EARLY ONLINE RELEASE

Note: This article was posted on the *Archives* Web site as an Early Online Release. Early Online Release articles have been peer reviewed, copyedited, and reviewed by the authors. Additional changes or corrections may appear in these articles when they appear in a future print issue of the *Archives*. Early Online Release articles are citable by using the Digital Object Identifier (DOI), a unique number given to every article. The DOI will typically appear at the end of the abstract.

The DOI for this manuscript is doi: 10.5858/arpa.2020-0520-OA

The final published version of this manuscript will replace the Early Online Release version at the above DOI once it is available.

A Deep Learning Convolutional Neural Network Can Differentiate Between *Helicobacter Pylori* Gastritis and Autoimmune Gastritis With Results Comparable to Gastrointestinal Pathologists

Michael M. Franklin, DO; Fred A. Schultz, MA; Marissa A. Tafoya, MD; Audra A. Kerwin, MD; Cory J. Broehm, MD; Edgar G. Fischer, MD, PhD; Rama R. Gullapalli, MD, PhD; Douglas P. Clark, MD; Joshua A. Hanson, MD; David R. Martin, MD

● *Context.*—Pathology studies using convolutional neural networks (CNNs) have focused on neoplasms, while studies in inflammatory pathology are rare. We previously demonstrated a CNN differentiates reactive gastropathy, *Helicobacter pylori* gastritis (HPG), and normal gastric mucosa.

Objective.—To determine whether a CNN can differentiate the following 2 gastric inflammatory patterns: autoimmune gastritis (AG) and HPG.

Design.—Gold standard diagnoses were blindly established by 2 gastrointestinal (GI) pathologists. One hundred eighty-seven cases were scanned for analysis by HALO-AI. All levels and tissue fragments per slide were included for analysis. The cases were randomized, 112 (60%; 60 HPG, 52 AG) in the training set and 75 (40%; 40 HPG, 35 AG) in the test set. A HALO-AI correct area distribution (AD) cutoff of 50% or more was required to credit the CNN with the correct diagnosis. The test set was blindly reviewed by pathologists with different levels of GI pathology expertise as follows: 2 GI pathologists, 2 general surgical pathologists, and 2 residents. Each pathologist rendered their preferred diagnosis, HPG or AG.

Results.—At the HALO-AI AD percentage cutoff of 50% or more, the CNN results were 100% concordant with the gold standard diagnoses. On average, autoimmune gastritis cases had 84.7% HALO-AI autoimmune gastritis AD and HP cases had 87.3% HALO-AI HP AD. The GI pathologists, general anatomic pathologists, and residents were on average, 100%, 86%, and 57% concordant with the gold standard diagnoses, respectively.

Conclusions.—A CNN can distinguish between cases of HPG and autoimmune gastritis with accuracy equal to GI pathologists.

(*Arch Pathol Lab Med.* doi: 10.5858/arpa.2020-0520-OA)

Deep learning is a new and emerging paradigm of artificial intelligence (AI) research. Convolutional neural networks (CNN) are an algorithmic form of deep learning that are highly accurate in image-based recognition. These methods are capable of unsupervised learning, either from labeled or unlabeled data sets. Numerous past studies have demonstrated that CNNs can correctly recognize morphologic patterns of disease in anatomic pathology using whole slide images.[1–3] In gastrointestinal (GI) pathology, most of these studies have been related to neoplasia with promising results.[4–8] However, deep-learning analyses of GI inflammatory conditions have not been well studied. Wei et al[9] successfully trained a deep learning model to distinguish between celiac disease, normal duodenum, and nonspecific duodenitis. In a prior study, our group trained a CNN to identify regions of normal gastric mucosa, *Helicobacter pylori* (HP) gastritis (HPG), and reactive gastropathy.[10] In the current study, we examine the utility of a CNN to differentiate 2 clinically significant and morphologically similar GI biopsy diagnose, HPG and autoimmune gastritis (AG).

Helicobacter Pylori Gastritis

Although most individuals infected with HP have only mild or no symptoms, this comma-shaped, urease-producing bacterium causes significant disease burden worldwide.[10–13] Prevalence of infection is as high as 80% in some parts of the world, and is inversely proportional to socioeconomic status.[12,14] Individuals are most commonly infected in childhood, with overcrowded living conditions being the most significant risk factor.[12] Infection is associated with dyspepsia, acute and chronic gastritis, and an increased risk for both gastric adenocarcinoma and mucosa associated lymphoid tissue lymphoma.[11,13] In the United States, HP is the most common cause of peptic ulcer

Figure 1. *Hematoxylin-eosin stain, ×5.5 on digital slide viewer. This antral biopsy from a case of* Helicobacter pylori *gastritis demonstrates the classic superficial band-like chronic inflammatory infiltrate seen in the majority of cases.*

Figure 2. *Hematoxylin-eosin stain, ×5.5 on digital slide viewer. This gastric body biopsy in a case of autoimmune gastritis shows complete atrophy/loss of the normal oxyntic mucosa. The parietal cells and chief cells are gone and are replaced by mucous-neck cells (antralization). The chronic inflammation is spread evenly throughout the mucosa as opposed to the typical top-heavy distribution often seen in* Helicobacter pylori *gastritis.*
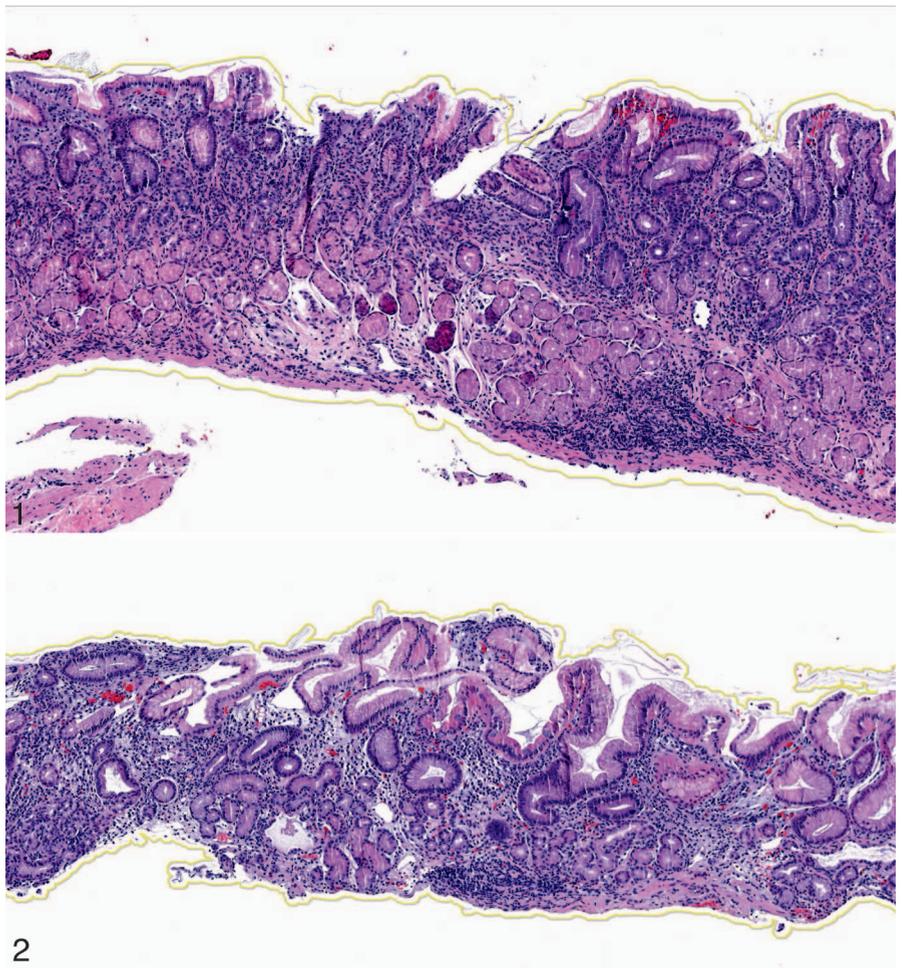
disease, leading to an annual health care cost of an estimated 6 billion dollars.[13]

HP typically resides in the mucin layer adjacent to the foveolar epithelium and induces a band-like, superficial lymphoplasmacytic infiltrate in the gastric mucosa, which is usually antral predominant (Figure 1).[11,13,15] Acute inflammation and lymphoid aggregates with germinal centers are also hallmarks of this infection, along with morphologic identification of the organisms in biopsy specimens.[13,15] In patients with chronic proton pump inhibitor use, the body of the stomach can also be heavily colonized and subsequently inflamed, often with organisms deep within oxyntic glands.[13,15,16]

Autoimmune Gastritis

Autoimmune gastritis (AG) is an immunologic disorder whose underlying pathophysiology has not been fully elucidated. It is believed to arise from 1 of the 2 following different etiologic mechanisms: primary autoimmune disease with antiparietal cell and anti-intrinsic factor antibodies, or a chronic immunologic process involving molecular mimicry, secondary to infection with HP.[16–18]

Regardless of the inciting mechanism, circulating autoantibodies to the parietal cell proton pump, as well as anti-intrinsic factor leads to destruction of oxyntic glands in the gastric body and fundus.[16–19] The gastric mucosa subsequently attempts to regenerate, and metaplastic glands replace native glands. Patients present with a variety of

clinical symptoms, including anemia or gastrointestinal (GI) complaints.[17,18] Diagnosis can be aided with serologic findings of hypergastrinemia, and the presence of anti-intrinsic factor and/or antiparietal cell antibodies. However, the diagnosis ultimately is confirmed on biopsy findings.

The histopathologic changes associated with AG can be broken into a spectrum of early- and late-histologic phases in the gastric body. Early-lymphoplasmacytic infiltration can be subtle and may be the only visible change in early disease. Unlike HPG, which creates a top-heavy infiltrate, the AG inflammatory pattern is full thickness within the lamina propria.[16,18]

Eventually, the full-thickness lamina propria chronic inflammation becomes dense and associated changes in oxyntic glands become evident, including destruction and metaplasia. The most common type of metaplasia has been referred to as antralization of oxyntic mucosa.[17,18] This is characterized by a patchy distribution of metaplastic glands with clear mucous-secreting cells, with an absence of native parietal cells (Figure 2).

Changes in the later phases are characterized by progressive gland destruction and continued metaplasia, including intestinal type and pancreatic acinar metaplasias. Eventually, atrophic gastric mucosa demonstrates complete replacement of native glands with metaplastic epithelium. Other associated changes include enterochromaffin-like–cell hyperplasia due to achlorhydria stimulating increased gastrin secretion from the antral G cells.[16,18,19] Reactive

changes are often seen in the antral foveolar epithelium, mimicking reactive antral gastropathy.

AG is considered a preneoplastic syndrome due to its association with both type-1 well-differentiated neuroendocrine tumors and gastric adenocarcinoma.[16,18,19] While the oncogenic pathway of adenocarcinoma is likely through chronic inflammation and intestinal metaplasia, well-differentiated neuroendocrine tumors arise from chronic stimulation of enterochromaffin-like cells by gastrin. In contrast to sporadic well differentiated neuroendocrine tumors (type 3), these demonstrate relatively benign behavior and an indolent clinical course.[16,19]

Diagnostic Challenges

Pathologists experienced with AG and HPG can distinguish these disease entities from one another, though confirmatory immunohistochemical stains are typically performed to demonstrate an absence of HP and the presence of enterochromaffin-like–cell hyperplasia in antralized oxyntic mucosa for cases of AG. However, many pathologists may struggle to make the diagnosis of AG in a chronically inflamed gastric biopsy lacking detectable HP. Too often, the search for a diagnosis stops with a negative HP stain and the case is signed out as "chronic active/chronic inactive gastritis, HP negative." We hypothesize that deep learning technology can discriminate between AG and HPG solely on hematoxylin-eosin (H&E)-stained slides, and therefore potentially serve as an aid for the general surgical pathologist in these challenging cases.

MATERIALS AND METHODS

This study was performed in accordance with the institutional review board requirements at the University of New Mexico (Albuquerque, New Mexico).

Gastric biopsy cases that carried a diagnosis of HPG or AG were obtained from our archives for review by 2 expert GI pathologists. Cases of HPG (n = 200) from 2018 were reviewed, and 100 classic examples were selected that demonstrated superficial band-like inflammation of antral and/or oxyntic mucosa and displayed at least several organisms on H&E or immunohistochemical stain. Cases of AG (n = 125) from 2012 to 2018 were reviewed, and 87 cases were selected that demonstrated classic morphologic features, such as the following: oxyntic gland atrophy, antral-type/intestinal-type/pancreatic acinar-type metaplasia, transmucosal inflammation and enterochromaffin-like–cell hyperplasia on immunostain, with relative noninflamed antral fragments. The GI pathologists' diagnoses based on this case selection were considered the gold standard diagnoses.

The 187 cases (100 HP and 87 AG) were scanned with the Aperio VERSA 200 slide scanner (Leica, Wetzlar, Germany) at ×40 magnification and imported into a computer containing a 12-core, 2.2-GHz Intel Xeon Processor E5-2650 chip and a Nvidia Titan XP graphics card. HALO-AI image analysis software (Indica Labs, Albuquerque, NM) was used to perform training and testing. HALO-AI uses a fully convolutional version of the VGG architecture with padding removed.

Each tissue fragment on each slide, including all levels and all fragments from both the gastric body and antrum, were annotated and included for analysis. Of the 187 cases, 112 (60%) were randomly selected for inclusion in the training set (60 HP, 52 AG) and 75 cases (40%) were randomly included in the test set (40 HP, 35 AG).

Training was performed by HALO-AI on cases labeled as either AG or HPG according to the gold standard diagnosis. These training cases were broken down into "image patches" of 400 × 400 pixels (where 1 pixel = 1 μm), at a resolution corresponding to a ×4 digital view magnification. This magnification corresponded to a

digital view in which the 2 GI pathologists could reliably confirm the gold standard diagnosis in a digital field. Approximately 30% to 50% of the biopsy fragment could be viewed at this magnification.

Within the confines of the previously annotated entire tissue fragments, the image patches (400 × 400 μm within a ×4 digital field view) analyzed by HALO-AI were generated by automated selection of random points and cropping a patch around the point. These patches were further augmented with random rotations and random shifts to hue, saturation, contrast, and brightness. Training was performed for a total of 125 396 analytic iterations using RMSProp (delta of 0.9) with a learning rate of 1^{e-3} reducing the learning rate by 10% every 2k iterations and an L2 regularization of 5^{e-4}. Because no padding was used, the tile size was increased to 1867 × 1867, thus enhancing performance without changing the output. During these iterations, the algorithm would change the node-weighted values based off the gold standard call of the 2 GI pathologists continuously. The HALO-AI operator stopped the algorithm once an error rate/cross entropy rate of less than 0.01 was achieved.

For the test set, HALO-AI analyzed these cases blindly, assigning each region (patch) a likelihood score for that area, which corresponded to the most probable diagnostic call. The output for each test case was ultimately X% area AG versus X% area HPG based on these calls. These percentages are termed "area distribution" (AD) in this paper, with each test case receiving an AG AD (% of the case favored to represent AG) and HP AD (% of the case favored to represent HP). HPG was labeled with a red label, while AG was labeled with a green label (Figure 3, A through D).

Four general surgical pathologists at the University of New Mexico (2 senior residents, 1 junior faculty, and 1 senior faculty; all without GI fellowship training) blindly reviewed the same digital test cases and put them into 1 or 2 of the following diagnostic categories: HPG or AG. The 2 GI pathologists also reviewed the digital cases after a 1-month washout period. During the testing phase, pathologists only reviewed scanned H&E slides and were not given access to previously performed immunohistochemical stains. All pathologists were given a 3-hour time restriction to complete their review. Their results were then compared with the gold standard diagnoses and the deep-learning results.

RESULTS

Complete HALO-AI AD percentages for the test sets (40 HPG cases and 35 AG cases) are documented in Tables 1 and 2, respectively. The tables are organized from the highest to the lowest HALO-AI AD determination for each diagnosis. Each table also includes the raw total tissue area evaluated by the CNN and the raw tissue area determined to represent HP AD and AG AD for each test case. At an AD percentage cutoff of 50% or more (meaning the CNN assigned the correct diagnosis to one half or more of the tissue analyzed), the CNN results were 100% concordant with the gold standard diagnoses for both test sets (Tables 1 and 2). For cases of HPG, the mean HP AD was 87.3% (range 59.2%–99.7%), and for cases of AG, the mean AG AD was 84.8% (range 52.5%–98.2%), showing substantial CNN agreement with the gold standard diagnoses. The GI pathologists, general surgical pathologists, and residents were on average, 100%, 86%, and 57% concordant with the gold standard diagnoses, respectively (Table 3).

DISCUSSION

We herein demonstrate that a CNN can distinguish between HPG and AG with accuracy equal to 2 expert GI pathologists. To our knowledge, this is the first study investigating the use of deep learning to discriminate between AG and HPG. Previously, we showed that deep learning could accurately discriminate between normal
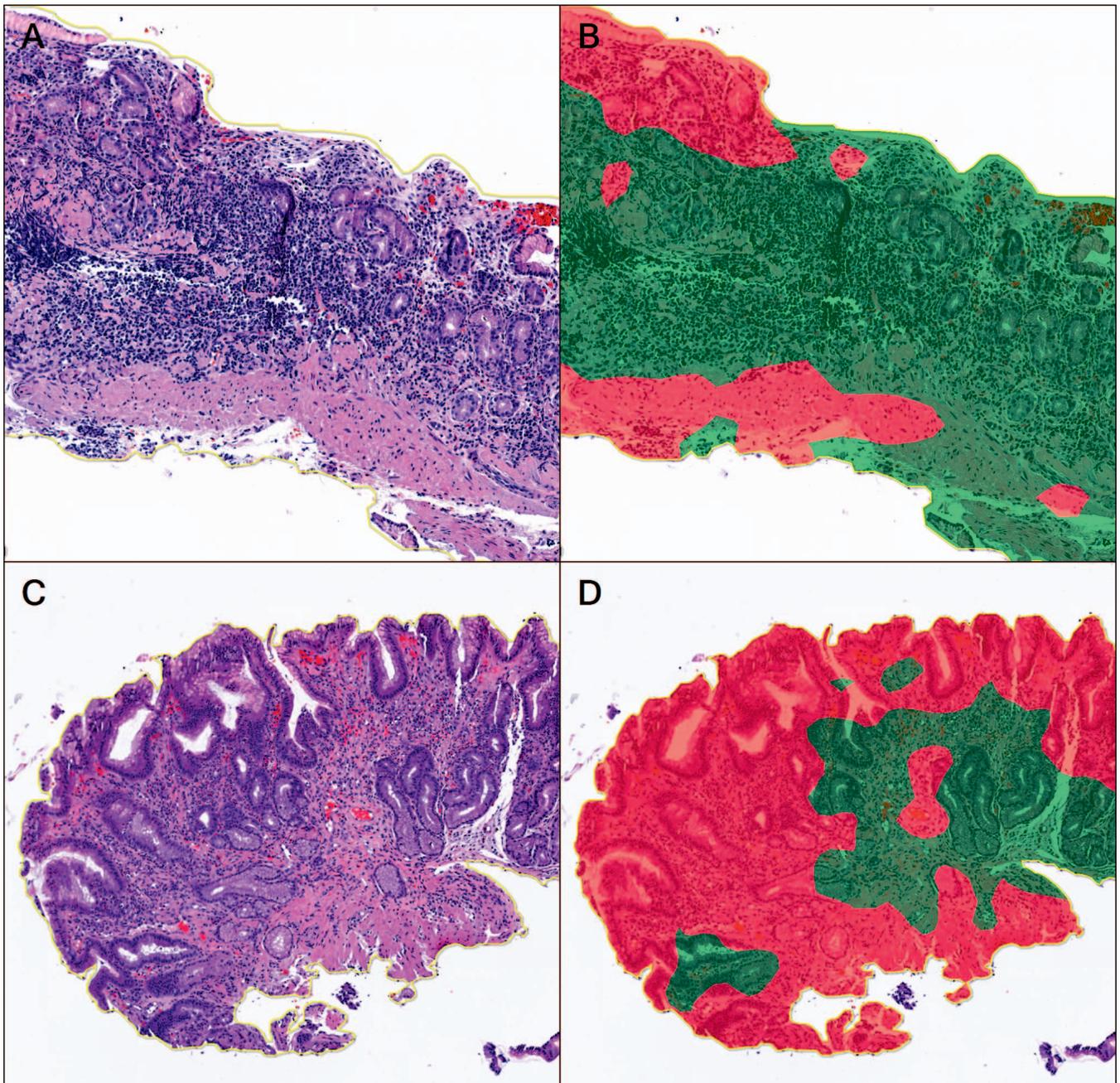
Figure 3. *Hematoxylin-eosin stain, ×4 on digital slide viewer. (A) Autoimmune gastritis demonstrating lymphoplasmacytic inflammation spread evenly throughout the lamina propria in atrophic/antralized gastric body mucosa. (B) The paired HALO-AI green area distribution (AD) shows a majority call of autoimmune gastritis, and thus the convolutional neural network (CNN) is credited for a correct diagnosis. (C) A case of* Helicobacter pylori *gastritis shows the classic top-heavy chronic inflammatory infiltrate. (D) HALO-AI correctly assigns a red* H pylori *AD to the majority of this tissue fragment and is credited with a correct diagnosis.*

gastric mucosa, reactive gastropathy, and HPG, performing best on cases of HPG.[10] In daily practice the distinction between reactive gastropathy and HPG is relatively straightforward, hence, we sought to challenge the CNN with inflamed gastric biopsies; the 2 most common patterns being AG and HPG. In our clinical experience, AG can be a difficult diagnosis to make, especially for pathologists with minimal training in inflammatory GI pathology.

We trained HALO-AI on examples of AG and HPG in which 2 expert GI pathologists concurred on the diagnosis. The 75 test cases (all cases in the study were randomized

into the training set and test set beforehand) were randomized, and 2 residents, 2 general surgical pathologists, and 2 expert GI pathologists were given the scanned H&E test set to blindly diagnose. To minimize any diagnostic variables, HALO-AI and the pathologists were tested in near identical conditions. All participants viewed digitally scanned slides, without immunohistochemistry, and had to choose a diagnosis of AG versus HPG per case. Given that HALO-AI highlights diagnostic regions according to like-lihood calls per image patch, the results were presented as a percentage of the biopsy favored to represent AG versus

Correct AD, %	Total Classified Area, mm²	HP AD, mm²	AG AD, mm²
99.7	12.7	12.7	0.0
98.9	39.1	38.7	0.4
98.0	13.5	13.2	0.3
96.8	27.7	26.8	0.9
96.5	38.1	36.8	1.3
96.0	47.2	45.3	1.9
95.4	59.9	57.1	2.7
94.9	30.8	29.2	1.6
94.9	41.0	38.9	2.1
94.2	35.0	32.9	2.0
94.1	10.8	10.2	0.6
94.0	63.8	60.0	3.8
93.6	11.4	10.7	0.7
93.3	58.7	54.8	3.9
93.0	48.0	44.7	3.3
92.3	39.2	36.2	3.0
92.2	39.3	36.3	3.1
91.7	12.4	11.4	1.0
91.4	11.5	10.5	1.0
91.3	16.0	14.6	1.4
90.7	16.7	15.1	1.5
89.3	18.7	16.7	2.0
87.6	19.7	17.3	2.4
87.1	10.2	8.9	1.3
86.4	52.2	45.1	7.1
84.4	41.1	34.7	6.4
82.7	26.2	21.6	4.5
82.2	30.4	24.9	5.4
81.5	25.2	20.6	4.7
81.0	54.5	44.1	10.3
80.6	56.6	45.7	11.0
79.9	31.1	24.8	6.2
79.8	60.4	48.2	12.2
79.6	47.0	37.4	9.6
79.4	26.6	21.1	5.5
78.3	38.1	29.9	8.3
75.4	33.5	25.3	8.3
74.9	42.9	32.1	10.8
61.2	9.3	5.7	3.6
59.2	43.1	25.5	17.6

Table 1. HALO-AI Area Distribution (AD) Calls for Gold Standard *Helicobacter pylori* Cases (n = 40)

Abbreviations: AG AD, autoimmune gastritis area distribution; HP AD, *Helicobacter pylori* area distribution.

Correct AD, %	Total Classified Area, mm²	AG AD, mm²	HP AD, mm²
98.2	22.6	22.2	0.4
97.1	22.3	21.6	0.7
96.7	25.3	24.5	0.8
93.5	14.8	13.8	1.0
93.4	8.7	8.2	0.6
92.2	23.8	21.9	1.8
91.9	17.4	16.0	1.4
91.7	29.3	26.8	2.4
91.5	12.6	11.5	1.1
91.5	41.5	38.0	3.5
91.4	43.1	39.4	3.7
90.8	18.0	16.3	1.7
90.2	173.3	156.3	17.0
90.2	48.8	44.0	4.8
88.7	33.4	29.7	3.8
88.4	23.4	20.7	2.7
88.3	11.2	9.9	1.3
87.6	29.2	25.6	3.6
87.5	37.1	32.5	4.6
87.4	21.6	18.9	2.7
87.3	10.4	9.1	1.3
87.3	22.1	19.3	2.8
86.4	27.3	23.6	3.7
85.4	28.8	24.6	4.2
83.7	33.7	28.2	5.5
79.9	13.3	10.6	2.7
79.8	39.8	31.7	8.1
78.0	14.5	11.3	3.2
77.0	41.7	32.1	9.6
75.5	71.0	53.6	17.4
75.3	57.6	43.4	14.2
72.2	18.0	13.0	5.0
65.9	34.5	22.8	11.8
53.2	9.9	5.3	4.6
52.5	23.6	12.4	11.2

Table 2. HALO-AI Area Distribution (AD) Calls for Gold Standard Autoimmune Gastritis Cases (n = 35)

Abbreviations: AG AD, autoimmune gastritis area distribution; HP AD, *Helicobacter pylori* area distribution.

HPG. Because of this, all pathologists were instructed to choose the diagnosis they favored in the majority of tissue fragments per case as well.

Under these parameters, the data are compelling. The CNN algorithm assigned the correct diagnosis in 100% of cases (ie, the majority of the biopsy was called in accordance with the gold standard diagnosis), in agreement with both expert GI pathologists. The general surgical pathologists were correct 129 of 150 (86%) of the time, and the pathology residents were correct 86 of 150 (57%) of the time.

Another interesting facet of this project was that all tissue fragments present on each slide were included for HALO-AI analysis. This means that if the endoscopist submitted gastric body and antrum in the same specimen jar, both

regions were included and labeled as either AG or HPG for analysis. Typically, HPG has antral predominant inflammation, whereas the body/fundus is usually less inflamed. Because entire fragments were labeled in accordance with the gold standard diagnosis for the training sets, HALO-AI incorporated a spectrum of inflammation/metaplastic change into its learning. Likewise, AG-associated inflammation/injury is restricted to the body/fundus, while the antrum is spared. It therefore follows that HALO-AI was trained to recognize the heterogeneous/zonal nature of AG. Despite this, the algorithm was not confused, and achieved more than 50% AD on every test case. Because this is how training/tested was performed, this algorithm requires minimal human annotation/interaction.

We have now trained 2 different CNN algorithms to recognize the most commonly encountered gastric pathologies as follows: normal gastric mucosa, reactive gastropathy, HPG, and AG.[10] While most gastric biopsies are diagnostically unambiguous, AG can be a challenge for

Table 3. Pathologist Versus CNN Concordance With Gold Standard Diagnoses

Reviewer	Average % Concordant Results With the Gold Standard Diagnosis
CNN (HALO-AI)	100%
Expert GI pathologists (n = 2)	100%
Non-GI fellowship trained surgical pathologists (n = 2)	86%
Pathology residents (n = 2)	57%

Abbreviations: CNN, convolutional neural network; GI, gastrointestinal.

generalists and trainees. If deep learning algorithms such as these were implemented into general surgical pathology clinical practice, this would likely improve diagnostic accuracy for generalists and trainees alike. Perhaps in the future, after a slide is scanned, algorithms such as these could rapidly "prescreen" cases to trigger the ordering of confirmatory immunohistochemical stains as follows: HP, gastrin, and chromogranin/synaptophysin. The pathologist could then arrive at the correct diagnosis on the same day they receive the H&E-stained slides. This could potentially obviate the wasteful practice of ordering up-front stains on all gastric biopsies while simultaneously ensuring rapid 24-hour turnaround on this very common surgical pathology specimen.

Deep learning may also serve as a training tool for residents. A resident on a busy GI pathology rotation could render a preview diagnosis and then have immediate access to an AI digital label classifier displaying the favored AI diagnosis. Such rapid feedback has the potential to provide on-demand instruction during case preview and may help residents refine their diagnostic skills before case review with their attending. However, training programs would have to ensure that the AI algorithms are used as a learning tool and not as a diagnostic crutch.

This study has a few areas of weakness. First, it is not completely indicative of "real-life practice." Resident and general surgical pathologist accuracies would likely improve if they had access to the immunohistochemical stains, patient information, and no time restriction. That said, the 2 expert GI pathologists were 100% concordant with gold standard diagnoses after testing solely on the scanned H&E slides, indicating that the morphologic differences between AG and HPG are reliably detected by pathologists with subspecialized training.

Another limitation is that cases of multifocal atrophic *Helicobacter pylori* (MAHP) gastritis, a pattern of chronic HP infection that is often confused with AG, were not included for analysis. MAHP gastritis cases were not included because we did not have enough examples to create training and test sets for this entity. For curiosity's sake, we tested the CNN on one recent case of MAHP gastritis from out

archives. Of note, the CNN generated a correct HP AD of 71%, thus arriving at the correct diagnosis.

In conclusion, we demonstrate that a CNN can distinguish between unequivocal cases of HPG and AG with accuracy equal to expert GI pathologists, albeit in a controlled research setting. This was accomplished while analyzing all/entire tissue fragments from the respective cases with minimal human annotation. We believe that deep learning may be able to serve as a diagnostic aid to pathologists at different levels of training in cases of inflammatory gastric pathology.

References

1. Khosravi P, Kazemi E, Imielinski M, Elemento O, Hajirasouliha I. Deep convolutional neural networks enable discrimination of heterogeneous digital pathology images. *EBioMedicine*. 2018;27(1):317–328. doi:10.1016/j.ebiom.2017.12.026

2. Xu Y, Jia Z, Wang LB, et al. Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features. *BMC Bioinformatics*. 2017;18(1):1–17. doi:10.1186/s12859-017-1685-x

3. Janowczyk A, Madabhushi A. Deep learning for digital pathology image analysis: a comprehensive tutorial with selected use cases. *J Pathol Inform*. 2016;7(1):29. doi:10.4103/2153-3539.186902

4. Qu J, Hiruta N, Terai K, Nosato H, Murakawa M, Sakanashi H. Gastric pathology image classification using stepwise fine-tuning for deep neural networks. *J Healthc Eng*. 2018;2018(1):1–13. doi:10.1155/2018/8961781

5. Ning Z, Luo J, Li Y, et al. Pattern classification for gastrointestinal stromal tumors by integration of radiomics and deep convolutional features. *IEEE J Biomed Heal Informatics*. 2019;23(3):1181–1191. doi:10.1109/JBHI.2018.2841992

6. Korbar B, Olofson AM, Miraflor AP, et al. Deep-learning for classification of colorectal polyps on whole-slide images. *J Pathol Inform*. 2017;8(1):1–30. doi:10.4103/jpi.jpi_34_17

7. Ribeiro E, Uhl A, Wimmer G, Häfner M. Exploring deep learning and transfer learning for colonic polyp classification. *Comput Math Methods Med*. 2016;2016(1):1–16. doi:10.1155/2016/6584725

8. Sirinukunwattana K, Raza SEA, Tsang YW, Snead DRJ, Cree IA, Rajpoot NM. Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. *IEEE Trans Med Imaging*. 2016;35(5):1196–1206. doi:10.1109/TMI.2016.2525803

9. Wei JW, Wei JW, Jackson CR, Ren B, Suriawinata AA, Hassanpour S. Automated detection of celiac disease on duodenal biopsy slides: a deep learning approach. *J Pathol Inform*. 2019;10(1):7. doi:10.4103/2153-3539.253722

10. Martin DR, Hanson JA, Gullapalli RR, Schultz FA, Sethi A, Clark DP. A deep learning convolutional neural network can recognize common patterns of injury in gastric pathology. *Arch Pathol Lab Med*. 2020;144(3):370–378.

11. Dubois A. Intracellular *Helicobacter pylori* and gastric carcinogenesis: an "old" frontier worth revisiting. *Gastroenterology*. 2007;83(2):1–29. doi:10.1053/j.gastro2007.01.068

12. Malaty H. Epidemiology of *Helicobacter pylori* infection. *Best Pract Res Clin Gastroenterol*. 2007;21(2):205–214. doi:10.1016/j.bpg.2006.10.005

13. Testerman TL, Morris J. Beyond the stomach: an updated view of *Helicobacter pylori* pathogenesis, diagnosis, and treatment. *World J Gastroenterol*. 2014;20(36):12781–12808. doi:10.3748/wjg.v20.i36.12781

14. Sonnenberg A, Lash RH, Genta RM. A National study of *Helicobacter pylori* infection in gastric biopsy specimens. *Gastroenterology*. 2010;139(6):1894–1901.e2. doi:10.1053/j.gastro.2010.08.018

15. Goldblumm J, Odze R. *Odze and Goldblum Surgical Pathology of the Gi Tract, Liver, Biliary Tract*. 3rd edition. Philadelphia, PA: Elsevier Health Sciences; 2014.

16. Montgomery E, Voltaggio L. *Biopsy Interpretation of the Gastrointestinal Tract Mucosa*. 3rd edition. Philadephia, PA: Wolters Kluwer; 2018.

17. Massironi S, Zilli A, Elvevi A, Invernizzi P. The changing face of chronic autoimmune atrophic gastritis: an updated comprehensive perspective. *Autoimmun Rev*. 2019;18(3):215–222. doi:10.1016/j.autrev.2018.08.011

18. Coati I, Fassan M, Farinati F, Graham DY, Genta RM, Rugge M. Autoimmune gastritis: pathologist's viewpoint. *World J Gastroenterol*. 2015;21(42):12179–12189. doi:10.3748/wjg.v21.i42.12179

19. Choi W-T, Lauwers GY. Patterns of gastric injury. *Surg Pathol Clin*. 2017;10(4):801–822. doi:10.1016/j.path.2017.07.003