Table 1: Menopausal Factors, Menopausal Hormone Therapy (MHT) Use, and Risk of Microscopic Colitis amongst post-menopausal women in the Nurses' Health Study (NHS) and NHSII

	No. of cases	No. of person-years	Age-adjusted hazard ratio [95% CI]	Multivariable-adjusted hazard ratio [95% CI][1]
MHT Status				
Never Use	22	1,145,132	(reference)	(reference)
Past Use	64	852,980	2.48 (1.51 - 4.08)	2.33 (1.41 - 3.84)
Current Use	46	847,793	2.95 (1.75 - 4.98)	2.79 (1.63 - 4.77)
P-trend			*<0.0001*	*0.0002*
Years of MHT Use				
Never	22	1,145,132	(reference)	(reference)
≤ 2 years	12	350,848	1.73 (0.85 - 3.50)	1.62 (0.80 - 3.29)
2.1 - 4 years	11	273,085	2.07 (0.99 - 4.31)	1.98 (0.95 - 4.14)
4.1 - 8 years	17	390,189	2.00 (1.05 - 3.80)	1.99 (1.04 - 3.80)
8.1 - 16 years	36	455,071	2.81 (1.63 - 4.85)	2.79 (1.60 - 4.87)
> 16 years	34	231,579	4.68 (2.68 - 8.19)	5.00 (2.71 - 9.22)
P-trend			*<0.0001*	*<0.0001*
Years since MHT discontinuation				
Current Use	46	847,793	(reference)	(reference)
≤ 4 years	26	310,024	0.98 (0.59 - 1.63)	0.98 (0.59 - 1.63)
4.1 - 8 years	27	319,674	0.87 (0.52 - 1.47)	0.85 (0.50 - 1.44)
> 8 years	11	223,282	0.54 (0.27 - 1.10)	0.53 (0.26 - 1.09)
Never Use	22	1,145,132	0.33 (0.20 - 0.56)	0.35 (0.21 - 0.60)
P-trend[2]			*0.073*	*0.055*
Age of menopause				
< 44	17	305,642	1.31 (0.76 - 2.24)	1.20 (0.69 - 2.08)
44 - 46.9	12	238,404	1.11 (0.60 - 2.06)	1.00 (0.54 - 1.86)
47 - 48.9	13	267,098	1.02 (0.56 - 1.85)	0.93 (0.51 - 1.69)
49 - 52.9	71	1,599,839	(reference)	(reference)
≥ 53	19	434,923	0.79 (0.47 - 1.32)	0.81 (0.48 - 1.37)
Type of Menopause				
Natural	73	1,502,825	(reference)	(reference)
Surgical/Radiation	59	1,343,080	0.95 (0.67 - 1.35)	0.82 (0.57 - 1.18)

[1] Adjusted for age (months), cohort (NHS, NHS2), body mass index (<20, 20-24.9, 25-29.9, ≥30), smoking (never, past, current), age of menarche (years), oral contraceptive use (never, ever), age of menopause (years), and menopause type (natural, surgical/radiation). [2] P-trend was estimated by entering Years since MHT discontinuation in the model as a continuous variable excluding the "Never Use" category

195

CONVOLUTIONAL NEURAL NETWORKS IMAGE ANALYSIS OF DUODENAL BIOPSIES ROBUSTLY DISTINGUISHES ENVIRONMENTAL ENTEROPATHY FROM HEALTHY CONTROLS AND IDENTIFIES SECRETORY CELL LINEAGES AS HIGH ACTIVATION LOCATIONS

Sana Syed, Mohammad Al-Boni, Kamran Sadiq, Najeeha T. Iqbal, Christopher A. Moskaluk, Asad Ali, Sean Moore, Don Brown

Background. Environmental Enteropathy (EE) is an acquired small intestinal inflammatory condition underlying high rates of linear growth stunting in children <5y of age in low- and middle-income countries. The histologic appearance of duodenal EE biopsies significantly overlaps with celiac enteropathy. To identify novel features that distinguish these two enteropathies from healthy tissue, we characterized intestinal mucosal alterations in Pakistani infants with EE using Convolutional Neural Networks (CNN). **Methods.** The proposed CNN consisted of four convolutions, one fully connected, and a soft-max layer **(Fig 1)**. We coupled each convolution layer with a deconvolution layer, which allowed us to trace back high activations to the corresponding positions in the input image. These positions highlighted sections of the image that serve as indicators of EE. We obtained 74 biopsies (32 from Pakistani children with EE; 42 from US normals). The biopsies were scanned with relatively high resolutions (e.g., ranging from 2288 x 1356 to 18304 x 14926 pixels). Although most discriminating features can only be observed at high resolution, it is impractical to feed such images as input to the network. Therefore, we segmented each biopsy into a number of 1360 x 1024 images. Furthermore, we adopted standard data augmentation methods to avoid overfitting problems caused by limited data. For each image, we randomly selected ten 1000 x 1000 patches and their horizontal and vertical reflections. This method increased the size of our data by a factor of 30, and helped in learning translation and rotation invariant features. While testing time, we generated 15 patches from each image: central patch, 4 corner patches, and their reflections. Then, we calculated the average probability of EE from these patches. Finally, the likelihood of EE in a biopsy was computed as the average of its segments' estimated probabilities. **Results.** We trained the proposed model on 1219 images from 51 biopsies, and tested it against 659 images from 23 biopsies. The model misclassified 4 EE images from 3 biopsies as control; however, the remaining images from each of the 3 biopsies were identified correctly. The model had 100% biopsy prediction based off average segment estimated probabilities. When analyzing the locations of high activations, the model identified secretory cellular indicators of EE **(Fig 2)**. **Conclusions.** We proposed a deep learning algorithm for: 1) predicting EE from duodenal biopsies; and 2) identifying secretory cells that correlate with the EE. Future directions include extending this model to distinguish between EE and celiac disease. Both diseases have overlapping histologic findings making this a more challenging problem for the machine learning model. **Funding:** UVA Engineering-in-Medicine seed grant (SSyed, DBrown); Bill & Melinda Gates Foundation (AA: OPP1066203).

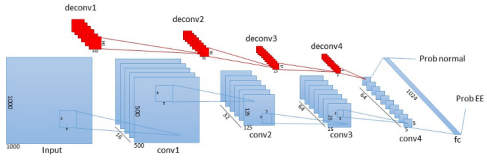

Figure 1. Illustration of our proposed convolutional and deconvolutional neural network. Note, each convolution layer consists of three sub-layers: convolution, ReLU activation, and max pooling layers.

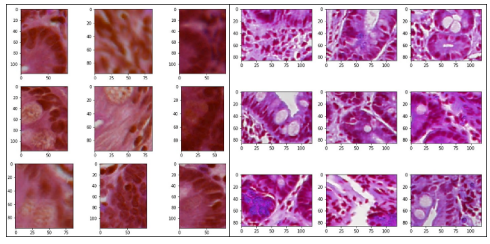

Figure 2. Areas with high activations for scanned biopsies of from normal duodenal tissue (right) and tissue diagnosed with Environmental Enteropathy (EE) (left). We observed Paneth cells and goblet cells in the mucosa. Our classification model identified secretory cells to be of high importance for distinguishing normal biopsies from EE.

196

COMPARISON OF ABILITY OF ELEVATED FECAL PRIMARY BILE ACIDS AND TOTAL FECAL BILE ACIDS TO DETECT ACCELERATED COLONIC TRANSIT AND HIGH STOOL WEIGHT IN SUSPECTED BILE ACID MALABSORPTION

Priya Vijayvargiya, Victor Chedid, Alan R. Zinsmeister, Irene Busciglio, Duane Burton, Leslie J. Donato, Michael Camilleri

Introduction: Bile acid malabsorption (BAM) is currently diagnosed with elevated total fecal bile acids (BAs) where ^{75}SeHCAT is not available. Other proposed tests for BAM are serum C4 and FGF19. Chenodoxycholic acid (CDCA) stimulates colonic transit (CT) and secretion. In healthy subjects, the proportion of fecal primary BAs, CDCA and cholic acid (CA) is <4%. A previous study identified a high concordance correlation coefficient [(CCC) r>0.80] for CA and CDCA between individual random stool samples and total 48h stool collection. **Aim**: To compare the associations of tests for diagnosing BAM with elevated CT and fecal weight in patients with irritable bowel syndrome (IBS). **Methods**: We identified a cohort of patients with IBS-diarrhea (D) and IBS-constipation (C) and healthy volunteers with recorded CT [geometric center at 24h (GC_{24})] and total and primary 48h fecal BAs while on a 4-day, 100g fat diet. We predefined GC_{24} >3.34 [75th %ile in 145 healthy females and 75 healthy males (PMID 24118658)] and fecal weight >400g/48h as objective markers of diarrhea. We performed ROC analysis to assess the accuracy, sensitivity and specificity to predict high GC_{24} (>3.34) or fecal weight (>400g/48h) with tests proposed for BAM diagnosis: total 48h fecal BAs, % primary BAs, fasting serum C4 and FGF19, individually and in combination. **Results**: We included 30 IBS-C, 64 IBS-D, and 30 healthy controls [majority females (59/64 IBS-D, 30/30 IBS-C, 22/30 healthy), mean age respectively 41.9±12y (SD), 44.6±10y, and 39.3±12 y]. One CT data and 7 fecal BAs were not available (2 IBS-C, 3 IBS-D and 2 healthy controls). Figure 1 represents the ROC curves associated with GC_{24}. The AUCs are comparable for total fecal BAs (AUC=0.65), % primary BAs alone (AUC 0.69), and the combination of total fecal BAs and % primary BAs (AUC 0.70). The AUC for serum C4 was 0.60, whereas AUC for FGF19 alone was 0.46 and did not improve with the combination of serum C4 and FGF19. ROC curves associated with fecal weight >400g (Fig. 2) demonstrate that total fecal BAs and combined total fecal BAs with % primary BAs have a similar AUC of 0.86, while % primary BAs alone has AUC of 0.73. The AUCs for serum C4 (0.57) and combined serum FGF19 and C4 (0.57) are lower. With proportion of primary BAs >8%, there is a specificity of 80% and sensitivity of 48% for GC_{24} >3.34, and specificity of 86% and sensitivity of 57% for fecal weight >400g. **Conclusions**: The proportion of primary BAs in stool alone is comparable to total fecal BA excretion in the association with increased CT and fecal weight. C4 is a reasonable serum biomarker of BAM. Given the documented high specificity and prior evidence of high CCC between single and random samples, and total 48h fecal collection, primary BAs >8% in a random stool sample may be an alternative to measure total BAs in 48h collection in the diagnosis of BAM.