INTRODUCTION

Environmental enteric dysfunction (EED) is a chronic, inflammatory, subclinical condition prevalent in low- and middle-income countries (LMICs) that is hypothesized to be a consequence of persistent environmental exposure to enteropathogens. Environmental enteric dysfunction is characterized by morphologic changes including intestinal mucosal inflammation and villus blunting, as well as functional damage, including altered gut permeability, and reduced intestinal absorption. Diminished intestinal absorption leads to growth faltering, neurodevelopmental disability, and poor response to oral vaccines and nutritional therapy. The persistence of stark nutritional disparities in LMIC settings versus their high-income counterparts directly impacts childhood growth rates, which may be secondary to defects in intestinal absorptive capacity due to the abovementioned environmental factors. Although health interventions are being deployed worldwide to tackle malnutrition, there is great disparity between the responses to these interventions in high-income and LMIC settings because of the prevalence of EED in resource-poor areas. Moreover, current efforts to establish diagnostic biomarkers for EED, as well as effective interventions to treat or prevent this enteropathy, are in their infancy.

As with other enteropathies, the current gold standard for the diagnosis of EED, as well as the evaluation of EED pathophysiology, is the histopathologic examination of duodenal biopsies. Upon evaluation of children with EED across multiple geographic sites, the EED histologic scoring index was developed to quantify the extent of morphologic changes associated with EED and is a step toward a better characterization of EED as an inflammatory enteropathy on a spectrum of severity. Environmental enteric dysfunction shares histopathologic features with celiac disease, such as altered villus-to-crypt ratios, increased intraepithelial T cells, lamina propria T-cell infiltrates, and B-cell aggregates. Kelly et al. reported a correlation between villus height and markers of intestinal permeability among an adult population in Zambia. Additionally, our group previously reported the presence of goblet cells and intraepithelial lymphocytes (IELs) as key distinguishing features of EED. Numerous studies show that the EED transcriptome exhibits the suppression of antioxidant and detoxification genes and the induction of antimicrobial response genes similar to celiac disease. In addition to duodenal inflammation, previous studies have implied that EED may also be associated with colonic mucosal and histomorphologic alterations. However, there is a dearth of data available to understand whether there is an environmental colonopathy partnering with EED.

In this study, we present the evaluation of small bowel enteropathy in children with EED using histomorphometry and machine learning–based biopsy image analysis methods and compare our results with those of patients with celiac disease and controls. We also characterize differences in celiac disease across geography. Furthermore, we had the...
unique opportunity to explore the prevalence of colonopathy in this population by evaluating alterations in the colonic mucosa in children with EED.

MATERIALS AND METHODS

Study design and selection of cohorts. This study used retrospective, archival samples from two different sites (Pakistan and the United States) and the prospective collection of samples from Pakistani children with suspected EED (Supplemental Figure 1). From the Pakistani site, clinical metadata and archival duodenal histopathology slides (Department of Pathology, Aga Khan University [AKU]) were collected from children under the age of 5 years who presented with abdominal pain, poor growth, diarrhea, and/or blood in stools and subsequently underwent esophagogastroduodenoscopy or colonoscopy with biopsy between 2007 and 2017. These slides were classified using prior clinical histopathologic evaluation into three categories: 1) celiac disease, 2) chronic duodenitis, and 3) histopathologically undiagnosed (based on the absence of any gastrointestinal diseases on tissue biopsy) after assessment by two institutional clinical pathologists.

Slides from the U.S. cohort were obtained from the clinical archives of the University of Virginia (UVA), in Charlottesville, VA, between 1992 and 2017 from children under 18 years of age who had duodenal and rectal biopsies available. Clinical diagnosis of celiac disease was considered a disease comparison group, whereas those with no histopathologic abnormalities on biopsy were selected as controls. Similarly, for rectal biopsy, clinical confirmation of ulcerative proctitis/crohnitis or cryptitis was chosen as the disease comparison group, and patients with no histopathologic abnormalities were selected as controls. Archival slides from patients with other gastrointestinal diseases (eosinophilic esophagitis, gastritis, inflammatory bowel disease) or slides with artifacts obscuring duodenal tissue were excluded from this analysis.

Informed consent was obtained from the parents of the children for whom archival tissue blocks were used for analysis. This study was approved by the AKU Ethical Review Committee (ERC#3836-Ped-ERC-15) and the UVA Institutional Review Board (HSR IRB# 20107). Data used for the prestrained image analysis platform have been described elsewhere.12,16

To explore causes of malnutrition, the samples from patients with EED in Pakistan included prospectively obtained duodenal biopsies from children refractory to nutritional intervention who were enrolled in the Study of Environmental Enteropathy and Malnutrition.17 The enrollment procedures and patient selection for endoscopic evaluation of EED have been published elsewhere.18 A subset of patients diagnosed with EED also underwent flexible sigmoidoscopy with rectal biopsy if they had persistent diarrhea and/or blood in the stool.

Hematoxylin-eosin (H&E)–stained biopsy glass slides were digitized at ×40 magnification using the Olympus VS120 scanner (Olympus Corporation Inc., Center Valley, PA) at AKU and the Leica SCN400 brightfield scanner (Leica Microsystems CMS GmbH, Mannheim, Germany) at UVA.

Histomorphometric analysis of digitized biopsies. All histomorphometric analyses were performed using the Olympus cellSens or Leica Aperio ImageScope software for biopsies digitzed on Olympus VS120 or Leica SCN400 brightfield scanners, respectively.

Duodenal histomorphometry. Villus and crypt measurements of the duodenal biopsies were conducted on well-oriented villus-crypt units that were representative of each biopsy image and for which the assessment of the depth of the mucosa was possible. Distances from the villus tip to the crypt-villus junction and then the crypt base were measured in micrometers (μm), as shown in Figure 1A. A horizontal line was placed at the crypt-villus junction (identified via shouldering; a villus “shoulder” was defined as the region where a scaffold was present, which marked the beginning of a crypt) as a point of reference to aid in accurately measuring villus length and crypt depth (Figure 1A). In cases where the villus was curved or bent, the measurements took the curvature into account as supported by previously published literature19 (Figure 1B). For crypt depth measurement in a subset where the muscularis mucosa or the beginning of the submucosa was visible without the complete crypts being evident, we extended the measurement from the crypt-villus junction to the muscularis mucosa or the start of the submucosa (Figure 1C). This was done to reduce the risk of underestimating crypts in the absence of suitable orientation, as reported elsewhere.20,21 Those sections without a visible submucosa for which an assessment of the depth of mucosa could not be evaluated were excluded.

To quantify cells in the surface epithelium, IELs, goblet cells, and neutrophils were measured per 100 epithelial cells; surface epithelial cells within the villi only (and not the crypts) were selected. Areas of damaged surface epithelium with breaks and nuclear crowding were avoided unless an increase in nuclear counts was present throughout the biopsy image. Those cells situated at the basal membrane of the surface epithelium were disregarded as it was difficult to distinguish whether they were truly within the surface epithelium or the lamina propria (Figure 1D).

For the quantification of cells in the lamina propria, mononuclear inflammatory cells (MICs), neutrophils, and eosinophils were quantified per 2,500 μm² of lamina propria in each duodenal biopsy.22 Overly or deficiently dense cellular areas were disregarded to avoid over-representation of cell counts when counting MICs, neutrophils, and eosinophils (Figure 1E).

Rectal histomorphometry. For the quantification of cells within crypt cross sections, a representative 100,000-μm² region was selected, and the number of crypts within this region was noted.23 If greater than 50% of the crypt was within the specified region, then the entire crypt was counted. Overly or defiently dense crypt epithelium was disregarded to avoid over-representing cell counts. The number of lymphocytes, eosinophils, and neutrophils within the selected crypts was counted. Inflammatory cells within the rectal biopsy lamina propria were quantified using methods analogous to those used for duodenal biopsies.

EED and celiac disease severity scoring. Gastrointestinal pathologists used the histologic scoring index developed by the Environmental Enteric Dysfunction Biopsy Initiative (EEDBI) consortium to score the prospective EED (Z. A. and R. I.) and AKU archival biopsies (C. M.).19 This scoring system comprises 11 variables with a total of 37 points (see Supplemental Appendix 1 for a detailed EED histologic scoring index). Celiac disease severity was assessed using the Marsh–Oberhü6er classification (modified Marsh score), which involves evaluating duodenal villus architecture and intraepithelial lymphocytosis.24 In the case where more than one biopsy

DUODENAL AND RECTAL MUCOSAL TISSUE OF CHILDREN WITH EED 673
fragment was present per image, each fragment was separately assessed using the Marsh–Oberhüner classification (see Supplemental Appendix 2 for a detailed illustration of the modified Marsh score classification).

**Data analysis of morphometric measurements.** Regarding data analysis, the morphometric measurements are presented in dot plots with the median (interquartile range [IQR]) marked. The Kruskal-Wallis test was performed to evaluate statistical differences between the groups using GraphPad Prism version 9.04 (GraphPad Software, La Jolla, CA; www.graphpad.com). Multiple comparisons were corrected using Dunn’s multiple comparisons test. Spearman correlation was applied to assess the association between histopathologic features and morphometric measurements. A Mann-Whitney test was used to compare the celiac (Pakistan) and celiac (U.S.) cohorts. A P value < 0.05 was considered statistically significant.

**Machine learning–based biopsy image analysis model.** Methods for the process of data preparation and training the machine learning image analysis model are summarized in Figure 2. High-resolution whole slide images (WSIs) that measure more than 20,000 × 20,000 pixels were divided into patches measuring 512 × 512 pixels for analysis within the computational constraints of the model. Each biopsy WSI generated an average of 290 image patches. Details of biopsy image patch creation are explained in Supplemental Appendix 3.

To eliminate bias due to color differences, stain color normalization used the structure-preserving method described by Vahadane et al.25 (details in Supplemental Appendix 4). Three
independent pathologists (Z. A., R. I., and L. C.) completed a blind review of the color-normalized biopsy images from different sites to assess the structure-preserving ability of the method, details of which have been previously published.

The patches were classified using a pretrained ResNet50 machine learning architecture, as shown in Figure 2. This machine learning–based classification model was used to evaluate both duodenal and rectal biopsy tissues to help identify patterns of distinguishing morphologic features characterizing duodenal and rectal inflammation. Details of the model architecture are described in Supplemental Appendix 5. The models were trained on biopsies from patients with EED, celiac disease, and control cases and were then tested on a separate cohort of celiac disease and chronic duodenitis for predictive accuracy. For the classification of rectal biopsies, models were trained using rectal biopsies from children with EED, diseased, and control rectal tissue to predict a separate set of EED, diseased, and control patients.

Gradient-weighted Class Activation Mappings (Grad-CAMs) were used to visualize the regions of interest used by the model for decision-making. Grad-CAMs were reviewed by a gastrointestinal pathologist (Shyam Raghavan) and a pediatric gastroenterologist (Sana Syed) to enable corroboration of model results with the EED and Marsh histologic scoring indices to assess whether features highlighted by the model as being of high significance in machine learning–based decision-making could be biologically explained (see Supplemental Appendix 6 for details).

**RESULTS**

The background characteristics of the study cohorts are summarized in Tables 1 and 2. Further details regarding retrieval of archival biopsies are mentioned in Supplemental Appendix 7.

**Morphometric features of EED and celiac disease.** The villus blunting on the duodenal sections was most prominent in celiac disease, with a median (IQR) length in controls of 324 (218, 454) μm. Pakistani cases of celiac disease demonstrated even shorter lengths than U.S. celiac cases, with
Table 1

## Distribution of cases with duodenal tissue hematoxylin–eosin–stained histology slides

<table>
<thead>
<tr>
<th>Variables</th>
<th>EED</th>
<th>Celiac</th>
<th>Chronic duodenitis</th>
<th>Undiagnosed</th>
<th>Celiac</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>63</td>
<td>18</td>
<td>17</td>
<td>7</td>
<td>63</td>
<td>61</td>
</tr>
<tr>
<td>Median (IQR) age, months</td>
<td>(15.3, 22.2)</td>
<td>(25.5, 48)</td>
<td>(42, 60)</td>
<td>(36, 60)</td>
<td>(92.5, 175.5)</td>
<td>(17, 41)</td>
</tr>
<tr>
<td>Male gender, %</td>
<td>69.8</td>
<td>22.2</td>
<td>52.9</td>
<td>42.9</td>
<td>66.0</td>
<td>54.1</td>
</tr>
<tr>
<td>LAZ/HAZ, median (IQR)</td>
<td>−2.04</td>
<td>−1.82</td>
<td>−1.45</td>
<td>−0.68</td>
<td>−0.34</td>
<td>−0.22</td>
</tr>
</tbody>
</table>

EED = environmental enteric dysfunction; IQR = interquartile range; LAZ/HAZ = length/height-for-age Z score. Missing length/height for duodenal biopsies: Pakistani celiac (N = 4), chronic duodenitis (N = 2), undiagnosed (N = 1).

Table 2

## Distribution of cases with rectal tissue hematoxylin–eosin–stained histology slides

<table>
<thead>
<tr>
<th>Variables</th>
<th>EED</th>
<th>Controls</th>
<th>Ulcerative proctitis</th>
<th>Cryptitis</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>14</td>
<td>9</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Median (IQR) age, months</td>
<td>21.5 (20.5, 22.7)</td>
<td>165 (50, 196.8)</td>
<td>141 (111.8, 193.8)</td>
<td>72 (67, 88.5)</td>
</tr>
<tr>
<td>Male gender, %</td>
<td>71.4</td>
<td>33.3</td>
<td>25.0</td>
<td>25.0</td>
</tr>
<tr>
<td>LAZ/HAZ, median (IQR)</td>
<td>−1.89 (−2.56, −0.98)</td>
<td>0.02 (−0.57, 0.54)</td>
<td>−0.02 (−0.12, 0.19)</td>
<td>0.49 (0.14, 1.02)</td>
</tr>
</tbody>
</table>

EED = environmental enteric dysfunction; IQR = interquartile range; LAZ/HAZ = length/height-for-age Z score. Missing length/height for rectal biopsies: cryptitis (N = 1), controls (N = 1).
Paneth cells, along with total EED scores. The presence of neutrophils in the epithelium and lamina propria was associated with some histologic features, yet it did not reach statistical significance. Regarding rectal morphometry, the presence of neutrophils within the epithelial lining of the crypt reported the most significant associations with histology features of the duodenum (Figure 6B).

Machine learning models for biopsies and Grad-CAMs. Regarding duodenal tissue, the ResNet50 machine learning model exhibited an accuracy of 49%, 21%, and 30% for predicting celiac disease biopsies as diseased, EED, and normal, respectively (Figure 7A). The model also exhibited 64%, 6%, and 30% accuracy for predicting chronic duodenitis biopsies as celiac disease, EED, and normal. Finally, it
exhibited 12%, 2%, and 86% accuracy for predicting chronic duodenitis biopsies as celiac disease, EED and normal. Grad-CAMs are shown in Figure 8A. The model was performed on rectal tissue. The model predicted diseased rectal biopsies with an accuracy of 98% and EED with an accuracy of 97% (Figure 7B). The normal biopsies overlapped with diseased (66%) with a 33% prediction accuracy for normal. Grad-CAMs for rectal classification are shown in Figure 8B.

DISCUSSION

This study aimed to characterize duodenal and rectal tissues among children with EED using histomorphometry. In addition to duodenal features, rectal inflammation previously reported in the last century has been observed in a subset of children who underwent colonoscopy in addition to upper endoscopy, highlighting the potential involvement of the large intestine in EED. Compared with celiac disease patients, duodenal tissue in patients with EED showed milder histopathologic changes. Archival samples with a lack of findings on histopathology were evaluated to serve as local controls. However, pathologic findings were seen on morphometry and EED scoring and were excluded from the analysis.

Widespread variations in morphometric measurements were observed across EED samples, suggesting a spectrum of severity that supports previous histopathology findings. Goblet cell depletion and IELs were marked in EED samples compared with controls, whereas severely altered villus architecture was a feature of celiac disease. Inflammatory cells in the lamina propria were commonly observed in both duodenitis and EED. Goblet cell depletion and IELs were marked in EED samples compared with controls, whereas severely altered villus architecture was a feature of celiac disease. Inflammatory cells in the lamina propria were commonly observed in both duodenitis and EED. Goblet cell depletion and IELs were marked in EED samples compared with controls, whereas severely altered villus architecture was a feature of celiac disease. Inflammatory cells in the lamina propria were commonly observed in both duodenitis and EED. Goblet cell depletion and IELs were marked in EED samples compared with controls, whereas severely altered villus architecture was a feature of celiac disease. Inflammatory cells in the lamina propria were commonly observed in both duodenitis and EED.
**Figure 5.** Rectal histomorphometry. (A and B) Rectal morphometry shown on photomicrographs taken at ×40 magnification. (C–H) Quantification of mononuclear cells, eosinophils, and neutrophils per 2,500 μm² of lamina propria (C–E) and crypt cross sections within a region of 100,000 μm² (F–H). The sample sizes of each group were EED = 14, diseased = 12, and controls = 9. The Kruskal-Wallis test was applied to assess the P value between various groups. The error bars represent median with interquartile ranges. *P < 0.05, **P < 0.005, ***P < 0.0005, ****P < 0.0001. EED = environmental enteric dysfunction. ROI = region of interest.
Previously published studies indicate an underestimation of the prevalence of even well-established enteropathies, such as celiac disease, in South Asia as a result of lack of resources for diagnosis, further supporting our findings.\textsuperscript{29,30}

Our rectal histomorphometry insights point toward more significant colonic involvement in EED than was previously understood. Mathan et al.\textsuperscript{15} demonstrated increased inflammatory cells in the rectal mucosa of healthy Indian volunteers, resembling the nonspecific inflammatory response in the small intestine associated with EED (known then as tropical enteropathy) and hypothesized the existence of tropical colopathy.\textsuperscript{15} In our study, rectal biopsies were collected only from children with fecal bleeding and demonstrated a significantly higher number of inflammatory cells in the mucosa. This suggests that EED may not be limited to the small intestine; thus, improving colonic health may be added to the...
strategies for treatment. In children with short bowel syndrome and subsequent malabsorption, the large intestine acts as a salvage organ for increasing caloric intake via the absorption of carbohydrates. Therefore, in small intestinal malabsorption syndromes such as EED, compromised ability of absorption through the large intestine may hinder the body’s efforts to compensate for absorption through colonic tissue. These findings must be further explored in rectal tissue collected from EED cases, with careful consideration to study colopathy in patients both with and without rectal bleeding to rule out other colonopathies.

Rectal tissue Grad-CAM saliency map analyses highlighted areas of goblet cells, lymphocytes in the lamina propria, and surface epithelial cells. These findings are similar to lymphocytic colitis, distinguished from celiac sprue by increased lamina propria cellularity, surface epithelial abnormalities, and fewer IELs. Even though our machine learning–based analysis model for rectal biopsy analysis demonstrated high classification accuracies for predicting both diseased (98%) and EED (98%) colon, our small sample size precludes us from drawing meaningful interpretations at this time. However, similar to our model for duodenal analysis, more data will...
strengthen our analyses and enable us to predict disease characteristics that will aid in elucidating tissue features of EED-related colonopathy.

One of the major strengths of our study is the expansion of our understanding of colonopathy among children with EED. We compared duodenal tissue to enteropathies other than celiac disease. Further, we demonstrated the ability of machine learning–based image analysis to identify key distinguishing features in both diseased and normal tissues. Grad-CAMs may increase physician confidence in machine learning–based decision-making and pave the way for discerning disease-specific histopathologic markers.

However, our study does have several limitations. First, the sample size for the rectal analysis was limited, and we acknowledge that more data will be required before meaningful conclusions can be drawn. Inflammation in the lamina propria was quantified by mononuclear cells that include lymphocytes, plasma cells, dendritic cells, and macrophages; this is a limitation of H&E–stained biopsies. Immunohistochemistry can be used to better characterize cells in the lamina propria. Third, histomorphometry can be subject to observer bias, and interobserver analyses were beyond our scope. Orientation of the tissue is also an important feature while performing morphometry, and as a result, we were able to evaluate villus architecture in only a subset of duodenal biopsies. Lastly, demographic data for archival samples were limited to gender and age (with a diverse range of ages up to 18 years), as gut tissue does not exhibit age-related differences.

In conclusion, EED comprises a spectrum of inflammation in the duodenum that has been observed in Pakistani children in comparison to their U.S. counterparts and needs further exploration in other settings. In addition, the involvement of the rectal mucosa merits further investigation, especially in the context of loss of compensatory absorption from the colon.

Received January 22, 2022. Accepted for publication December 21, 2022.

Published online March 13, 2023.

Note: Supplemental figures and appendices appear at www.ajtmh.org.

Acknowledgments: We thank Pat Pramoonjago, Tad Kolterman, and Rebecca Blackwell from the UVA Biorepository and Tissue Research Facility and all field staff from the Study of Environmental Enteroopathy and Malnutrition at Aga Khan University, Karachi for their contributions to this work.

Financial Support: Research reported in this manuscript was supported by the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health under award numbers K23DK117061-01A1 (to S. S.) and D43TW007585 (to S. A. A.) and the Bill & Melinda Gates Foundation under award numbers OPP1144149 (to S. M.) and OPP1138727 (to S. A. A.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.

Authors’ addresses: Marium Khan, Lubaina Ehsan, Fatima Zukarnain, Sanjana Srivastava, Philip Fernandes, Muhammad Raghib, and Sean Moore, Department of Pediatrics, School of Medicine, University of Virginia, Charlottesville, VA, E-mails: mk2he@hscmail.mcc.virginia.edu, lubaina.ehsan@gmail.com, fatimav@virginia.edu, eqpp6pg@uvaheroth.org, pf2uxc@hscmail.mcc.virginia.edu, farazraghib7@gmail.com, and srmi5@hscmail.mcc.virginia.edu. Zehra Jamil, Sheraz Ahmed, and Fayaz Umrani, Department of Biological and Biomedical Sciences, Aga Khan University, Karachi, Pakistan, E-mails: zehra.jamil@aku.edu, sherazahmed@aku.edu, and fayaz.umrani@aku.edu. Saman Siddiqui, Zia Mujahid, Najeeha Iqbal, and Syed Asad Ali, Department of Pediatrics and Child Health, Aga Khan University, Karachi, Pakistan, E-mails: samansidd@hotmail.com, zia.mujahid@aku.edu, najeha.iqbal@aku.edu, and asad.ali@aku.edu. Saurav Sengupta, School of Data Science, University of Virginia, Charlottesville, VA, E-mail: ss44y@virginia.edu. Zubair Ahmed and Romana Idrees, Department of Pathology and Laboratory Medicine, Aga Khan University, Karachi, Pakistan, E-mails: zubair.ahmed@aku.edu and romana.idrees@aku.edu. Christopher Moskaluk and Shyam Raghavan, Department of Pathology, University of Virginia, Charlottesville, VA, E-mails: cam5p@virginia.edu and srw7q@hscmail.mcc.virginia.edu. Lin Cheng, Department of Pathology, Rush University, Chicago, IL, E-mail: lin.cheng@rush.edu. Junaid Iqbal, Department of Biological and Biomedical Sciences, Aga Khan University, Karachi, Pakistan, and Department of Pediatrics and Child Health, Aga Khan University, Karachi, Pakistan, E-mails: j.iqbal@aku.edu, sana.syed@aku.edu, and Department of Pediatrics, School of Medicine, University of Virginia, Charlottesville, VA, Department of Pediatrics and Child Health, Aga Khan University, Karachi, Pakistan, School of Data Science, University of Virginia, Charlottesville, VA, E-mail: sana.syed@virginia.edu.

This is an open-access article distributed under the terms of the Creative Commons Attribution (CC-BY) License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

REFERENCES