (0%)	1/16 (6%)	2/18 (11%)	1/13 (8%)	1/13 (8%)	0.12	
Chest pain	0/30 (0%)	4/35 (11%)	3/30 (10%)	0/16 (0%)	1/18 (6%)	3/13 (23%)	3/13 (23%)	0.02	
Epigastric pain	13/30 (43%)	5/35 (14%)	11/30 (37%)	6/16 (38%)	5/18 (28%)	3/13 (23%)	3/13 (23%)	0.04	
Heartburn	2/30 (7%)	7/35 (20%)	6/30 (20%)	2/16 (13%)	2/18 (11%)	4/13 (31%)	4/13 (31%)	0.17	
Reflux	2/30 (7%)	14/35 (40%)	16/30 (53%)	3/16 (19%)	2/18 (11%)	7/13 (54%)	7/13 (54%)	<0.001	
Feeding difficulties	4/30 (13%)	5/35 (14%)	1/30 (3%)	3/16 (19%)	3/18 (17%)	1/13 (8%)	1/13 (8%)	0.22	
Vomiting	4/30 (13%)	9/35 (26%)	12/30 (40%)	4/16 (25%)	7/18 (39%)	3/13 (23%)	3/13 (23%)	0.15	
Endoscopy									
Pallor	0/30 (0%)	12/35 (34%)	4/30 (13%)	1/16 (6%)	0/18 (0%)	0/13 (0%)	0/13 (0%)	0.005	
Edema	0/30 (0%)	3/35 (9%)	1/30 (3%)	1/16 (6%)	1/18 (6%)	0/13 (0%)	0/13 (0%)	0.38	
Loss of vascularity	0/30 (0%)	13/35 (37%)	4/30 (13%)	2/16 (13%)	5/18 (28%)	1/13 (8%)	1/13 (8%)	<0.001	
Furrowing	3/30 (10%)	27/35 (77%)	8/30 (27%)	1/16 (6%)	15/18 (83%)	1/13 (8%)	1/13 (8%)	<0.001	
Exudate	1/30 (3%)	10/35 (29%)	0/30 (0%)	0/16 (0%)	8/18 (44%)	0/13 (0%)	0/13 (0%)	<0.001	
Allergic diatheses									
IgE (median, range)	21 (2-308)	195 (4-1768)	38 (0-1658)	6 (2-99)	136 (28-714)	69 (7-957)	69 (7-957)	<0.001	
Eczema	6/29 (21%)	16/35 (46%)	13/29 (44%)	3/16 (19%)	5/18 (28%)	5/13 (38%)	5/13 (38%)	0.046	
Asthma	6/30 (20%)	9/35 (26%)	11/29 (38%)	5/16 (31%)	4/18 (22%)	4/13 (31%)	4/13 (31%)	0.39	
Seasonal allergies	7/30 (23%)	20/35 (57%)	15/29 (52%)	5/16 (31%)	13/18 (72%)	4/13 (31%)	4/13 (31%)	0.002	
Food allergy	4/30 (13%)	18/35 (51%)	6/29 (21%)	3/16 (19%)	11/16 (69%)	4/11 (36%)	4/11 (36%)	<0.001	
Positive RAST or skin prick test against food antigens	0/6 (0%)	22/30 (73%)	9/20 (45%)	1/4 (25%)	16/17 (94%)	1/5 (20%)	1/5 (20%)	<0.001	
Tissue eosinophilia (<i>peak value</i>)									
Proximal (median, range)	0 (0-0)	30 (5-150)	4 (2-10)	0 (0-0)	40 (2-150)	2 (1-5)	2 (1-5)	<0.001	

PARAMETER	TRAINING SET			PREDICTIVE SET			p-value *
	NH	EoE	RE	NH	EoE	RE	
Distal (median, range)	0 (0-0)	40 (20-150)	3 (1-20)	0 (0-0)	50 (2-150)	1.5 (1-2)	<0.001

* comparing patients from all three diagnostic groups in combined training and predictive sets, calculated using Fisher's exact test or ANOVA.

Table 2

Significantly regulated genes between all three conditions.

Comparison:	EoE / NH		RE / NH		EoE / RE	
	t-value	p-value	t-value	p-value	t-value	p-value
Proximal						
CD4	-6.61	7.39×10^{-9}	-3.65	0.001	-2.43	0.018
CCL22	-4.72	1.22×10^{-5}	-2.84	0.006	-2.13	0.037
FCER1A	-7.70	7.96×10^{-11}	-3.10	0.003	-3.59	0.001
CCL26	10.56	5.96×10^{-16}	2.13	0.037	7.38	3.00×10^{-10}
FCER1B	7.42	2.51×10^{-10}	2.43	0.018	4.73	1.17×10^{-5}
IL13	8.05	1.85×10^{-11}	2.71	0.009	4.45	3.25×10^{-5}
CD40LG	-5.29	1.42×10^{-6}	-2.54	0.013	-2.45	0.017
Distal						
CD4	-7.90	3.92×10^{-11}	-3.53	0.001	-2.84	0.006
CCL17	-7.13	9.23×10^{-10}	-2.63	0.011	-3.77	3.45×10^{-4}
CD207	-6.80	3.68×10^{-9}	-2.01	0.049	-3.63	5.41×10^{-4}
FCER1B	11.48	2.14×10^{-17}	2.66	0.010	6.55	9.37×10^{-9}
FCER1A	-9.46	6.60×10^{-14}	-3.57	0.001	-3.39	0.001
CPA3	11.43	2.60×10^{-17}	2.29	0.026	7.38	3.17×10^{-9}
CCL26	14.70	1.60×10^{-22}	2.87	0.006	8.66	1.51×10^{-12}
CCR3	7.54	1.74×10^{-10}	2.62	0.011	5.15	2.47×10^{-6}
IL13	10.25	2.69×10^{-15}	4.28	6.80×10^{-5}	4.59	2.02×10^{-5}
LGALS3	-5.40	9.73×10^{-7}	-3.48	0.001	-2.35	0.022