

Computational pathology reveals unique spatial patterns of immune response in H&E images from COVID-19 autopsies: preliminary findings

Germán Corredor^{a,b,†}, Paula Toro^{a,†}, Kaustav Bera^a, Dylan Rasmussen^a,
Vidya Sankar Viswanathan^a, Christina Buzzy^a, Pingfu Fu^c,
Lisa M. Barton^d, Edana Stroberg^d, Eric Duval^d, Hannah Gilmore^e,
Sanjay Mukhopadhyay^f and Anant Madabhushi^{a,b,*}

^aCase Western Reserve University, Center for Computational Imaging and Personalized Diagnostics, Cleveland, Ohio, United States

^bLouis Stokes Cleveland VA Medical Center, Cleveland, Ohio, United States

^cCase Western Reserve University, Department of Population and Quantitative Health Sciences, Cleveland, Ohio, United States

^dOklahoma Office of the Chief Medical Examiner, Oklahoma City, Oklahoma, United States

^eUniversity Hospitals, Department of Pathology, Cleveland, Ohio, United States

^fCleveland Clinic, Department of Pathology, Cleveland, Ohio, United States

Abstract

Purpose: We used computerized image analysis and machine learning approaches to characterize spatial arrangement features of the immune response from digitized autopsied H&E tissue images of the lung in coronavirus disease 2019 (COVID-19) patients. Additionally, we applied our approach to tease out potential morphometric differences from autopsies of patients who succumbed to COVID-19 versus H1N1.

Approach: H&E lung whole slide images from autopsy specimens of nine COVID-19 and two H1N1 patients were computationally interrogated. 606 image patches (~55 per patient) of 1024 × 882 pixels were extracted from the 11 autopsied patient studies. A watershed-based segmentation approach in conjunction with a machine learning classifier was employed to identify two types of nuclei families: lymphocytes and non-lymphocytes (i.e., other nucleated cells such as pneumocytes, macrophages, and neutrophils). Based off the proximity of the individual nuclei, clusters for each nuclei family were constructed. For each of the resulting clusters, a series of quantitative measurements relating to architecture and density of nuclei clusters were calculated. A receiver operating characteristics-based feature selection method, violin plots, and the t-distributed stochastic neighbor embedding algorithm were employed to study differences in immune patterns.

Results: In COVID-19, the immune response consistently showed multiple small-size lymphocyte clusters, suggesting that lymphocyte response is rather modest, possibly due to lymphocytopenia. In H1N1, we found larger lymphocyte clusters that were proximal to large clusters of non-lymphocytes, a possible reflection of increased prevalence of macrophages and other immune cells.

Conclusion: Our study shows the potential of computational pathology to uncover immune response features that may not be obvious by routine histopathology visual inspection.

© 2021 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: [10.1117/1.JMI.8.S1.017501](https://doi.org/10.1117/1.JMI.8.S1.017501)]

Keywords: computational pathology; coronavirus disease 2019; H1N1; image processing; immune response; machine learning.

Paper 20337SSRR received Dec. 10, 2020; accepted for publication Jun. 28, 2021; published online Jul. 13, 2021.

*Address all correspondence to Anant Madabhushi, anant.madabhushi@case.edu

†These authors contributed equally to this work.

1 Introduction

Coronavirus disease 2019 (COVID-19) is caused by the severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2). SARS-CoV-2 is absorbed by the mucosae of the respiratory tract and targets bronchial cells and pneumocytes. Once SARS-CoV-2 is internalized into the host cell, it is sensed by receptors especially present in native cells and macrophages, a subtype of immune cell.^{1,2} The binding of viral particles to their receptors leads to a pro-inflammatory response by the infected cells. The innate immune response is the immediate mechanism of defense by the immune system and is led by neutrophils and macrophages, which release molecules (cytokines) that act locally to stop the aggressor as well as attract lymphocytes responsible for viral recognition and regulation of the immune cell response.^{2,3} This phase, led by both T and B lymphocytes, is known as the adaptive immune response.⁴

In lung tissue from severely ill COVID-19 patients, some cytokines are overexpressed⁴ (cytokine storm), which is thought to result in severe collateral damage in the affected tissues, driving organ failure.² This cell network between immune and surrounding cells has been morphologically explored previously, mainly in routine histopathology postmortem tissue from people who have died on account of COVID-19.^{5–13} Additionally, the most common immune cell response consists of a predominance of interstitial T-lymphocytes.¹⁴ In contrast, B cells are consistently few while neutrophils and macrophages are heterogeneously reported.¹⁴

H1N1 is a viral disease that is still causing thousands of deaths.¹⁵ Previous works^{15–17} have highlighted the importance of comparing the differences between COVID-19 and H1N1 since they share common etiologies and similar clinical symptoms, occur in the same season, and were considered pandemics in recent times. Given the novelty of COVID-19 and the non-specific nature of clinical presentation of both diseases, there is a need for better understanding how these diseases impact the immune system.

When comparing COVID-19 with H1N1, it has been found that subjects with COVID-19 present (1) higher median age of onset; (2) more prevalence of symptoms such as non-productive cough, fatigue, and gastrointestinal disturbances; (3) more ground-glass opacities in radiology images; (4) worse response to some oxygen therapies; and (5) wider spectrum of pharmacotherapy with different responses.¹⁵ A possible reason for these could be on account of the differentially expressed immune response between H1N1 and SARS-CoV-2.^{15,18,19} In the lung, the spectrum of histopathological findings in COVID-19 includes diffuse alveolar damage, fibrin microthrombi, desquamative pneumocytes, and superposed changes of bronchopneumonia; all of them accompanied by T lymphocytic infiltration, in mild to moderate degree. In H1N1, additional findings such as necrosis of main airways, extensive hemorrhage,²⁰ superimposed acute bronchopneumonia, and a considerable amount of interstitial macrophages accompanying lymphocytic infiltration^{21,22} have been more consistently found. However, none of these studies used objective metrics for comparing differences in architectural patterns of the immune response in both diseases using pathology slides.

Although visual assessment of hematoxylin–eosin (H&E)-stained tissues and immunohistochemistry of autopsied specimens in both COVID-19 and H1N1^{14,23} provide a gestalt interpretation, it involves some degree of subjectivity and does not yield quantitative measurements. Computational pathology, image analysis, and machine learning approaches have been previously employed for quantitative characterization of the immune response in digitized H&E tissue pathology images of diseases such as breast and lung cancer.^{24–27}

In this preliminary work, we introduce a new computational imaging framework to characterize the disease microenvironment from H&E whole slide images (WSIs) of postmortem lung tissue belonging to nine patients who succumbed to COVID-19. First, image processing and machine learning were employed to automatically identify two nuclei types in each image tile: “lymphocytes” and “non-lymphocytes.” Although non-lymphocytes in lung tissue include pneumocytes, macrophages, neutrophils, among other nucleated cells, in this study, these individual cell types were not differentiated but treated as corresponding to a single family. Then nuclei clusters for both types were built based on the proximity of the individual nuclei. In addition, we present a new set of quantitative metrics that capture the density and spatial interplay between clusters of the same and different nuclei types. Unlike the previous work employing nuclei-graph methods,²⁸ this approach takes into account inter- and intraprimitive relationships

(i.e., relationships between nuclei types of the same or of different families), thereby enabling the exploration of complex cellular interactions within the microenvironment. Although, in this work, we showcase the utility of this approach in characterizing the interplay and spatial relationships between two different nuclei types (i.e., lymphocytes and non-lymphocytes), the approach can be extended to other nuclei types and also to a larger number of nuclei families (i.e., >2). Additionally, we apply this approach to characterize the architectural nuclei arrangement found in two patients who died due to H1N1 and identify a set of image features relating to the immune morphology that were most different between H1N1 and COVID-19 patients.

2 Materials and Methods

2.1 Dataset

The dataset consisted of WSIs from 11 patients who died either due to H1N1 ($N = 2$) or COVID-19 ($N = 9$). These were obtained from the National Institutes of Health (NIH) COVID-19 Digital Pathology Repository²⁹ ($N_{\text{h1n1}} = 1$ and $N_{\text{covid}} = 6$), Cleveland Clinic (CC) ($N_{\text{h1n1}} = 1$), the Oklahoma Office of the Chief Medical Examiner (OCME) ($N_{\text{covid}} = 1$), and University Hospitals (UH) ($N_{\text{covid}} = 2$). All cases were anonymized. The NIH slides were digitized at 20 \times while the others were digitized at 40 \times magnification. CC and UH slides were digitized using a Ventana iScan HT scanner; OCME slides using an Aperio AT Turbo.

Due to the different magnification levels of the images, the CC, OCME, and UH images were downsampled to 20 \times to make them comparable to the NIH images. From this dataset, 606 1024 \times 882-pixel patches (~ 55 /patient) were randomly extracted (patches with $\geq 70\%$ background were discarded). Random extraction of patches has been used in the previous studies aiming to model the heterogeneity of the lesion. For example, Khan and Yuan³⁰ developed a “virtual biopsy approach,” in which non-overlapping tumor regions were randomly sampled to characterize the degree of lymphocytic infiltration in breast cancer samples.

For the NIH cases, as we were unable to download the WSIs, screenshots of the WSIs were directly taken from the web browser using Microsoft Windows’ native screenshot function. Although extracting screenshots may sound unorthodox, it is worth mentioning that the NIH repository employs the HALO Link web microscope, which offers a high quality of visualization. Image quality was not altered by taking screenshots since the tissue structures were preserved adequately. In fact, the screenshots were shown to our pathology collaborators, and they confirmed they were equally valid as the patches directly obtained from image files corresponding to tissue slides digitized at 20 \times magnification.

2.2 Automatic Identification of Lymphocytes

For automatic identification of lymphocytes (Fig. 1), image color normalization³¹ was applied to compensate staining variations of slides acquired from different institutions [Fig. 1(c)]. Next, the method of Veta et al.,³² which has been used by previous works,^{26,33,34} was used for segmenting all image nuclei [Fig. 1(d)]. This method is a watershed-based model specifically tuned for nuclei segmentation of H&E images. In an independent testing set containing 18 histopathology cases, it showed a sensitivity of 0.85 and mean predictive value of 0.89.³²

Finally, the model by Corredor et al.³⁵ was employed to classify each nucleus as either “lymphocyte” or “non-lymphocyte.” This model consists in a support vector machine with linear kernel that employs image-derived features of each segmented nucleus related to texture, shape, and color (Fig. 1, bottom row). Authors of the method report that in an independent testing set containing 1026 lung nuclei (manually annotated and validated by an expert pathologist), it yielded a precision of 0.89 and recall of 0.83.³⁵

2.3 Characterizing the Spatial Architecture of Lymphocytes and Surrounding Nucleated Cells

Once lymphocytes and non-lymphocytes were automatically identified, different clusters for each family were built based on the distance between nuclei. First, lymphocytes were connected

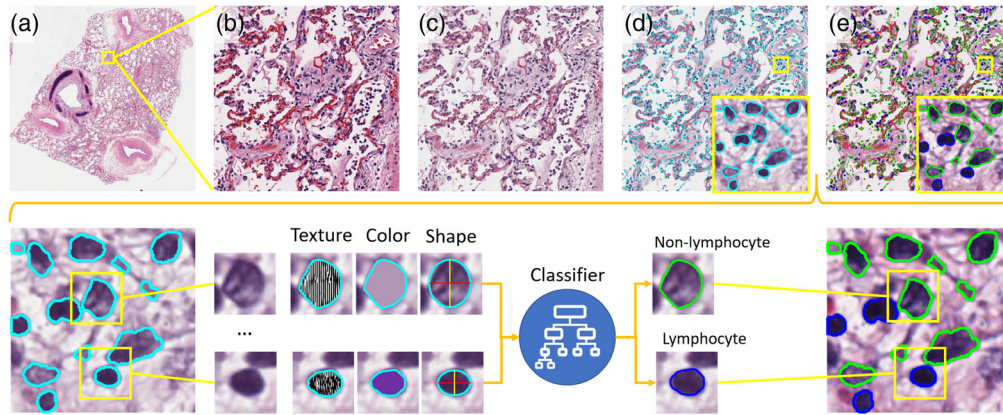


Fig. 1 Procedure for automatic detection of lymphocytes. Top row: (a) A WSI; (b) an enlarged field of view extracted from the WSI; (c) field of view after applying normalization of color; (d) automatic segmentation of nuclei; and (e) detection of lymphocytes. Bottom row: For each segmented nucleus, a set of features related to texture, color, and shape are extracted. Then a previously trained classifier (a support vector machine) uses such features to label each nucleus as either a lymphocyte or a non-lymphocyte.

with each other if the distance between them is below a predefined threshold. A very large threshold value generates just a single complete graph, which limits analyzing interactions within the disease microenvironment; whereas a small value (close to zero) produces multiple sparse graphs, which approximates to analyze individual nuclei. For this reason, in this study, the value of the threshold was empirically determined to be 50 pixels ($\sim 12.5 \mu\text{m}$) because it showed reasonable trade-off in the number of generated graphs. This action was carried out for each lymphocyte until all lymphocytes have been visited and potentially connected with corresponding proximal lymphocytes. This results in n disconnected subgraphs of lymphocytes [Fig. 2(b)]. The same process is carried out for non-lymphocytes, thereby generating m disconnected subgraphs of non-lymphocytes. Then a set P of polygons is constructed and defined by $P = \{P_1^L, P_2^L, \dots, P_n^L, P_1^{NL}, P_2^{NL}, \dots, P_m^{NL}\}$, where P_i^L is the convex hull³⁶ that contains all the nuclei of the i 'th lymphocyte subgraph and P_j^{NL} is the convex hull of the j 'th non-lymphocyte subgraph [Fig. 2(c)]; each polygon represents a cluster. The convex hull has been used previously to model the area of influence, for example, the spatial extent of epidemics at the outbreak stage.^{37,38} Given the dynamics of the disease microenvironment, in this work, the convex hull concept was employed to model the area of influence of nuclei, taking into account their movements within the tissue, the associated cytoplasm, and connecting stroma.

Based on the built clusters, six groups of features were derived.

1. *Number of clusters of each nuclei type.* N_f represents the number of clusters of the f 'th nuclei type, where $f \in \{1, 2, \dots, F\}$, where F is the total number of families ($F \geq 2$).
2. *Density of each cluster.* $d_k = n_k/A_k$, where d_k is the density of the k 'th cluster, n_k is the number of nodes constituting the cluster k , and A_k is the area of cluster k , i.e., the total number of pixels covered by the convex hull C of the cluster nodes.
3. *Intersection between clusters.* $A_i \cap A_j, \frac{A_i \cap A_j}{A_i}, \frac{A_i \cap A_j}{A_j}, \frac{A_i \cap A_j}{(A_i + A_j)/2}$, where $A_i \cap A_j$ is the intersection between clusters i and j , A_i is the area of cluster i , and A_j is the area of cluster j [see Fig. 3(a)].
4. *Diversity of clusters surrounding a specific cluster.* $S_{f,k}/N_k$, where $S_{f,k}$ represents the numbers of clusters of nuclei type f surrounding the k 'th cluster ($f \in \{1, \dots, F\}$) and S_k represents the total number of clusters surrounding cluster k . This measure is applied for a predefined number of neighbors [see Fig. 3(b)].
5. *Measures from global graphs built for each nuclei type.* A complete and undirected graph $G = (O, E, W)$ was defined for each nuclei type, where $O = \{o_1, \dots, o_F\}$ is the set of nodes corresponding to the convex hull centroids of each nuclei type, $E = \{e_1, e_2, \dots, e_m\}$ is the set of edges connecting such centroids such that

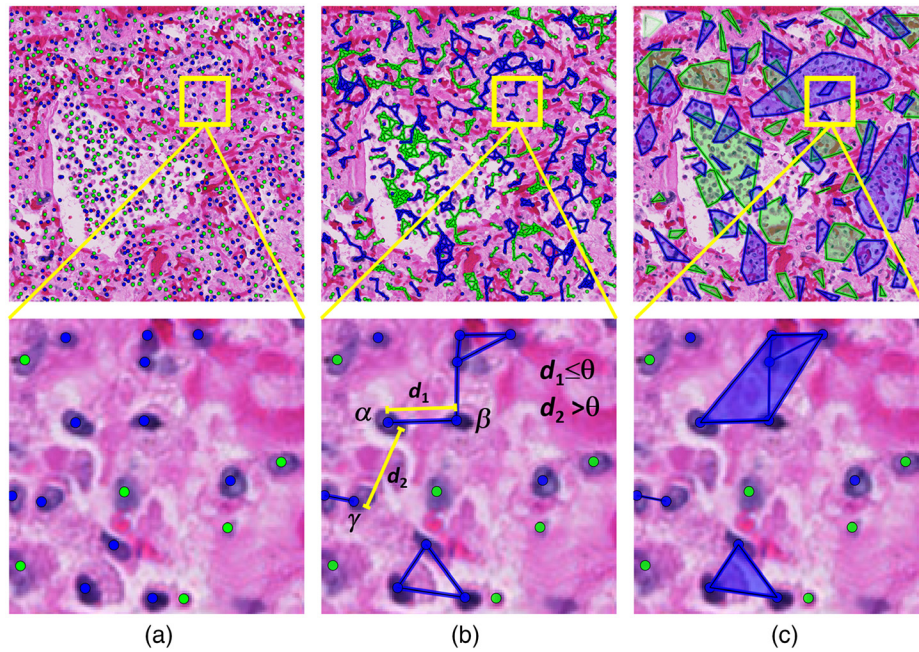


Fig. 2 Construction of nuclei clusters: (a) identification of two families of nuclei (blue: lymphocytes and green: non-lymphocytes); (b) linking of nuclei based on proximity; and (c) resulting nuclei clusters after drawing the convex hull that contains all the linked nuclei. Bottom cell of (b) illustrates three nuclei α , β , and γ . Since the distance between α and β (d_1) is less than the threshold θ , they are linked. On the contrary, the distance between α and γ (d_2) is larger than θ , so they are not connected. As a result, they belong to different clusters. The nuclei that are not linked to other nuclei cannot conform clusters, so they are not considered for the analysis.

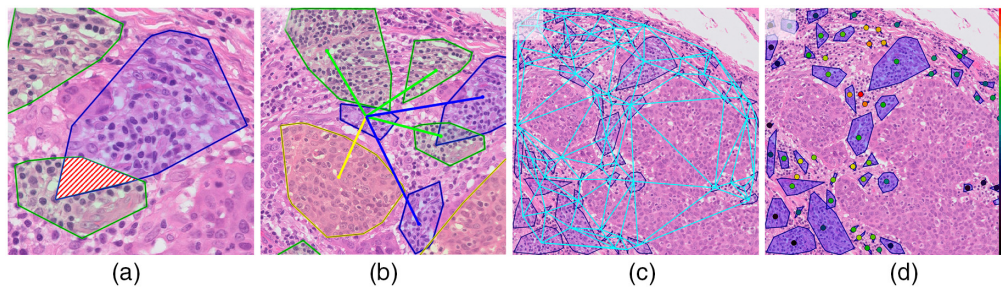


Fig. 3 Computation of some representative features: (a) intersecting area between clusters; (b) cluster neighborhood diversity; (c) Delaunay graph built for a specific nuclei type; and (d) node compactness for a specific nuclei type. The color bar represents the grouping measurement, in which H stands for nodes highly grouped, i.e., very close to multiple nodes, while L stands for sparsely clustered nodes, i.e., isolated or far from other nodes.

$\{(o_i, o_j) \in E : \forall o_i, o_j \in O, i, j \in \{1, \dots, F\}, i \neq j\}$, and $W = \{w_1, w_2, \dots, w_m\}$ is a set of weights proportional to the length of each $e \in E$. To extract information about the arrangement of cluster centroids of a nuclei type, subgraphs representing the Voronoi diagram G_V , Delaunay triangulation G_D [Fig. 3(c)], and MST G_{MST} were constructed. From such subgraphs, different features are computed, such as area, perimeter, and chord length of Voronoi polygons, area and side length of Delaunay triangles, edge length of MST, among others.

6. *Compactness of cluster centroids of each nuclei type.* Given the graph G previously defined, we computed the compactness of centroids as $c_i = \sum_{j=1}^{N_f-1} \frac{1}{D_{i,j}} * b_\theta$, where c_i is the compactness measure of the i 'th node, N_f is the number of nodes of nuclei type

f , $D_{i,j}$ is the Euclidean distance between nodes i and j , θ is a threshold value, and b_θ is a binary value that is 0 when $D_{i,j} > \theta$ and 1 otherwise [see Fig. 3(d)].

For characterizing each patch, a set of statistics (total, mean, standard deviation, median, minimum, maximum, skewness, and kurtosis) were computed for each feature.

2.4 Statistical and Visual Analysis

Violin and split violin plots were used to visualize the distribution of the spatial immune architecture features and the probability distributions associated with these features.

The t-distributed stochastic neighbor embedding (t-SNE) algorithm was used for embedding the features extracted from the spatial architecture of clusters for visualization in a two-dimensional space. This algorithm has been previously used for visualization in different studies related to single-cell transcriptomics,³⁹ human genetic data,⁴⁰ single nucleotide polymorphisms,⁴¹ and among others.

Additionally, an algorithm that maximizes the area under the receiver operating characteristics (ROC) curve was used to find the most relevant/discriminating features.⁴² Finally, spatial heat maps were built to illustrate the magnitude change of a specific feature across the patches of WSIs from COVID-19 and H1N1 patients.

3 Results

3.1 Experiment 1: Analyzing Spatial Patterns of Immune Response in COVID-19

Figure 4 presents the resulting t-SNE plot for all the patches of COVID-19 patients. No obvious grouping of patient patches was observed, suggesting that spatial architecture features are sparsely distributed and there is not a drastic difference from one patient to another.

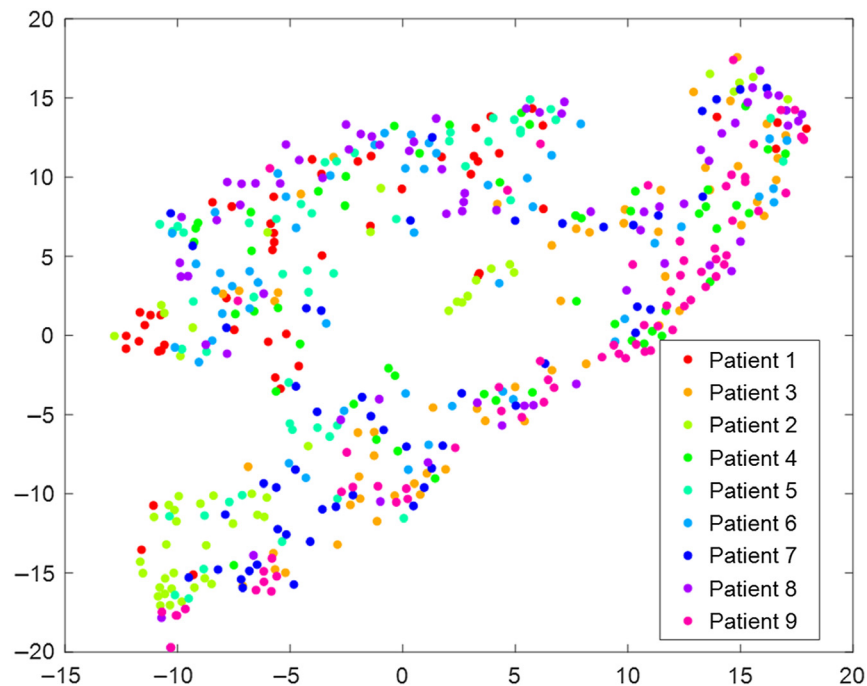


Fig. 4 Unsupervised clustering of spatial architectural features of immune clusters of image patches from COVID-19 patients. Each point represents a patch and each color represents a different patient. This plot was generated using the t-SNE algorithm that started out with 350 features and reduced the dimensionality to 2 for facilitating visualization.

Some lymphocyte-related features such as the area and density of lymphocyte clusters, intersection between clusters of lymphocytes non-lymphocytes, graph-based measures, and compactness of clusters were also analyzed using violin plots (Fig. 5). The median area of clusters of lymphocytes is around 3000 pixels ($\sim 0.33\%$ of the patch area), indicating that clusters of lymphocytes are generally small [Fig. 5(a)]. Similarly, the median intersected area between clusters of lymphocytes and non-lymphocytes was ~ 2500 pixels ($\sim 83\%$ of the median area of lymphocyte clusters), suggesting that a considerable area of the lymphocyte clusters was overlapping with non-lymphocyte clusters [Fig. 5(c)]. Finally, graph-based and compactness measures [Figs. 5(d)–5(f)] revealed that lymphocyte clusters tend to be circumscribed by other lymphocyte clusters.

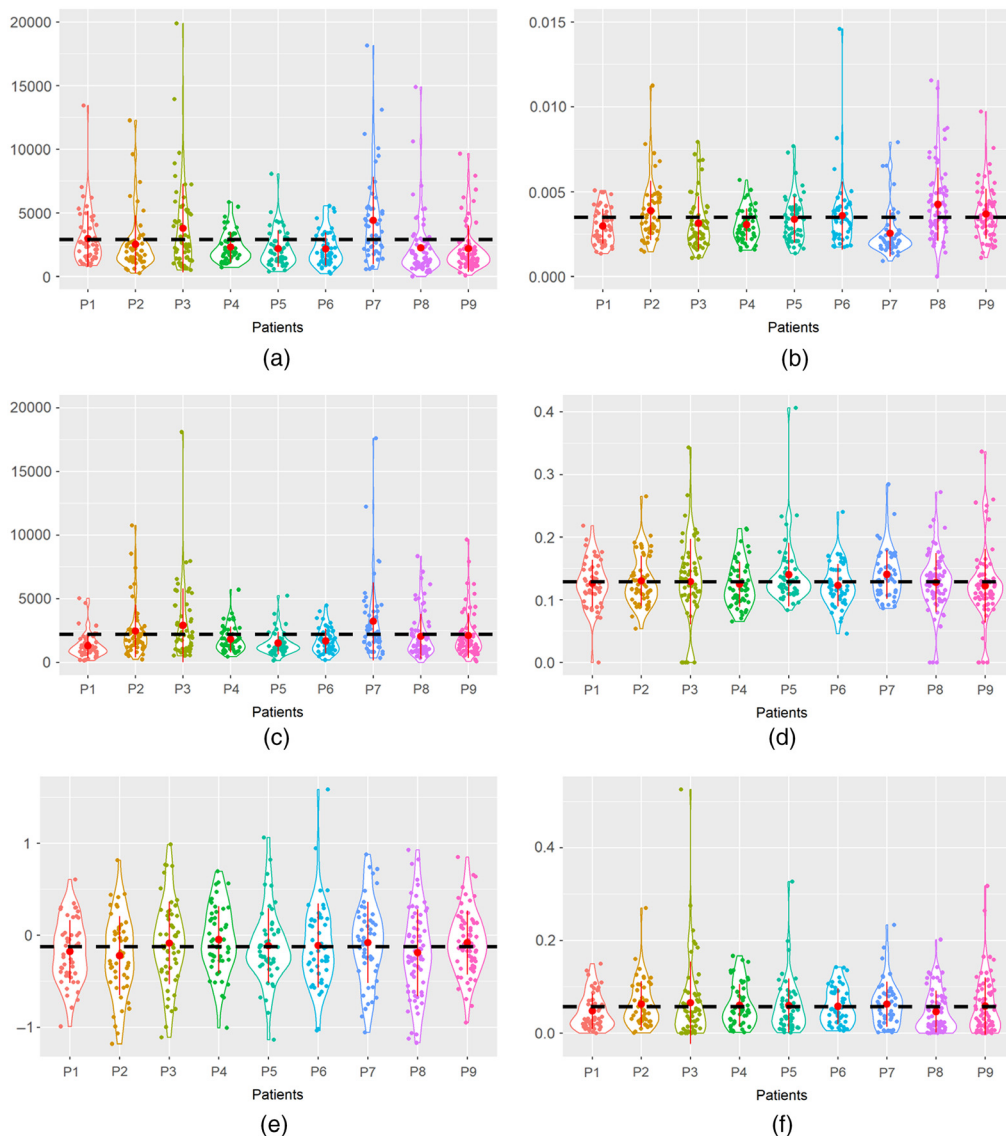


Fig. 5 Violin plots comparing the variation of some lymphocyte-related features across different COVID-19 patients ($N = 9$). For the sake of visualization, some extreme outliers were removed from subfigures (a)–(c). The dashed line indicates the mean value over all nine patients. (a) Median area (pixels) of lymphocyte clusters per patch; (b) median density of lymphocyte clusters per patch; (c) median inters. area b/w clusters of lymph. and non-lymph.; (d) side length min/max (Delaunay triang.) of lymph. clust.; (e) skewness of compactness of lymphocyte clusters; and (f) triangle area min/max (Delaunay) of lymphocyte clusters.

Table 1 Benjamin-Hochberg procedure using a FDR of 10% for the top discriminating features (as dictated by an ROC-based feature selection algorithm) and four lymphocyte-related features (quantity of lymphocytes, area of lymphocyte clusters, and compactness of cluster centroids).

Variable	p value	Rank	Adjusted p value using FDR approach
Top feat 1	2.12×10^{-25}	1	0.0071
Top feat 2	7.24×10^{-19}	2	0.0143
Top feat 5	2.65×10^{-18}	3	0.0214
Top feat 6	2.94×10^{-18}	4	0.0286
Top feat 3	4.58×10^{-16}	5	0.0357
Top feat 4	6.27×10^{-16}	6	0.0429
Top feat 7	4.48×10^{-15}	7	0.0500
Top feat 9	6.16×10^{-15}	8	0.0571
Top feat 10	2.19×10^{-14}	9	0.0643
Top feat 8	1.82×10^{-13}	10	0.0714
Lymp feat 4	1.84×10^{-5}	11	0.0786
Lymp feat 3	5.02×10^{-5}	12	0.0857
Lymp feat 2	1.44×10^{-4}	13	0.0929
Lymp feat 1	2.22×10^{-4}	14	0.1000

3.2 Experiment 2: Comparing Patterns of Immune Response in COVID-19 and H1N1

A ROC-based feature selection algorithm⁴² was employed to find the 10 most discriminating features. The Wilcoxon test was employed to determine whether differences with respect to the top features were statistically significant. This process was also carried out for four features related to lymphocytes, namely quantity of lymphocytes, area of lymphocyte clusters, and compactness of cluster centroids. Since multiple comparisons were made at the patch level and there were many features per patient, the p values corresponding to a comparison of image patches between COVID-19 versus H1N1 were adjusted using the Benjamini-Hochberg procedure using a false discovery rate (FDR) of 10% (Table 1). As all the assessed p values were smaller than the critical value, they were considered significant.

Figure 6 illustrates the top four most discriminating features, which are related to overlap between clusters of lymphocytes and non-lymphocytes. Although interpretation of these features is subtle, they indicate that very large sized clusters of non-lymphocytes are more frequently found in H1N1 compared to COVID-19. Figure 7 illustrates the four lymphocyte-related features analyzed. We can deduce that in COVID-19 there are more clusters of lymphocytes, of smaller size, that are closer to each other (high compactness) compared to H1N1, lymphocyte clusters tending to be fewer and larger.

Figure 8 shows spatial heat maps illustrating the variation of the number of non-lymphocytes per patch across 2 WSIs, one for a patient who died of COVID-19 (OCME) and another of H1N1 (CC). The density of non-lymphocytes per patch is higher for H1N1.

Finally, t-SNE was run to evaluate whether spatial-architecture-related features were able to separate H1N1 and COVID-19 patches. Figure 9(a) shows no evident clusters across the different sites, suggesting that data are similarly distributed. Figure 9(b) presents the distribution of patches according to diagnosis, exhibiting an area with a high concentration (top middle) of H1N1 points and a subtle grouping of H1N1 points (left). This could imply that some architectural patterns are more commonly found in H1N1 than in COVID-19. When looking at the architectural patterns of the studied nuclei families, in general, lymphocyte clusters are larger in

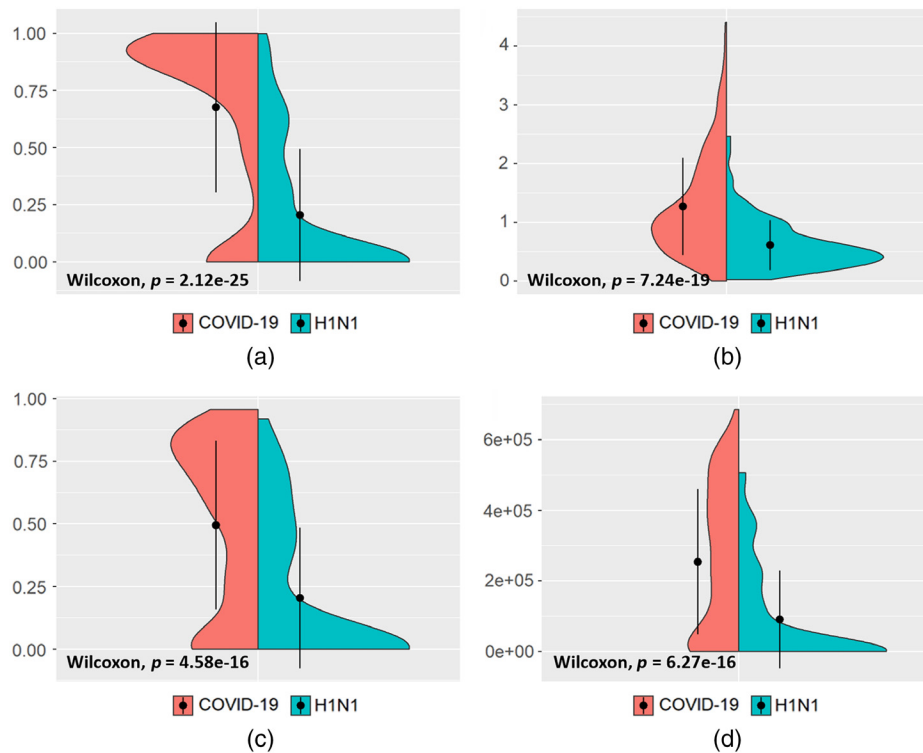


Fig. 6 Split violin plots comparing the variation of the 4 top discriminating features in patches of COVID-19 (9 patients, 496 patches) and H1N1 (2 patients, 110 patches). These features are related to the intersection of clusters of lymphocytes and non-lymphocytes and are defined as follows. (a) $\frac{AF1 \cap AF2}{AF2}$, where $AF1$ is the area of a convex hull containing the centroids of all the lymphocyte clusters and $AF2$ is the area of a convex hull containing the centroids of all the non-lymphocyte clusters. (b) $Total \cdot 2 \cdot \frac{A1 \cap A2}{A1 + A2}$, where $A1$ is the area of clusters of lymphocytes and $A2$ is the area of clusters of non-lymphocytes. (c) $2 \cdot \frac{AF1 \cap AF2}{AF1 + AF2}$. (d) $AF1 \cap AF2$.

H1N1. Similarly, non-lymphocytes are more numerous and compact, tending to generate larger clusters than COVID-19 (Fig. 9, red/yellow-border patches).

4 Discussion and Conclusion

COVID-19 causes a wide spectrum of signs and symptoms.^{43,44} In severely ill cases, some authors mentioned an exuberant cytokine production along with an intrinsic viral capacity to limit the innate immune response, which causes severe organ damage and perpetuates viral replication.^{1,2,19,45–48} Previous works have employed peripheral blood for describing the immune behavior in COVID-19 patients. The most common findings include low counts of total number of T lymphocytes and their CD4 subtype.^{15,46,49–51} These lymphocyte counts have been even proposed as predictors of outcome.^{49–51}

In tissue, visual pathologic assessment of lymphocytes in autopsied samples of patients who have died of COVID-19 has been reported. This estimation of lymphocytes is, however, time-consuming^{26,52} and has been subjectively estimated as: low-versus-medium, mild-versus-moderate, or patchy-versus-diffuse accompanying changes such as diffuse alveolar damage, acute bronchopneumonia, and microthrombi.^{6,7,10,12,53} No comparison of immune cell response among different stages in lung tissue of COVID-19 or other viral diseases has been performed to date.

To our knowledge, this is the first study in: (1) describing quantitatively immune nuclei patterns in lung tissue from patients who died of COVID-19 and (2) exploring the differences in the morphologic immune architecture of patients who succumbed to COVID-19 and H1N1, a comparable viral infection whose clinical spectrum, laboratory test results, pharmacotherapy, and prevention strategies have been extensively explored.^{15,54}

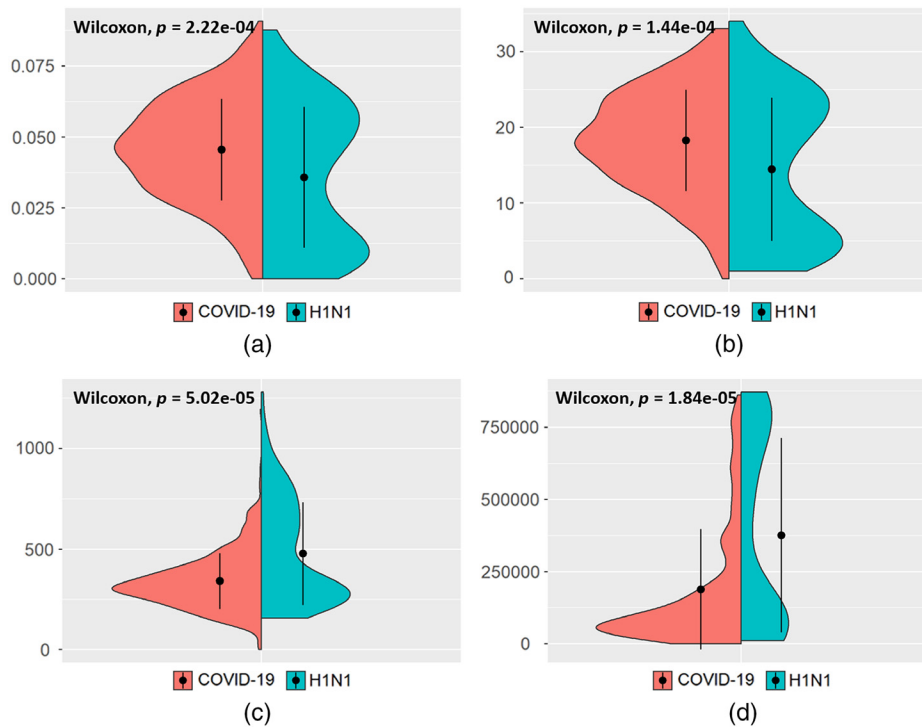


Fig. 7 Split violin plots comparing the variation of lymphocyte-related features in patches of COVID-19 (9 patients, 496 patches) and H1N1 (2 patients, 110 patches) patients. (a) Median compactness of lymphocyte clusters per patch; (b) number of clusters of lymphocytes per patch; (c) number of lymphocytes per patch; and (d) max area of lymphocyte clusters per patch.

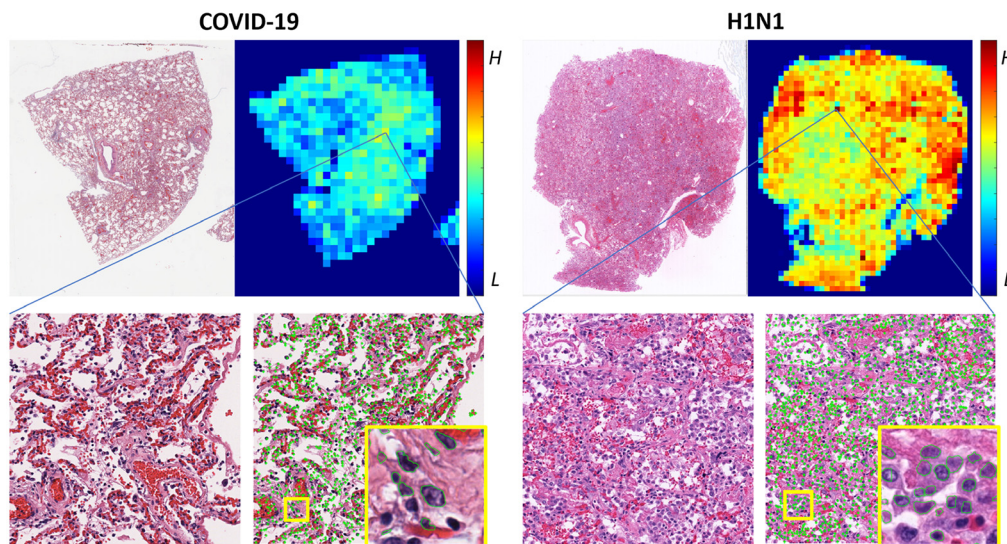


Fig. 8 Spatial heat maps illustrating the number of non-lymphocytes per patch and representative patches of a patient who died of COVID-19 and another who died of H1N1. *H* stands for a high number of non-lymphocytes and *L* for low. The patches in the bottom row show a hot spot for each disease (i.e., containing a high number of non-lymphocytes), revealing that in H1N1 there is a higher density of non-lymphocyte nuclei (green dots) compared to COVID-19.

Using an ROC-based feature selection method, violin plots, and t-SNE on patches of COVID-19 and H1N1, we found that spatial arrangement of lymphocytes per tile was different in both diseases (Figs. 5, 6, and 9). In COVID-19, we found three important patterns: (1) multiple smaller-sized lymphocyte clusters, (2) immune clusters typically surrounded by

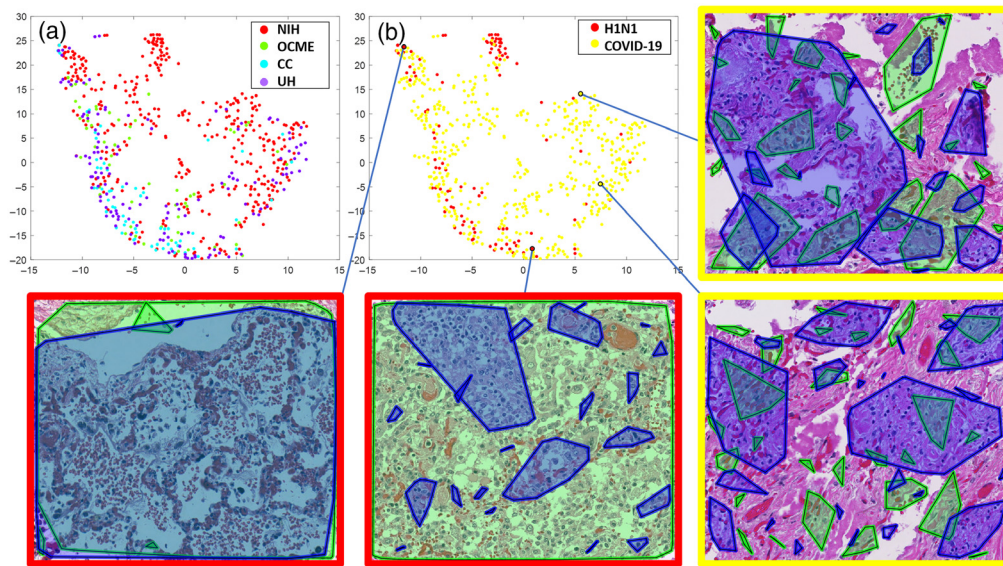


Fig. 9 Unsupervised clustering of spatial architectural features of immune clusters of COVID-19 and H1N1 patches using t-SNE and illustration of representative patches. (a) Each point represents an image patch, and each color corresponds to a different institution. The t-SNE algorithm started out with 350 features and reduced these to 2. (b) The same clustering but in this case each point represents a patch and each color a disease (i.e., either H1N1 or COVID-19). The red-border patches correspond to H1N1 and exhibit large clusters of non-lymphocytes (green border), which suggest that these nuclei are numerous and compact. The yellow-border patches correspond to COVID-19 and present as smaller groups of non-lymphocytes.

other lymphocyte clusters, and (3) regions encompassed by the lymphocyte clusters typically overlapping with other non-lymphocyte clusters. These patterns appear to suggest that the lymphocyte response in patients who died of COVID-19 is relatively modest. A possible explanation for this phenomenon is the widely reported lymphocytopenia;^{49–51} however, a more focused study needs to be performed to assess any correlation between lymphocytic infiltration in lungs and lymphocyte levels in the blood.

In H1N1 patients, by contrast, we noticed (1) fewer and larger clusters of lymphocytes and (2) significant presence of few but extremely large clusters of non-lymphocytes. This appears to be concordant with previous findings reporting a considerable quantity of macrophages accompanying the lymphocytic infiltration in H1N1, as well as a higher prevalence of acute immune cell infiltration.^{21,22} By contrast, fewer neutrophils have been found in COVID-19 cases given the potential of the *N* protein (found in SARS-CoV-2) to block IFN-1 production, a cytokine involved in neutrophil attraction.^{45,55} This imbalanced proportion of immune cells in solid tissues such as the lung is a potential explanation for the differences we found when comparing lymphocyte spatial distribution and their interaction with other nuclei in COVID-19 versus H1N1 patients.

In fatal cases of COVID-19, the inability of lymphocytes to limit viral damage despite possibly being extremely activated could be a consequence of abnormal lymphocyte–lymphocyte or lymphocyte–non-lymphocyte crosstalk, concordant with a prior hypothesis by Tan et al.⁵¹ In contrast, in H1N1, the response of non-lymphocytes appears to be flourishing, possibly reflecting either better coordination between lymphocytes and their surrounding nuclei, or a key role of superimposed acute inflammation.

We acknowledge the limitations of this work. Quantitative differences in a small number of cases might not hold up when larger numbers of cases are studied; however, it is worth mentioning that studies involving autopsies usually draw conclusions from fewer cases. Calabrese et al.⁵⁶ presented a work summarizing the pathological findings described in 23 articles reporting autopsies in suspected/known COVID-19 patients, evidencing that the median number of cases per study is 2 patients. Of course, as more cases are published, a more comprehensive analysis of the spatial architecture of the immune response for COVID-19 and H1N1 could be performed. Some studies⁶ have employed immunohistochemistry and have been able to analyze the

importance of subfamilies of lymphocytes (e.g., CD4 or CD8); however, in those works, the spatial architecture of such families was not analyzed. Our findings using standard H&E samples suggest that spatial relationships between immune and non-immune nuclei could play a role in characterizing these diseases. Identification of subgroups of lymphocytes or other immune nuclei could meaningfully enrich a similar work in the future. Future work may also benefit from including an extensive comparison against effective immune response as well as with normal lymphoid tissue associated to airway mucosa. Additionally, a comprehensive analysis of the implications of the threshold value employed to build nuclei clusters for characterizing both diseases could meaningfully enrich this work in the future.

In summary, we used a computational-pathology-based image analysis approach to characterize the immune nuclei architecture in digitized autopsied H&E lung samples from patients who died of COVID-19 and H1N1. As far as we know, this is the first attempt to quantitatively capture morphology for characterizing lymphocytic infiltrates and their relationship to surrounding nuclei in COVID-19 and simultaneously compare this arrangement in H1N1. In this preliminary study, we found that the immune response in COVID-19 consistently showed multiple lymphocyte clusters of small size, suggesting that the lymphocyte response is rather modest, possibly due to lymphocytopenia. In contrast, in H1N1, we found larger lymphocyte clusters accompanied by large clusters of non-lymphocytes, a possible reflection of increased prevalence of macrophages, and other immune cells. In combination with cytokine profiling⁵⁷ and after testing this algorithm in a larger population, our approach could potentially help enrich and inform understanding in the process of developing therapeutic involving immune response modulation in COVID-19 patients.

Disclosures

Dr. Madabhushi is an equity holder in Elucid Bioimaging and in Inspirata Inc. In addition, he has served as a scientific advisory board member for Inspirata Inc., Astrazeneca, Bristol Meyers-Squibb, and Merck. Currently, he serves on the advisory board of Aiforia Inc. He also has sponsored research agreements with Philips and Bristol Meyers-Squibb. His technology has been licensed to Elucid Bioimaging. He is also involved in an NIH U24 grant with PathCore Inc., and three different R01 grants with Inspirata Inc. No potential conflicts of interest were disclosed by the other authors.

Acknowledgments

Research reported in this publication was supported by the National Cancer Institute of the National Institutes of Health (Award Nos. 1U24CA199374-01, R01CA249992-01A1, R01CA202752-01A1, R01CA208236-01A1, R01CA216579-01A1, R01CA220581-01A1, R01CA257612-01A1, 1U01CA239055-01, 1U01CA248226-01, and 1U54CA254566-01); the National Institute for Biomedical Imaging and Bioengineering (No. 1R43EB028736-01); the National Heart, Lung, and Blood Institute (No. R01HL151277-01A1); the National Center for Research Resources under Award No. 1 C06 RR12463-01; the VA Merit Review Award No. IBX004121A from the United States Department of Veterans Affairs Biomedical Laboratory Research and Development Service; the Office of the Assistant Secretary of Defense for Health Affairs, through the Breast Cancer Research Program (W81XWH-19-1-0668), the Prostate Cancer Research Program (W81XWH-15-1-0558 and W81XWH-20-1-0851), the Lung Cancer Research Program (W81XWH-18-1-0440 and W81XWH-20-1-0595), and the Peer Reviewed Cancer Research Program (W81XWH-18-1-0404); the Wallace H. Coulter Foundation Program in the Department of Biomedical Engineering and the Clinical and Translational Science Award Program (CTSA) at Case Western Reserve University; the Kidney Precision Medicine Project (KPMP) Glue Grant; the Ohio Third Frontier Technology Validation Fund; the Clinical and Translational Science Collaborative of Cleveland (UL1TR0002548) from the National Center for Advancing Translational Sciences (NCATS) component of the National Institutes of Health (NIH) and NIH Roadmap for Medical Research; and sponsored research agreements from Bristol Myers-Squibb, Boehringer-Ingelheim, and Astrazeneca. The content is solely the responsibility of the authors and does not necessarily represent the official views

of the National Institutes of Health, the U.S. Department of Veterans Affairs, the Department of Defense, or the United States Government.

Code, Data, and Materials Availability

Part of the images employed for this study are publicly available at <https://halolink.covid19pathology.nih.gov>. The remaining images are available from Cleveland Clinic and University Hospitals upon request.

References

1. L. F. García, “Immune response, inflammation, and the clinical spectrum of COVID-19,” *Front Immunol* **11**, 1441 (2020).
2. F. Yazdanpanah, M. R. Hamblin, and N. Rezaei, “The immune system and COVID-19: Friend or foe?” *Life Sci.* **256**(32502542), 117900 (2020).
3. D. D. Chaplin, “Overview of the immune response,” *J. Allergy Clin. Immunol.* **125**(2), S3–S23 (2010).
4. A. D. Luster, “The role of chemokines in linking innate and adaptive immunity,” *Curr. Opin. Immunol.* **14**(1), 129–135 (2002).
5. P. W. Harms et al., “Autopsy findings in eight patients with fatal H1N1 influenza,” *Am. J. Clin. Pathol.* **134**(1), 27–35 (2010).
6. S. E. Fox et al., “Pulmonary and cardiac pathology in African American patients with COVID-19: an autopsy series from New Orleans,” *Lancet Respir. Med.* **8**(7), 681–686 (2020).
7. L. M. Barton et al., “COVID-19 autopsies, Oklahoma, USA,” *Am. J. Clin. Pathol.* **153**(6), 725–733 (2020).
8. S. Tian et al., “Pulmonary pathology of early-phase 2019 novel coronavirus (COVID-19) pneumonia in two patients with lung cancer,” *J. Thorac. Oncol.* **15**(5), 700–704 (2020).
9. S. Tian et al., “Pathological study of the 2019 novel coronavirus disease (COVID-19) through postmortem core biopsies,” *Mod. Pathol.* **33**(6), 1007–1014 (2020).
10. D. Wichmann et al., “Autopsy findings and venous thromboembolism in patients with COVID-19,” *Ann. Intern. Med.* **173**(4), 268–277 (2020).
11. S. F. Lax et al., “Pulmonary arterial thrombosis in COVID-19 with fatal outcome: results from a prospective, single-center, clinicopathologic case series,” *Ann. Intern. Med.* (32422076), M20-2566 (2020).
12. T. Menter et al., “Post-mortem examination of COVID-19 patients reveals diffuse alveolar damage with severe capillary congestion and variegated findings of lungs and other organs suggesting vascular dysfunction,” *Histopathology* **77**(2), 198–209 (2020).
13. K. E. Konopka et al., “Diffuse alveolar damage (DAD) from coronavirus disease 2019 infection is morphologically indistinguishable from other causes of DAD,” *Histopathology* **77**(4), 570–578 (2020).
14. S. B. Polak et al., “A systematic review of pathological findings in COVID-19: a pathophysiological timeline and possible mechanisms of disease progression,” *Mod. Pathol.* **33**, 2128–2138 (2020).
15. X. Tang et al., “Comparison of hospitalized patients with ARDS caused by COVID-19 and H1N1,” *Chest* **158**(1), 195–205 (2020).
16. A. M. Foust et al., “Pediatric SARS, H1N1, MERS, EVALI, and now coronavirus disease (COVID-19) pneumonia: what radiologists need to know,” *Am. J. Roentgenol.* **215**(3), 736–744 (2020).
17. Z. Yin et al., “A comparison of clinical and chest CT findings in patients with influenza A (H1N1) virus infection and coronavirus disease (COVID-19),” *Am. J. Roentgenol.* **215**(5), 1065–1071 (2020).
18. F. Zhou et al., “Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study,” *Lancet* **395**(10229), 1054–1062 (2020).
19. X. Li et al., “Molecular immune pathogenesis and diagnosis of COVID-19,” *J. Pharm. Anal.* **10**(2), 102–108 (2020).

20. T. Mauad et al., "Lung pathology in fatal novel human influenza A (H1N1) infection," *Am. J. Respir. Crit. Care Med.* **181**(1), 72–79 (2010).
21. S. Vangeti, M. Yu, and A. Smed-Sørensen, "Respiratory mononuclear phagocytes in human influenza A virus infection: their role in immune protection and as targets of the virus," *Front. Immunol.* **9**(30018617), 1521–1521 (2018).
22. J. R. Gill et al., "Pulmonary pathologic findings of fatal 2009 pandemic influenza A/H1N1 viral infections," *Arch. Pathol. Lab. Med.* **134**(2), 235–243 (2010).
23. N. Nakajima et al., "Histopathological and immunohistochemical findings of 20 autopsy cases with 2009 H1N1 virus infection," *Mod. Pathol.* **25**(1), 1–13 (2012).
24. J. Saltz et al., "Spatial organization and molecular correlation of tumor-infiltrating lymphocytes using deep learning on pathology images," *Cell Rep.* **23**(1), 181–193.e7 (2018).
25. K. AbdulJabbar et al., "Geospatial immune variability illuminates differential evolution of lung adenocarcinoma," *Nat. Med.* **26**, 1054–1062 (2020).
26. G. Corredor et al., "Spatial architecture and arrangement of tumor-infiltrating lymphocytes for predicting likelihood of recurrence in early-stage non-small cell lung cancer," *Clin. Cancer Res.* **25**(5), 1526–1534 (2019).
27. A. N. Basavanhally et al., "Computerized image-based detection and grading of lymphocytic infiltration in HER2+ breast cancer histopathology," *IEEE Trans. Biomed. Eng.* **57**(3), 642–653 (2010).
28. J. S. Lewis et al., "A quantitative histomorphometric classifier (QuHbIC) identifies aggressive versus indolent p16-positive oropharyngeal squamous cell carcinoma," *Am. J. Surg. Pathol.* **38**(1), 128–137 (2014).
29. <https://halolink.covid19pathology.nih.gov>.
30. A. M. Khan and Y. Yuan, "Biopsy variability of lymphocytic infiltration in breast cancer subtypes and the ImmunoSkew score," *Sci. Rep.* **6**(1), 36231 (2016).
31. M. Macenko et al., "A method for normalizing histology slides for quantitative analysis," in *IEEE Int. Symp. Biomed. Imaging: From Nano to Macro*, pp. 1107–1110 (2009).
32. M. Veta et al., "Automatic nuclei segmentation in H&E stained breast cancer histopathology images," *PLoS One* **8**(7), e70221 (2013).
33. C. Lu et al., "An oral cavity squamous cell carcinoma quantitative histomorphometric-based image classifier of nuclear morphology can risk stratify patients for disease-specific survival," *Mod. Pathol.* **30**(12), 1655–1665 (2017).
34. X. Wang et al., "Prediction of recurrence in early stage non-small cell lung cancer using computer extracted nuclear features from digital H&E images," *Sci. Rep.* **7**(1), 13543 (2017).
35. G. Corredor et al., "A watershed and feature-based approach for automated detection of lymphocytes on lung cancer images," *Proc. SPIE* **10518**, 105810R (2018).
36. G. Swart, "Finding the convex hull facet by facet," *J. Algorithms* **6**(1), 17–48 (1985).
37. B. Cowled et al., "The equine influenza epidemic in Australia: Spatial and temporal descriptive analyses of a large propagating epidemic," *Prev. Vet. Med.* **92**(1–2), 60–70 (2009).
38. E. Dumonteil et al., "Spatial extent of an outbreak in animal epidemics," *Proc. Natl. Acad. Sci. U. S. A.* **110**(11), 4239–4244 (2013).
39. D. Kobak and P. Berens, "The art of using t-SNE for single-cell transcriptomics," *Nat. Commun.* **10**(1), 5416 (2019).
40. W. Li et al., "Application of t-SNE to human genetic data," *J. Bioinform. Comput. Biol.* **15**(04), 1750017 (2017).
41. A. Platzer, "Visualization of SNPs with t-SNE," *PLoS One* **8**(2), e56883 (2013).
42. S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, 4th ed., Elsevier Academic Press, Amsterdam (2009).
43. W. Guan et al., "Clinical characteristics of coronavirus disease 2019 in China," *N. Engl. J. Med.* **382**(18), 1708–1720 (2020).
44. Y. Liu et al., "Viral dynamics in mild and severe cases of COVID-19," *Lancet Infect. Dis.* **20**(6), 656–657 (2020).
45. D. Blanco-Melo et al., "Imbalanced host response to SARS-CoV-2 drives development of COVID-19," *Cell* **181**(5), 1036–1045 (2020).

46. X. Zhang et al., “Viral and host factors related to the clinical outcome of COVID-19,” *Nature* **583**(7816), 437–440 (2020).
47. W. Zhang et al., “The use of anti-inflammatory drugs in the treatment of people with severe coronavirus disease 2019 (COVID-19): the perspectives of clinical immunologists from China,” *Clin Immunol* **214**, 108393 (2020).
48. Y. Shi et al., “COVID-19 infection: the perspectives on immune responses,” *Cell Death Differ.* **27**(5), 1451–1454 (2020).
49. Q. Zhao et al., “Lymphopenia is associated with severe coronavirus disease 2019 (COVID-19) infections: a systemic review and meta-analysis,” *Int. J. Infect. Dis.* **96**, 131–135 (2020).
50. B. M. Henry, “COVID-19, ECMO, and lymphopenia: a word of caution,” *Lancet Respir. Med.* **8**(4), e24 (2020).
51. L. Tan et al., “Lymphopenia predicts disease severity of COVID-19: a descriptive and predictive study,” *Signal Transduct. Target Ther.* **5**(1), 33 (2020).
52. K. A. Schalper et al., “Objective measurement and clinical significance of TILs in non-small cell lung cancer,” *J. Natl. Cancer Inst.* **107**(3), dju435 (2015).
53. J. L. Sauter et al., “Insights into pathogenesis of fatal COVID-19 pneumonia from histopathology with immunohistochemical and viral RNA studies,” *Histopathology* **77**(6), 915–925 (2020).
54. R. Jhaveri, “Echoes of 2009 H1N1 influenza pandemic in the COVID pandemic,” *Clin. Ther.* **42**(5), 736–740 (2020).
55. H. Chu et al., “Comparative replication and immune activation profiles of SARS-CoV-2 and SARS-CoV in human lungs: an ex vivo study with implications for the pathogenesis of COVID-19,” *Clin. Infect. Dis.* **71**, 1400–1409 (2020).
56. F. Calabrese et al., “Pulmonary pathology and COVID-19: lessons from autopsy. The experience of European pulmonary pathologists,” *Virchows Arch.* **477**(3), 359–372 (2020).
57. R. M. Boehler, J. G. Graham, and L. D. Shea, “Tissue engineering tools for modulation of the immune response,” *BioTech.* **51**(4), 239–254 (2011).

Germán Corredor received his BS, MS, and PhD degrees from the Universidad Nacional de Colombia in 2010, 2012, and 2019, respectively. He is a research associate at the Case Western Reserve University and the Louis Stokes VA Medical Center, Cleveland, Ohio, USA. His current research interests include the development of strategies for predicting patient diagnosis and prognosis using histopathological samples. He is also interested in machine learning, computer vision, software engineering, and medical information systems.

Paula Toro received her MD from the Universidad de Caldas in 2014, and she graduated as an anatomic and clinical pathologist from the Universidad Nacional de Colombia in 2020. She is a research coordinator with focus on pathology at the Case Western Reserve University. Her current professional interests are digital pathology, dermatopathology, and pediatric pathology.

Vidya Sankar Viswanathan, MD, is a post-doctoral fellow at the Center for Computational Imaging and Personalized Diagnostics (CCIPD), Department of Biomedical Engineering at Case Western Reserve University. He completed his medical education from the University College of Medical Sciences, Delhi, India, and graduated from the University of Delhi in 2019. As the research coordinator for CCIPD, he is the clinical lead for radiomics projects and works closely with collaborators in providing clinical domain inputs for the projects.

Christina Buzzy is the director of research operations of the Center for Computational Imaging and Personalized Diagnostics in the Department of Biomedical Engineering at Case Western Reserve University. She earned her PhD in pathology where she delineated a novel inflammatory pathway for interleukin-1 beta in immune cells activated by chemotherapeutic drugs. She has focused her early career on developing a translational clinical research knowledge base, serving as director of urologic oncology research at University Hospitals for three years and two years as assistant director of research programs at the NCI-designated Case Comprehensive Cancer Center.

Sanjay Mukhopadhyay is director of pulmonary pathology at the Cleveland Clinic and an associate editor (pulmonary) of the *American Journal of Clinical Pathology*. He attended medical school in India and completed pathology residency training at the All India Institute of Medical Sciences followed by a second pathology residency at SUNY Upstate Medical University, Syracuse, NY, USA. After a pulmonary pathology fellowship at the Mayo Clinic, Rochester, MN, USA, he was recruited back to Syracuse and stayed on as faculty for next six years, rising to the rank of tenured associate professor in 2011. In 2013, he joined the Department of Pathology at the Cleveland Clinic. His interests include granulomatous lung diseases (including infections), interstitial lung disease, and lung cancer.

Anant Madabhushi is director of the Center for Computational Imaging and Personalized Diagnostics and Donnell Institute professor, Department of Biomedical Engineering at the Case Western Reserve University. He is a research health scientist at Louis Stokes of Cleveland Veterans Administration Medical Center. He has authored more than 400 peer-reviewed publications and more than 100 patents issued or pending. He is a fellow of the American Institute of Medical and Biological Engineering, the Institute for Electrical and Electronic Engineers, and the National Academy of Inventors. His work on “Imaging computers for identifying lung cancer patients who need chemotherapy” was called out by *Prevention* magazine as one of the top 10 medical breakthroughs of 2018. In 2019, *Nature* hailed him as one of the five scientists developing “offbeat and innovative approaches for cancer research.” He was named to the pathologist’s Power List in 2019 and 2020.

Biographies of the other authors are not available.