AMERICAN COLLEGE
of RHEUMATOLOGY
Empowering Rheumatology Professionals

Rheumatoid Arthritis Synovial Inflammation Quantification Using Computer Vision

Steven Guan,[1] [iD] Bella Mehta,[2,3] David Slater,[1] James R. Thompson,[1] Edward DiCarlo,[2] Tania Pannellini,[2] Diyu Pearce-Fisher,[2] Fan Zhang,[4,5,6,7,8] Soumya Raychaudhuri,[4,5,6,7,8,9] Caryn Hale,[10] Caroline S. Jiang,[10] Susan Goodman,[2,3] [iD] and Dana E. Orange[2,10] [iD]

Objective. We quantified inflammatory burden in rheumatoid arthritis (RA) synovial tissue by using computer vision to automate the process of counting individual nuclei in hematoxylin and eosin images.

Methods. We adapted and applied computer vision algorithms to quantify nuclei density (count of nuclei per unit area of tissue) on synovial tissue from arthroplasty samples. A pathologist validated algorithm results by labeling nuclei in synovial images that were mislabeled or missed by the algorithm. Nuclei density was compared with other measures of RA inflammation such as semiquantitative histology scores, gene-expression data, and clinical measures of disease activity.

Results. The algorithm detected a median of 112,657 (range 8,160-821,717) nuclei per synovial sample. Based on pathologist-validated results, the sensitivity and specificity of the algorithm was 97% and 100%, respectively. The mean nuclei density calculated by the algorithm was significantly higher ($P < 0.05$) in synovium with increased histology scores for lymphocytic inflammation, plasma cells, and lining hyperplasia. Analysis of RNA sequencing identified 915 significantly differentially expressed genes in correlation with nuclei density (false discovery rate is less than 0.05). Mean nuclei density was significantly higher ($P < 0.05$) in patients with elevated levels of C-reactive protein, erythrocyte sedimentation rate, rheumatoid factor, and cyclized citrullinated protein antibody.

Conclusion. Nuclei density is a robust measurement of inflammatory burden in RA and correlates with multiple orthogonal measurements of inflammation.

INTRODUCTION

Rheumatoid arthritis (RA), the most prevalent autoimmune arthritis, is characterized by inflammation in synovial tissue. There is great interest in using synovial tissue samples to optimize treatment decisions for patients with RA in the long term (1–3) and in the short term; rigorous histologic assessments of synovial tissue are needed as benchmarks for molecular approaches such as

single cell RNA sequencing (RNA-seq) (4). Many groups have compared and validated various methods of quantifying antibody-stained cells such as T cells, B cells, plasma cells, and macrophages in RA synovium (5). Three general methods of measurement include semiquantitative scoring (whereby a pathologist generates an ordinal summary score ranging from 0 to 4 to describe the level of infiltrate), manual counting of individual cells (which is laborious and not well suited for large-scale studies),

[1]Steven Guan, PhD, David Slater, PhD, James R. Thompson, PhD: The MITRE Corporation, McLean, Virginia; [2]Bella Mehta, MBBS, MS, Edward DiCarlo, MD, Tania Pannellini, MD, PhD, Diyu Pearce-Fisher, BS, Susan Goodman, MD, Dana E. Orange, MD, MS: Hospital for Special Surgery, New York, New York; [3]Bella Mehta, MBBS, MS, Susan Goodman, MD: Weill Cornell Medicine, New York, New York; [4]Fan Zhang, PhD, Soumya Raychaudhuri, MD, PhD: Center for Data Sciences, Brigham and Women's Hospital, Boston, Massachusetts; [5]Fan Zhang, PhD, Soumya Raychaudhuri, MD, PhD: Division of Genetics, Department of Medicine, Brigham and Women's Hospital, Boston, Massachusetts; [6]Fan Zhang, PhD, Soumya Raychaudhuri, MD, PhD: Department of Biomedical Informatics, Harvard

Medical School, Boston, Massachusetts; [7]Fan Zhang, PhD, Soumya Raychaudhuri, MD, PhD: Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, Massachusetts; [8]Fan Zhang, PhD, Soumya Raychaudhuri, MD, PhD: Division of Rheumatology, Inflammation and Immunity, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts; [9]Soumya Raychaudhuri, MD, PhD: Centre for Genetics and Genomics Versus Arthritis, Manchester Academic Health Science Centre, University of Manchester, Manchester, UK; [10]Caryn Hale, PhD, Caroline S. Jiang, MS, Dana E. Orange, MD, MS: Rockefeller University, New York, New York.

No potential conflicts of interest relevant to this article were reported.

and digital image assessment (whereby computer software is used to quantify cell types). Although ordinal semiquantitative scores are considered valid, the more granular continuous data generated from manual counts and digital image analysis are more sensitive for discriminating clinically involved and uninvolved joints (6). However, these approaches are currently limited by the need for special antibody stains and field selection bias. Careful studies comparing digital image analysis of either 60, 20, 6, or 1 high-power fields (representing 7.5 mm^2, 2.5 mm^2, 0.75 mm^2, and 0.125 mm^2 of tissue, respectively) found that the variance increases when the number of high-power fields drops from 20 to 6 (7).

Hematoxylin and eosin (H&E) stain is the most widely used method for assessing any type of tissue. Some groups, including our group, have also validated semiquantitative assessments of H&E-stained synovium in RA (8,9). A key limitation of these semiquantitative methods is that they rely on the subjective selection of a subset of high-power fields from the whole slide image. Given that synovial inflammation tends to be heterogeneous, variability in the selected high-power fields for assessment can lead to different scores. In a recent study from our group, the weighted kappa for interrater reliability between two highly experienced pathologists assessing synovial lymphocyte infiltration on the same slides was only 67% (9).

Advances in computational pathology have been applied with great success in oncology to predict cancer outcomes and treatment response (10–13). In general, computational pathology seeks to extract meaningful information from histology samples such as detecting, segmenting, and characterizing key features. By automating these fundamental tasks, the entire sample can be efficiently analyzed, and the burden on the pathologist and other human biases in selecting a subset of high-power fields from the whole slide image can be reduced (14). However, these computer vision approaches have not yet been widely adopted for H&E-stained synovial tissue analysis. Considering that hospitals are beginning to incorporate digital pathology images of whole slides into medical records (15), and our site alone performs over 5,000 arthroplasties per year, there is an opportunity to apply computer vision at an enterprise scale to enable precise, rapid, and cost-effective quantification of inflammation in synovial tissue samples.

In recent years, deep learning methods, such as convolutional neural networks, have been successfully applied for quantifying histological features and phenotypes in H&E-stained tissue samples (16,17). Deep learning is a powerful tool that often outperforms traditional computer vision methods. However, the main drawback of data-driven methods, like deep learning, is the need for large datasets with labeled annotations. Acquiring a sufficiently large training dataset is often prohibitively time consuming and expensive. Therefore, labeled data are often unavailable, and conventional computer vision approaches, such as our proposed pipeline, are needed.

Synovial samples with dense inflammatory infiltrates typically have more nuclei (18). We hypothesized that the density of nuclei in H&E-stained synovial tissue samples can be used as a simple, fast, quantitative measurement of inflammation. Here, we share an algorithm (Figure 1) using computer vision techniques including color deconvolution, local adaptive thresholding, watershed segmentation, and shape filtering. We compare the output of computer vision assessments of RA synovial nuclei density, which yield hundreds of thousands of nuclei per whole slide image, with semiquantitative histology scores, bulk and single cell RNA-seq, and clinical measures of disease activity. These comparisons provide proof of concept that nuclei density is a robust measurement of synovial inflammation, thereby laying the groundwork for advancing computer vision in rheumatic disease.

METHODS

Patient data. This study includes data from 170 patients with RA undergoing arthroplasty at the Hospital for Special Surgery (HSS) in New York. All patients met either the American College of Rheumatology (ACR)/European League Against Rheumatism 2010 classification criteria and/or the ACR 1987 criteria for RA (19,20). Patient data, including age, duration of disease, body mass index, tender and swollen joint counts, Disease Activity Score in 28 joints (DAS28), erythrocyte sedimentation rate (ESR), C-reactive protein (CRP), rheumatoid factor (RF), and anti-cyclic citrullinated peptide (CCP), were collected (9). This study was approved by the HSS Institutional Review Board (approval number 2014-233), the Rockefeller University Institutional Review Board (approval number DOR0822), and the Biomedical Research Alliance of New York (approval number 15-08-114-385). All participating patients provided their signed informed consent. The patients and the public were not involved in the design of this study.

This study was approved by the HSS Institutional Review Board (approval no. 2014-233), the Rockefeller University Institutional Review Board (approval no. DOR0822), and the Biomedical Research Alliance of New York (approval no. 15-08-114-385). All participating patients provided their signed informed consent.

Tissue quantification. Whole slide images were split into a set of smaller 1,024 × 1,024 pixel-sized images, termed "tiles," representing an area of 0.25 mm^2 (Figure 1A). To equitably compare nuclei counts between tiles, it is necessary to quantify and normalize for tissue. The RBG (Red Blue Green) image was transformed into a grayscale image such that each pixel contains an intensity value between 0 *and* 255. Next, we employ*ed* an intensity value threshold of 30 to annotate pixels above this value as tissue. Finally, a tissue metric was calculated based on the proportion of the tile determined to contain tissue.

Color deconvolution and local adaptive thresholding. Hematoxylin (blue) preferentially stains cell nuclei, whereas eosin (magenta-red) stains cytoplasm (21). To separate the nuclei from the cytoplasm, color deconvolution was applied to transform
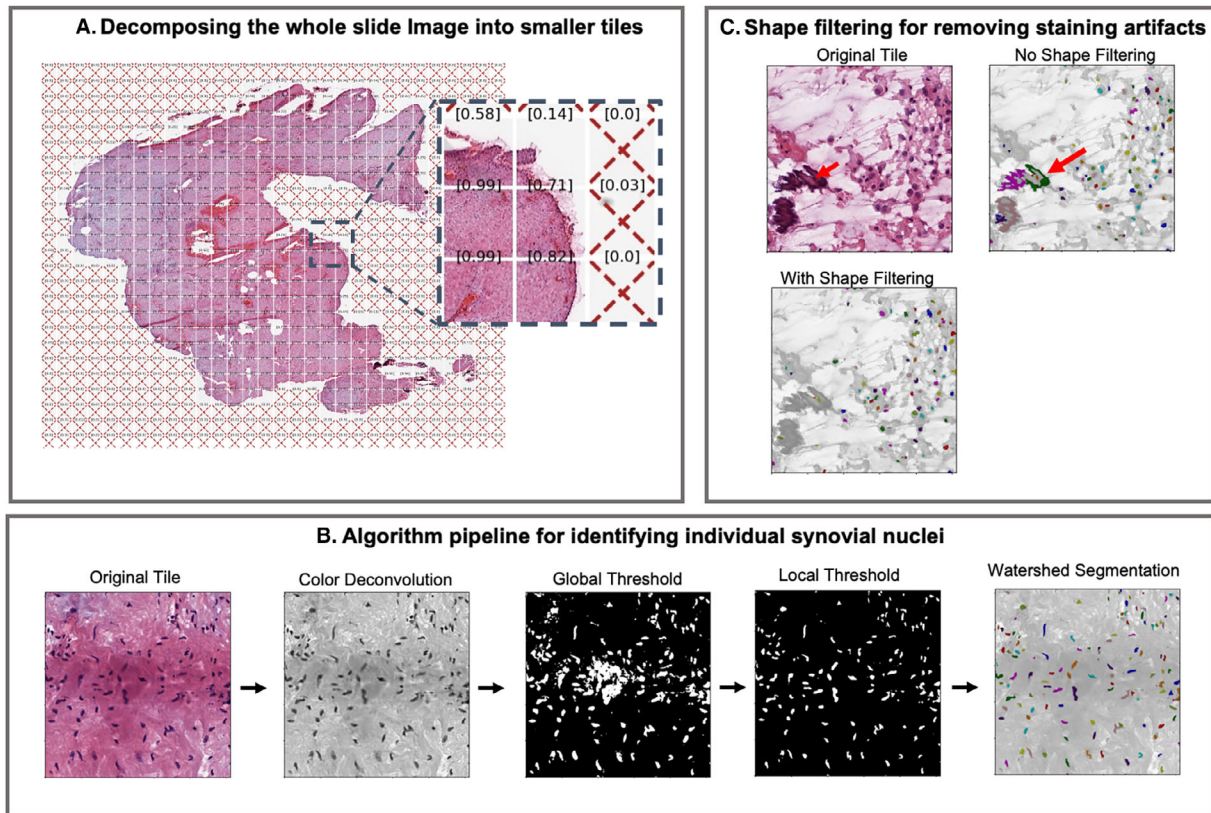
Figure 1. Algorithm pipeline for automated quantification of synovial nuclei. **A,** The whole slide image is decomposed into 0.25 mm^2 tiles with an isotropic resolution of 0.5 μm. Numbers in brackets indicate the amount of tissue identified by the algorithm in the tile (between 0 and 1). Tiles with a tissue metric less than 0.10 are excluded from further analysis. **B,** Visualization of the algorithm's image processing pipeline to identify synovial nuclei. This process is repeated for all tiles meeting the tissue threshold. 1) Original colored tile with H&E staining, 2) grayscale image representing the intensity of hematoxylin staining from color deconvolution, 3) global threshold image provides a rough estimate of nuclei locations, 4) local threshold image provides a refined estimate of nuclei locations, and 5) individual nuclei are found using the watershed segmented image. **C,** Shape filtering removes irregularly shaped artifacts as indicated by the red arrow that would be otherwise mistakenly identified as synovial nuclei.

the color image into grayscale images representing the concentration of each dye (22). Color deconvolution was performed using RBG channel weighting vectors of (0.65, 0.70, 0.29) for hematoxylin and (0.07, 0.99, 0.11) for eosin. The resulting grayscale images have normalized pixel intensities ranging from 0 to 255 (Figure 1B).

Using the hematoxylin image, it is easier to separate cell nuclei from the background. Pixels in the image were classified as background (eg, empty space, cytoplasm, and stroma) or foreground (ie, nuclei) based on a threshold. Otsu's method was used to automate selecting a threshold to separate the nuclei from the background (23). Due to the heterogeneous staining intensity, local adaptive thresholding was used to select threshold values for different regions of an image. First, a global threshold was applied resulting in a binary image with a rough estimate for the locations of nuclei (Figure 1B). Local regions were then defined based on an enlarged bounding box for each connected component in the global binary image. Local threshold values were then computed and applied to each local region to provide a refined estimate for the locations of nuclei.

Nuclei segmentation and shape filtering. In regions with densely clustered nuclei, an object in the binary image can contain multiple adjacent or overlapping nuclei. Nuclei segmentation was thus performed using the watershed method to find boundary lines between individual nuclei (24). We applied the distance transform to the binary image resulting in a map that represents the shortest Euclidean distance between the foreground and background (25). The distance map and the hematoxylin grayscale image were normalized and summed together (18). The watershed method was then applied to the combined distance and image intensity map to identify individual nuclei.

Staining artifacts incorrectly identified as nuclei were removed by filtering objects based on their shape with metrics like circularity and Euler number (the total number of objects minus the total number of holes in an object). Objects with a circularity less than 0.4 and a Euler number less than 0 were removed (Figure 1C). Nuclei density was calculated by normalizing the total count of individual nuclei by the tissue metric. Whole slide image-level mean nuclei densities were also calculated for each patient by averaging all tile-level nuclei densities.

Pathologist validation of nuclei detected. Nuclei identified by the algorithm were validated by a pathologist on a sample of 10 tiles (two tiles from each lymphocyte grade). We used the open-source software QuPath to annotate nuclei that were either mislabeled or missing (26). Algorithm performance was evaluated by calculating sensitivity, specificity, and the F1 score.

Histologic scoring. Given that the whole synovium is available for direct visual inspection after arthroplasty, the pathologist selected the area that appeared most grossly inflamed. If there were no obvious niduses of inflammation, synovial samples were preferentially obtained from areas of inflamed synovium or standard locations as previously described (9). Histology features were scored by a pathologist using a systematic approach described in (9) and is also available at www.hss.edu/pathology-synovitis. Chi-squared tests were conducted to determine whether there were significant relationships among nuclei density and the histology features. The Spearman correlation coefficient was calculated for features with a significant relationship.

Gene expression analysis. RNA was extracted from paired bulk synovial tissue samples and previously sequenced as described in (Immport Accession #SDY1299) (9). In brief, these libraries were prepared using TruSeq messenger RNA (mRNA) Stranded Library kits; 50-bp paired-end reads were sequenced on a HiSeq2500 platform, and reads were aligned to hg19 using STAR (27). Samples with more than 0.1% globin mRNA were excluded from further analysis. Consensus clustering identified three gene expression clusters (low, mixed, and high inflammatory). A principal component analysis was performed on the top 5,000 most variably expressed genes. The limma R package was used to identify differentially expressed genes using mean nuclei count as a continuous variable (28). Differentially expressed genes were defined as those with a false discovery rate (FDR) less than 0.05.

An enrichment analysis of 18 synovial cell type subsets (200 marker genes per cell) was performed (Immport Accession #SDY998) (4). Genes were ranked according to the moderated t-statistic, and the fgsea R package was used for enrichment analysis using the marker genes as gene sets (28,29).

RESULTS

Clinical characteristics. Among the 170 patients with RA, 81% were female, 41% were seropositive for RF, and 64% were seropositive for anti-CCP. Disease activity was moderate to high in 70% of patients, with 50% DASs of 3.2 or more and 17% DASs of 5.1 or more. Patient characteristics are presented in Supplemental Table 1.

Automated quantification of synovial nuclei. Our algorithm was used to identify synovial nuclei in 170 whole slide images of patients with RA and detected a median of 112,657 (range 8,160 to 821,717) nuclei per whole slide image. This large range of nuclei detected was partially due to the varying synovial tissue sample sizes. The whole slide images were sectioned into a median of 579 (range 108 to 1,850) tiles. After tissue quantification and removal of tiles containing mostly empty space, the whole slide images had a remaining median of 249 (range 29 to 1,103) tiles, corresponding to a median tissue area of 1.25 cm^2. A median of 430 (range 1 to 3,437) individual nuclei was detected in each tile. To account for varying synovial tissue sample sizes, we normalized the whole slide image nuclei count by tissue area per tile and refer to this value herein as the nuclei density.

Algorithm validation and robustness. Through a series of experiments, algorithm performance was evaluated under varying conditions to determine its accuracy and robustness. The spatial distribution of nuclei (ie, sparse, dense, and clustered) did not have a large impact on the algorithm performance provided because there was some separation between the nuclei (Figure 2A). There were instances of multiple overlapping nuclei being identified as one, but in most cases, even a trained expert would have difficulty in separating the nuclei. The algorithm's performance under varying H&E staining concentrations was evaluated by initially selecting multiple tiles and then digitally modifying them to generate a broad range of realistic H&E stain variations (30). The algorithm appears to be robust to a range of staining variations (Figure 2B). An expert pathologist reviewed 10 sample tiles ranging from low to high inflammatory grades. Each tile was annotated for mislabeled or missing nuclei. Using that assessment as the gold standard, the algorithm detected nuclei with a sensitivity of 97% and a specificity of 100%, as shown in Fig. 2C.

Comparison of nuclei density with synovial histology features. Synovial tissue samples with a high histological score for lymphocytic inflammation typically contained more nuclei (Figure 3A and 3B). Nuclei density varies between tiles in a whole slide image, and the distribution of nuclei density can provide useful insights into inflammatory burden. For example, the nuclei density histograms demonstrated that increased lymphocyte infiltration is associated with increased nuclei density (Figure 3C and 3D). This pattern is consistent when considering all patients with RA across each lymphocyte grade (Figure 3E). With increasing lymphocyte grade, the peak of the histogram shifted to the right, and the histogram had a longer tail. Interestingly, lymphocyte grade 4 had a wide distribution with two peaks indicating that the most inflamed samples were the most heterogenous. This heterogeneity can be seen in the representative example shown for grade 4 inflammation, in which a large range of nuclei densities were observed over adjacent tiles.

Nuclei density was found to have the strongest correlation with lymphocytic inflammation and was also correlated with many
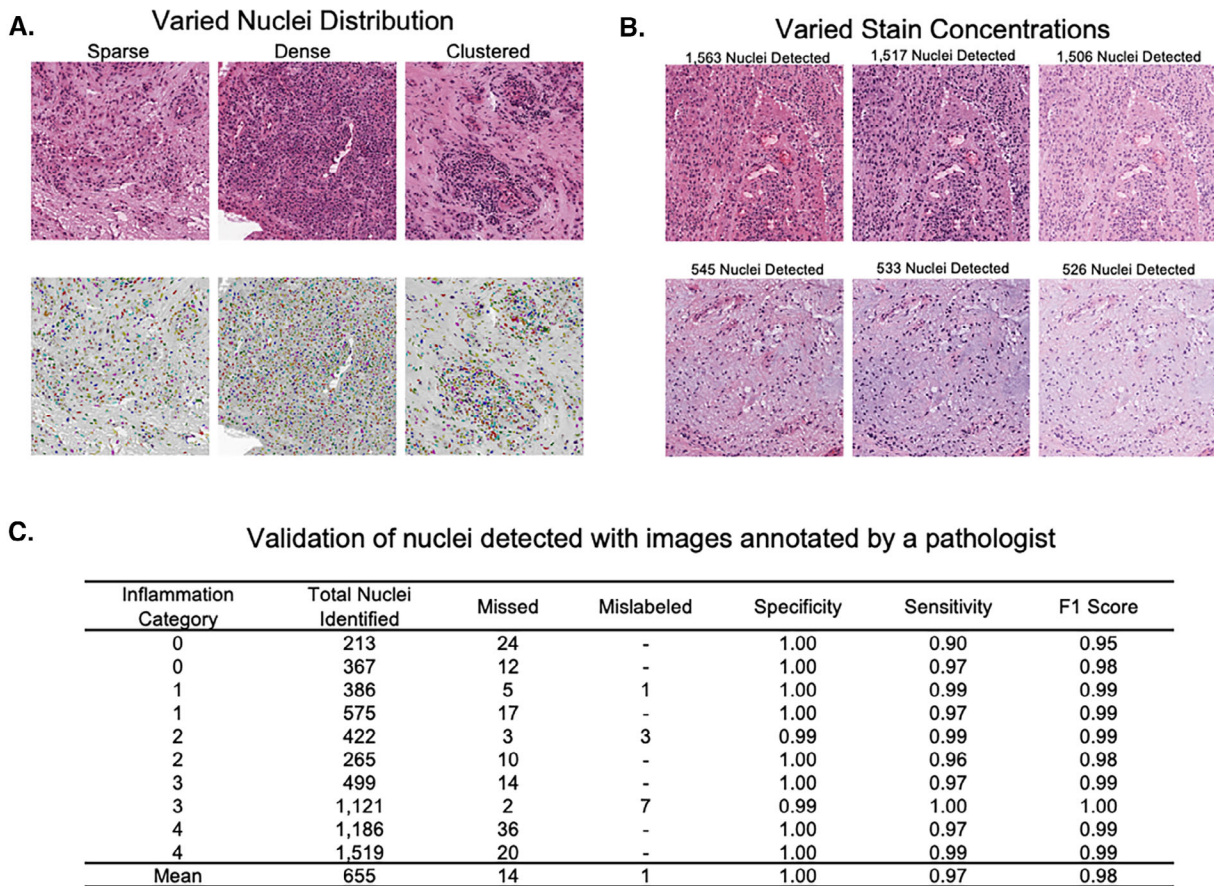
A. Varied Nuclei Distribution

B. Varied Stain Concentrations

C. Validation of nuclei detected with images annotated by a pathologist

Inflammation Category	Total Nuclei Identified	Missed	Mislabeled	Specificity	Sensitivity	F1 Score
0	213	24	-	1.00	0.90	0.95
0	367	12	-	1.00	0.97	0.98
1	386	5	1	1.00	0.99	0.99
1	575	17	-	1.00	0.97	0.99
2	422	3	3	0.99	0.99	0.99
2	265	10	-	1.00	0.96	0.98
3	499	14	-	1.00	0.97	0.99
3	1,121	2	7	0.99	1.00	1.00
4	1,186	36	-	1.00	0.97	0.99
4	1,519	20	-	1.00	0.99	0.99
Mean	655	14	1	1.00	0.97	0.98

Figure 2. Validation of algorithm nuclei detection under varying conditions. **A,** Algorithm performance when nuclei are spatially distributed in a sparse, dense, and clustered manner. Top row contains the original H&E-stained tiles, and the bottom row shows the nuclei detected by the algorithm. Note that the same color can be reused to display multiple separate nuclei in the image. **B,** Number of nuclei detected for two representative examples under a range of realistic staining concentrations. The first image in each row is the original tile and staining. **C,** Comparing nuclei detected by the algorithm to nuclei annotated by a pathologist in 10 tiles with different grades of lymphocyte inflammation. H&E, hematoxylin and eosin.

other histology features (Figure 4A). Sublining giant cells and stromal features such as detritus, fibrosis, and mucin were not significantly related to nuclei density. The one-way Kruskal-Wallis test demonstrated that there was a statistically significant difference in the mean nuclei density ($P < 0.05$) between grades of lymphocytes and other histology features correlated with lymphocytes (Figure 4B).

Comparison of nuclei density with gene expression measurements. RNA-seq gene expression analysis was previously performed on a subset (n = 36) of synovial samples and characterized as having either low, mixed, or high inflammatory gene expression (9). Mean nuclei density was statistically different among RNA subtypes ($P < 0.05$) and was increased in high inflammatory RNA subtype samples (Figure 5A). From the principal components analysis, the first principal component (PC1) explained 43% of the variance and was lowest in low inflammatory subtypes and highest in high inflammatory subtypes (Figure 5B). The PC1 score can therefore be interpreted as a

summary bulk gene expression score for general inflammatory burden. Nuclei density was significantly correlated with the PC1 score (r = 0.45, $P < 0.05$) (Figure 5C). We next used linear modeling to identify 915 genes that were significantly associated with nuclei density (FDR less than 0.05) (Figure 5D). Genes that were significantly increased in association with nuclei density included chemokines, chemokine receptors (such as CXCL9, CXCL10, and CCR5), and genes important for B cell activation and ectopic lymphoid structure formations (such as IL21R, TNFSF13B, and CD38), as well as JAK2, EOMES, and CTLA4. Genes that were significantly decreased in association with nuclei density included genes in the TGFB pathway such as TGFBR3, SMAD3, BMP4, and TIMP2, as well as cartilage intermediate layer protein 1 (CILP) and nerve growth factor receptors and neural guidance molecules such as SEMA4C, NTRK1, and NTRK2. A comparison of gene expression changes supervised according to nuclei density (log Fold Change vs. significance in the volcano plot) highly overlap with previously identified high inflammatory RNA subset genes identified using unsupervised clustering (presented as orange
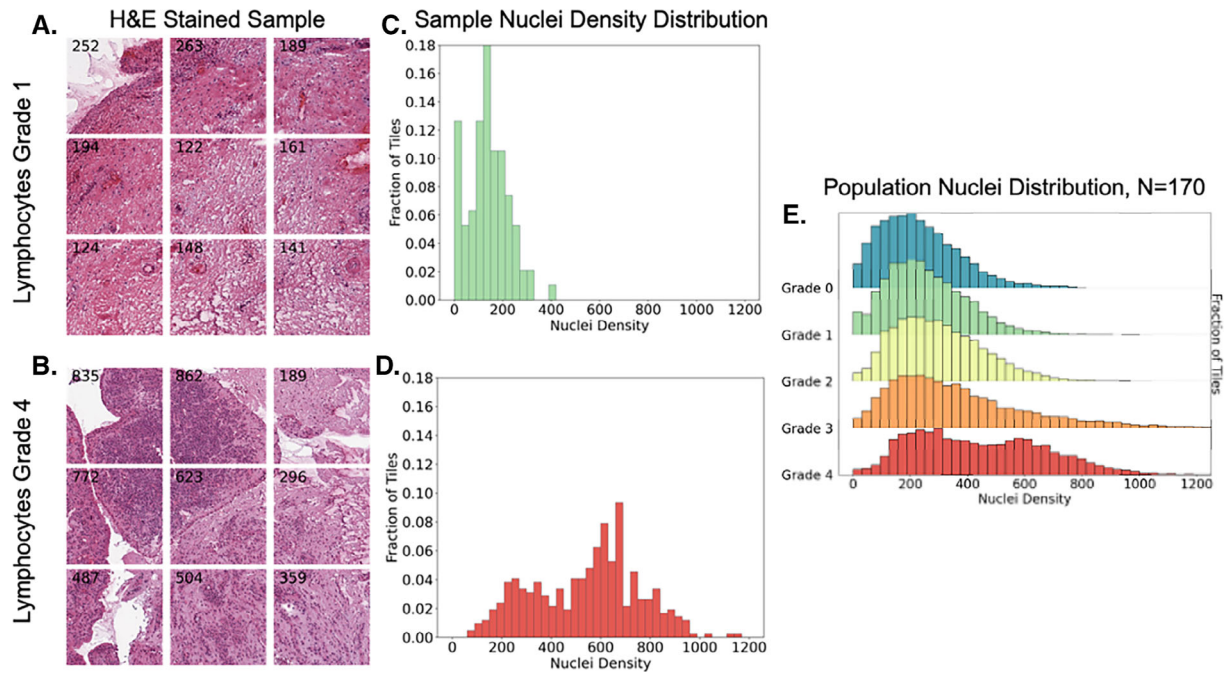
Figure 3. Quantification of nuclei density in 170 RA patient synovial samples. **A,B,** Representative images of H&E-stained synovium with pathologist lymphocytic inflammation scores of grade 1 (low inflammation) and grade 4 (high inflammation). The nuclei density metric for each tile is shown at the top left of the image. **C,D,** Normalized distribution of nuclei density across all tiles in a whole slide image for each representative sample of inflammation. **E,** Normalized distribution of nuclei density across all tiles in 170 patients with RA synovial samples grouped by inflammation grade. H&E, hematoxylin and eosin.

vs. green in the volcano plot in Figure 5D). Of the 2,460 genes that were increased in high inflammatory samples, 486 were also increased in association with nuclei density, and none were decreased in association with nuclei density. Of the 3,106 genes that were decreased in high inflammatory samples, 360 were also decreased in association with nuclei density, and none were increased in association with nuclei density.

Eighteen synovial cell subsets were previously defined in the RA synovium using single cell RNA-seq (4). We used gene-set enrichment analysis on 200 marker genes for each of the synovial cell type subsets in our paired bulk tissue RNA-seq data to determine whether nuclei density was associated with gene expression changes in any of these subsets. Consistent with the prior report, cells associated with leukocyte-rich RA synovium, such as
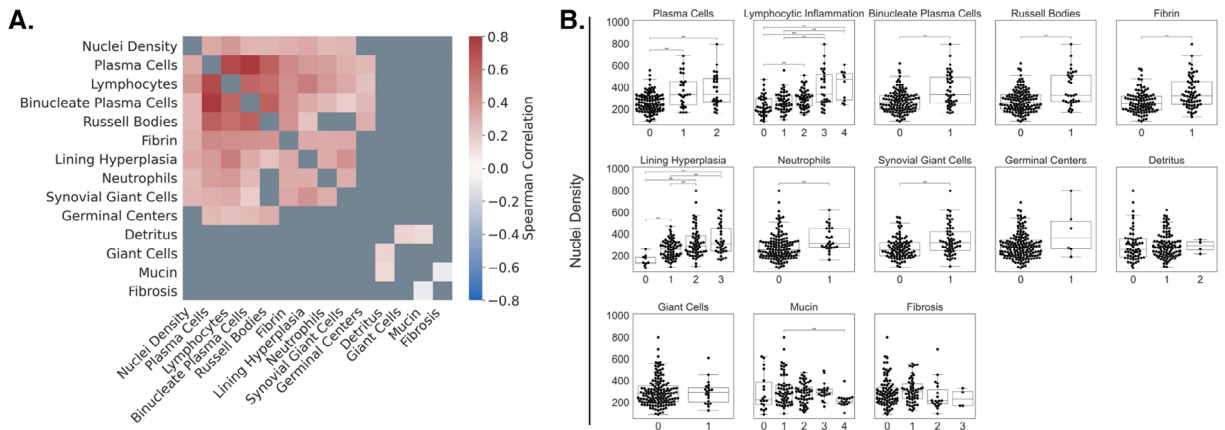

Figure 4. Comparison of computer assessments of nuclei density with human pathologist assessments of synovial inflammatory features. **A,** Heatmap of the Spearman correlation for 13 histological features and nuclei density. Correlation coefficients were only calculated for pairs of features that were determined to have a significant relationship as determined by the Chi-squared test. **B,** Computer assessments of mean nuclei density by pathologist assessments of histological features. The bar and ** indicate that there is a statistical difference according to the Kruskal-Wallis test with a Dunn's test corrected with Holm's adjustment for multiple comparisons.
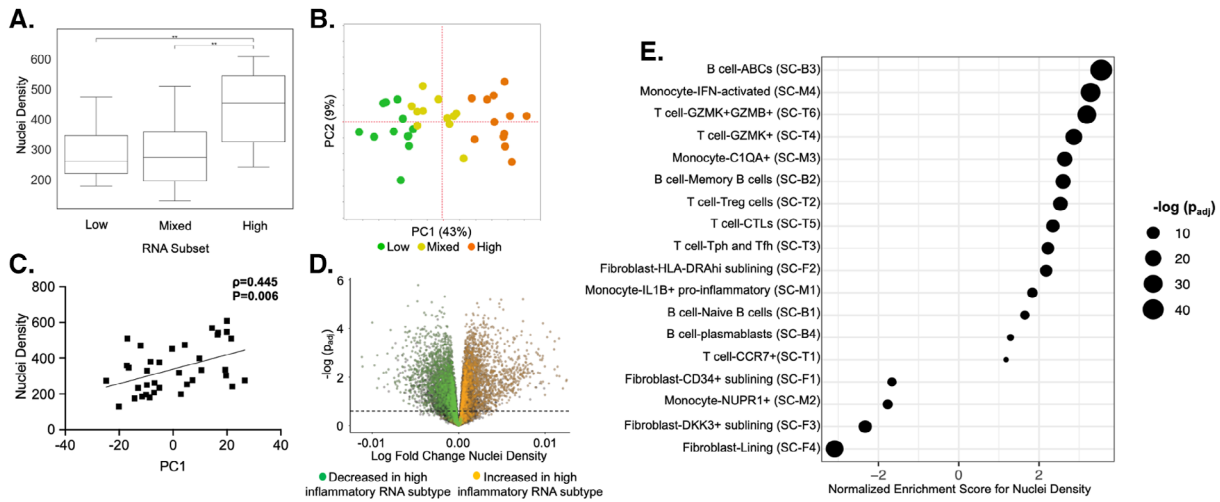
Figure 5. Comparison of computer assessments of nuclei density with gene expression measurement of inflammation. **A,** Mean nuclei density according to RNA subset. The bar and ** indicate that there is a statistical difference according to Kruskal-Wallis test with a Dunn's test corrected with Holm's adjustment for multiple comparisons. **B,** Principal component analysis of the top 5,000 most variably expressed genes in 36 patients with RA colored based on their characterization of having low, mixed, or high inflammatory gene expression. **C,** Mean nuclei density according to PC1 score. Data presented indicate Spearman correlation and P value. **D,** Log fold change in gene expression according to nuclei density versus significance ($-log(p_{adj})$). Individual genes are colored according to whether they were significantly ($P < 0.05$) increased or decreased in high inflammatory RNA subtype. Dashed line indicates the threshold for significance at $-log(p_{adj}) = 1.3$. **E,** Normalized enrichment scores for 200 marker genes from 18 synovial single cell RNA-seq cell type subsets according to mean nuclei density. Size of bubble indicates significance. Location on x-axis indicates enrichment according to nuclei density. PC1, first principal component; PC2, second principal component; RA, rheumatoid arthritis; RNA-seq, RNA sequencing.
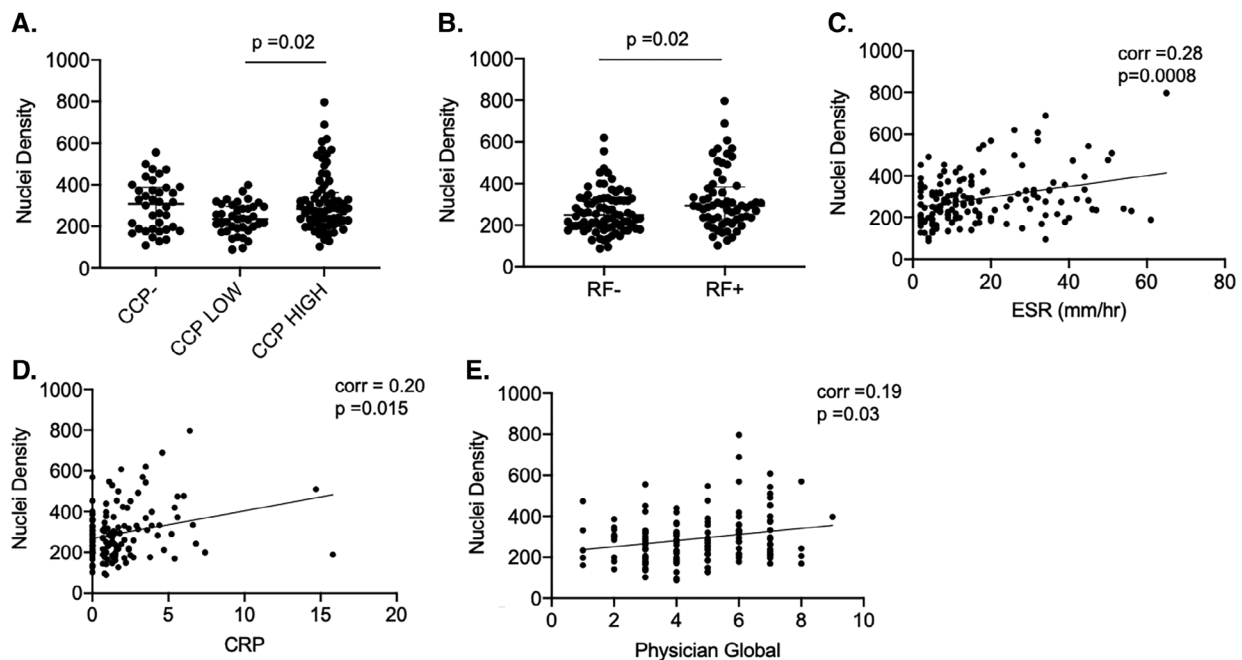

Figure 6. Comparison of computer assessments of nuclei density with blood markers of RA. Mean nuclei density according to **A,** CCP, **B,** RF, **C,** ESR, **D,** CRP, and **E,** Physician global assessment of disease activity. For categorical variables, P values reflect results of Kruskal-Wallis test with a Dunn's test corrected with Holm's adjustment for multiple comparisons. CCP high indicates of more than 250. For continuous variables, data presented indicate Spearman correlation and P value. CCP, anti-cyclic citrullinated peptide; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; RA, rheumatoid arthritis; RF, rheumatoid factor.

age-related memory B cells, interferon (IFN) activated and IL1B+ monocytes, and HLA-DR+ (Human Leukocyte Antigen - DR Isotype) sublining fibroblasts were enriched in samples with increased nuclei density, whereas cells that were previously associated with leukocyte poor synovium were negatively associated with nuclei density, such as lining fibroblasts and M2-like NUPR1+ monocytes (Figure 5E). Our analysis additionally detected significant enrichment in T-cell subsets such as GZMK+/GZMB+ T cells, GZMK+ T cells, C1QA+ monocytes, and cytotoxic T cells in accordance with increased nuclei density.

Comparison of nuclei density with clinical measures. Nuclei density was compared with clinical measures of disease activity. Mean nuclei density was significantly increased in samples labeled as RF+ or having high titer CCP (Figure 6A and 6B). Similarly, nuclei density was significantly correlated with elevated acute phase reactants, ESR and CRP, and global physician assessment of disease (Figure 6C-6E). There was no significant correlation between nuclei density and age, duration of disease, body mass index, tender joint counts, swollen joint counts, patient global assessment of disease, or pain (data not shown).

DISCUSSION

Here, we describe the development and application of a computer vision algorithm to automatically count hundreds of thousands of cell nuclei within H&E-stained RA synovium whole slide images. We assessed the validity of the algorithm by comparing nuclei density with histology features scored by a pathologist, gene expression analysis of inflammation, and clinical features of RA. We demonstrated that nuclei density was associated with histology features, such as lymphocytes, that are features of inflammation in RA. Nuclei density was also significantly increased in high inflammatory RNA gene expression RA subsets and significantly correlated with the PC1 inflammation score. We also found that marker genes from single cell RNA-seq subsets, such as age-related memory B cells, IFN activated and IL1B+ monocytes, and HLA-DR+ sublining fibroblasts, that were previously found to be increased in samples with increased leukocyte infiltration as assessed by flow cytometry were enriched in samples with increasing nuclei density. Our analysis further identified additional significant associations with other T and monocyte cell types. This discrepancy was likely due to nuclei density potentially being a more accurate assessment of leukocyte infiltration by not requiring tissue freezing or dissociation, which can lead to cell death. Nuclei density was associated with RA clinical features (ie, RF, CCP, ESR, and CRP) related to synovial inflammation. Results demonstrate that whole slide nuclei density is a robust measurement of inflammatory burden and correlates with multiple orthogonal measurements of inflammation.

Computer vision nuclei counting provides the potential to transform assessment of inflammation in synovial and other tissue biopsies by providing a precise, automatic, cost-efficient, and quantitative approach amenable to universal distribution and use. Automated nuclei detection is an important step in developing standardized and reproducible assessments of synovial inflammation, which are essential to build guidelines for incorporating histological findings in diagnosing RA and optimizing treatment decisions. Highly inflamed synovium tended to be heterogeneously distributed with some areas having approximately 200 nuclei per tile and others with upward of approximately 1,000 nuclei per tile (Figure 3C). This finding illustrates a key challenge in interrater reliability when pathologists subjectively choose a subset of high-power fields for assessment. Moreover, nuclei density is a continuous measurement and likely can better characterize and discriminate among inflammatory burden types.

The synovial samples assessed in this study were acquired at arthroplasty, and some suggest that these are less informative because the patients at arthroplasty have "end-stage RA." However, we have demonstrated that patients with RA have active disease at the time of arthroplasty, with moderate to high disease activity present in 70% of patients. Moreover, we propose that studies of inflammation in long-standing RA deserve consideration, because this population represents the vast majority of patients with RA and their ongoing management is an enormous financial burden and warrants optimization.

Another limitation of this study was the assessment of only one synovial biopsy per joint. Consistent with the patchiness of inflammation in a single whole slide image reported here, others have shown that there is a high degree of patchiness across the whole joint, and the accepted standard in the field is assessment of a minimum of six synovial biopsies for any one joint in order to accurately represent the inflammation in that joint. As opposed to ultrasound guided synovial biopsies, arthroplasty yields the entire joint explant, which was inspected by our pathologist, and the site of the most obvious inflammation (opaque and dull) was selected for biopsy. This selection bias may enrich for detection of inflammation. It is also important to note that the purpose of this algorithm is to address interrater variability between two pathologists rating the same image not to address intra-joint variability. Studies comparing the variance within individual inflamed joints indicate that six biopsy samples across a joint are needed to decrease the variance to less than 10% (31). This need for multiple samples across a joint compounds the time needed for a pathologist assessment as well as interrater variability, underscoring the need to develop automated computer vision to standardize assessments of synovial inflammation.

The algorithm uses classical computer vision techniques to identify synovial nuclei based on color, shape, and size. Other histology features and artifacts, such as excess staining with a similar appearance, can be incorrectly identified as nuclei. Dense clusters of nuclei, especially those with adjacent or overlapping nuclei, can be incorrectly labeled as a single nucleus. Typically, the nuclei count far exceeds the number of falsely identified

objects as demonstrated in the validation study. Thus, whole slide image statistics are robust to this type of error, but local statistics on smaller areas might be affected.

This work is an initial step toward the application of artificial intelligence and computer vision for understanding RA and developing automated tools for analyzing H&E-stained RA synovial samples. The proposed algorithm was developed for detecting individual nuclei but can be refined to identify patterns or groups of nuclei such as follicular structures, and it differentiate between cell types such as lymphocytes and other stromal cells. Inflammation and inflammatory infiltrates are common to many pathologies, and the proposed methodologies can be adapted for analyzing whole slide images of other tissue types and diseases.

ACKNOWLEDGMENTS

AUTHOR CONTRIBUTIONS

Study conception and design. Susan Goodman, Dana E. Orange.
Acquisition and analysis of data. Bella Mehta, Diyu Pearce-Fisher.
Algorithm design and analysis. Steven Guan, David Slater, James R. Thompson.
Analysis of data. Edward DiCarlo, Tania Pannellini, Fan Zhang, Soumya Raychaudhuri, Caryn Hale, Caroline S. Jiang.

DATA SHARING

The methods for scoring synovial histology features are available at https://www.hss.edu/pathology-synovitis.asp. The single-cell RNA-seq data, bulk RNA-seq data, and the clinical and histological data for this study are available at ImmPort (https://www.immport.org, study accession codes SDY998 and SDY1299). The source code for the nuclei counting algorithms is available at https://github.com/sgmitre/ai-histology.

REFERENCES

1. Dennis Jr G, Holweg CTJ, Kummerfeld SK Choy DF, Setiadi AF, Hackney JA, et al. Synovial phenotypes in rheumatoid arthritis correlate with response to biologic therapeutics. Arthritis Res Ther 2014; 16:R90.

2. Pitzalis C, Kelly S, Humby F. New learnings on the pathophysiology of RA from synovial biopsies. Curr Opin Rheumatol 2013;25:334–44.

3. Hogan VE, Holweg CTJ, Choy DF, Kummerfeld SK, Hackney JA, Onno Teng YK, et al. Pretreatment synovial transcriptional profile is associated with early and late clinical response in rheumatoid arthritis patients treated with rituximab. Ann Rheum Dis 2012;71:1888–94.

4. Zhang F, Wei K, Slowikowski K, Fonseka CY, Rao DA, Kelly S, et al. Defining inflammatory cell states in rheumatoid arthritis joint synovial tissues by integrating single-cell transcriptomics and mass cytometry. Nat Immunol 2019;20:928–42.

5. Smith MD, Baeten D, Ulfgren AK, McInnes IB, Fitzgerald O, Bresnihan B, et al. Standardisation of synovial tissue infiltrate analysis: how far have we come? How much further do we need to go? Ann Rheum Dis 2006;65:93–100.

6. Kraan MC, Haringman JJ, Ahern MK, Breedveld FC, Smith MD, Tak PP. Quantification of the cell infiltrate in synovial tissue by digital image analysis. Rheumatology 2000;39:43–9.

7. Kraan MC, Smith MD, Weedon H, Ahern MJ, Breedveld FC, Tak PP. Measurement of cytokine and adhesion molecule expression in synovial tissue by digital image analysis. Ann Rheum Dis 2001;60:296–8.

8. Krenn V, Morawietz L, Burmester GR, Häupl T. [Synovialitis score: histopathological grading system for chronic rheumatic and non-rheumatic synovialitis]. Z Rheumatol 2005;64:334–42.

9. Orange DE, Agius P, DiCarlo EF, Robine N, Geiger H, Szymonifka J, et al. Identification of three rheumatoid arthritis disease subtypes by machine learning integration of synovial histologic features and RNA sequencing data. Arthritis Rheumatol 2018;70:690–701.

10. Pantanowitz L, Sharma A, Carter A, Kurc T, Sussman A, Saltz J. Twenty years of digital pathology: an overview of the road travelled, what is on the horizon, and the emergence of vendor-neutral archives. J Pathol Inform 2018;9:40.

11. Cooper LAD, Kong J, Wang F, Kurc T, Moreno CS, Brat DJ, et al. Morphological signatures and genomic correlates in glioblastoma. Proc IEEE Int Symp Biomed Imaging 2011;30:1624–7.

12. Luo X, Zang X, Yang L, Huang J, Liang F, Rodriguez-Canales J, et al. Comprehensive computational pathological image analysis predicts lung cancer prognosis. J Thorac Oncol 2017;12:501–9.

13. Bera K, Schalper KA, Rimm DL, Velcheti V, Madabhushi A. Artificial intelligence in digital pathology—new tools for diagnosis and precision oncology. Nat Rev Clin Oncol 2019;16:703–15.

14. Aeffner F, Wilson K, Martin NT, Black JC, Hendricks C, Bolon B, et al. The gold standard paradox in digital image analysis: manual versus automated scoring as ground truth. Arch Pathol Lab Med 2017;141: 1267–75.

15. Nation's top orthopedic hospital launches the first FDA approved digital pathology platform. Hospital for Special Surgery. February 9, 2021. URL: https://news.hss.edu/nations-top-orthopedic-hospital-launches-the-first-fda-approved-digital-pathology-platform/.

16. Sirinukunwattana K, Snead D, Epstein D, Aftab Z, Mujeeb I, Tsang YW, et al. Novel digital signatures of tissue phenotypes for predicting distant metastasis in colorectal cancer. Sci Rep 2018;8:13692.

17. Verma R, Kumar N, Patil A, Kurian NC, Rane S, Graham S, et al. MoNuSAC2020: a Multi-organ nuclei segmentation and classification challenge. IEEE Trans Med Imaging 2021;40:3413–23. https://doi.org/10.1109/tmi.2021.3085712

18. ABdolhoseini M, Kluge MG, Walker FR, Johnson SJ. Segmentation of heavily clustered nuclei from histopathological images. Sci Rep 2019; 9:4551.

19. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham III CO, et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 2010;62:2569–81.

20. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315–24.

21. Gurcan MN, Boucheron LE, Can A, Madabhushi A, Rajpoot NM, Yener B. Histopathological image analysis: a review. IEEE Rev Biomed Eng 2009;2:147–71.

22. Ruifrok AC, Johnston DA. Quantification of histochemical staining by color deconvolution. Anal Quant Cytol Histol 2001;23:291–9.

23. Nobuyuki O. A threshold selection method from Gray-Level histograms. IEEE Journals & Magazine 1979;9:62–6.

24. Irshad H, Veillard A, Roux L, Racoceanu D. Methods for nuclei detection, segmentation, and classification in digital histopathology: a review—current status and future potential. IEEE Rev Biomed Eng 2014;7:97–114.

25. Maurer CR, Rensheng Qi, Raghavan V. A linear time algorithm for computing exact Euclidean distance transforms of binary images in arbitrary dimensions. IEEE Trans Pattern Anal Mach Intell 2003;25:265–70. https://doi.org/10.1109/tpami.2003.1177156

26. Bankhead P, Loughrey MB, Fernández JA, Dombrowski Y, McArt DG, Dunne PD, et al. QuPath: open source software for digital pathology image analysis. Sci Rep 2017;7:16878. https://doi.org/10.1038/s41598-017-17204-5

27. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29:15–21.

28. Ritchie ME, Phipson B, Wu D. Hu Y, Law CY, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43:e47.

29. Korotkevich G, Sukhov V, Budin N, Shpak B, Artyomov M, Sergushichev A. An algorithm for fast pre-ranked gene set enrichment analysis using cumulative statistic calculation. BioRxiv 2016; 060012.

30. Tellez D, Balkenhol M, Otte-Holler I, van de Loo R, Vogels R, Bult, P, et al. Whole-slide mitosis detection in HE breast histology using PHH3 as a reference to train distilled stain-invariant convolutional networks. IEEE Trans Med Imaging 2018;37: 2126–36.

31. Dolhain RJ, Ter Haar NT, De Kupoer R, Nieuwenhuis IG, Zwinderman AH, Breedveld FC, et al. Distribution of T cells and signs of T-cell activation in the rheumatoid joint: implications for semiquantitative comparative histology. Br J Rheumatol 1998;37: 324–30.