Artificial Intelligence-based Histopathological Analysis Identifies Unique Correlates of Tuberculosis

[1]Lucas Stetzik, [1]Thomas Westerling-Bui, [1]Hanna-Kaisa Sihvo, [1]Saana Rekinen, [1]Sami Blom, [1]Richard Fox, [2]Anas Alsharaydeh, [2]Gillian Beamer

[1]Aiforia, Helsinki, Finland and Cambridge, MA, USA; [2]Texas Biomedical Research Institute, San Antonio, TX, USA

INTRODUCTION

Tuberculosis (TB) remains a global health threat, with one person dying every 20 minutes of this infectious disease. Aerosol infection with *Mycobacterium tuberculosis* induces lung granulomas, which are complex cellular structures with critical microenvironments. A cornerstone of basic and clinical research to study granulomas and spatial relationships therein, is visual examination by pathologists using histochemical stains. We developed and applied a novel artificial intelligence (AI)-driven model to detect and quantify important pathological features of lung granulomas, such as necrosis, immune cell infiltrates, and acid-fast bacilli (AFB). Here, we (i) describe the AI model; (ii) show results from applying the model to hundreds of lung tissue sections; and (iii) perform statistical analyses on image analysis results to identify granuloma-level correlates of disease.

METHODS

We used Aiforia® Cloud 5.5 to develop a model trained on 124 representative digital images from 250 Diversity Outbred mice experimentally infected with ~25 Colony Forming Units (CFUs) of *M. tuberculosis.* Real-world heterogeneity included 8 independent experiments, 5 batches of histology stain, and 2 scanners (Aperio (Leica) ScanScope and AT2). Images were acquired at 40X magnification and uploaded. Following training and verification, the AI model was applied to lung tissue sections from 911 mice to quantify granuloma histological features.

LAYER & CLASS DEFINITIONS

CNN Layer	Class	Features	Area annotated (mm^2)	Area for training (mm^2)	Total Images
1	Tissue	All lung tissue	932.831	1306.155	124
2	Granuloma	All inflamed regions of lung tissue	417.785	670.257	124
	Non-granuloma	Normal lung tissue	239.698		
3	Bronchioles	Bronchiolar epithelium and lumens	12.903	93.248	121
	Bacterial mat	Mats of acid-fast bacilli (AFB)	0.055		
	Cells	Viable cells, fluid, fibrotic collagen	44.504		
	Lymphocytic cuffs	Lymphoplasmacytic cuffs	20.811		
	Necrosis cell poor	Eosinophilic cellular debris	4.768		
	Necrosis pyknotic	Basophilic nuclear debris	8.089		
4	Fibrotic collagen	Fibrotic collagen	0.944	2.89	121
	Fluid	Acellular fluid in alveoli	1.055		
	Foamy macs	Foamy macrophage cells/foci	0.156		
	Lymphocytes	Lymphocyte cells/foci	0.04		
	Multinucleated macs	Bi- and multinucleated macs	0.08		
	Neutrophils	Neutrophil cells/foci	0.013		
	Other macs	Activated macrophage cells/foci	0.381		
	Plasma cells	Plasma cells/foci	0.065		
5	Epithelium	Epithelial lining of bronchioles	1.115	0.443	57
	Lumen	Lumens of bronchioles	0.665		
6	Air	Air in bronchiole lumens	0.034	0.081	39
	Obstructed	Debris/fluid in bronchiole lumens	0.047		

Table 1. Convolutional neural network (CNN) layers, features, training data for region (segmentation) type. Six CNNs detect 21 different segmentation regions in digital images from mouse lung tissue sections stained by carbol fuschin and counter-stained with hematoxylin and eosin. (macs = macrophages)

CNN Layer	Class	Features	Total objects annotated	Total images
7	Acid Fast Bacilli (AFB)	Single AFB in 5µm diameter Clusters of AFB in 8µm diameter	3388	118
8	Lymphocyte nuclei	Nuclei of lymphocytes		124
9	Macrophage AFB	Single AFB in 5µm diameter Clusters of AFB in 8µm diameter	4217	124
	Macrophage nuclei	Nuclei of macrophages		
10	Multinucleated mac AFB	Same features as CNN 9	1969	123
	Multinucleated mac nuclei	Two or more macrophage nuclei in same cytoplasm		
11	Neutrophil nuclei	Neutrophil nuclei	1520	118
12	Plasma cell nuclei	Nuclei of plasma cells	2298	105

Table 2. Convolutional neural network (CNN) layers, features, training data for object detection type. Six CNNs detect seven different types of objects including 5 types of immune cell nuclei from lymphocytes, macrophages, multinucleated macrophages, neutrophils and plasma cells; and two types of AFB objects: Single and clusters of AFBs. (mac = macrophage)

LAYER TREE & EXAMPLE IMAGES

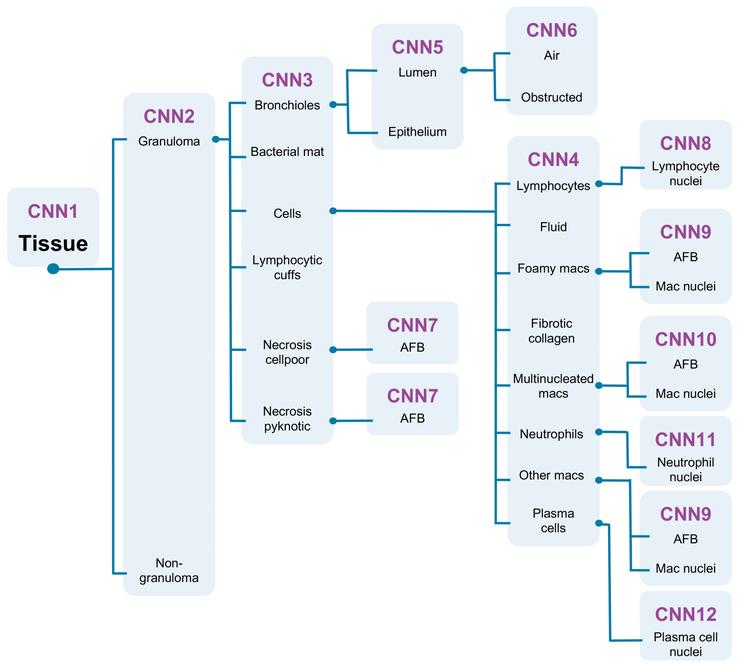


Figure 1. CNN layer tree and class names. Twelve CNNs were used to segment areas, detect/count AFB, and detect/count immune cells.

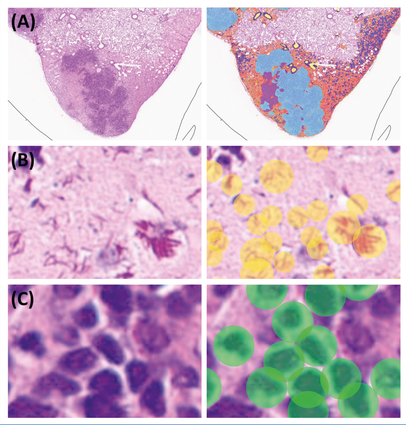


Figure 2. Representative lung images. (A) Segmentation masks of epithelium, lumen, fibrotic collagen, fluid, foamy macs, multinucleated giant macs, other macs, lymphocytes, neutrophils, plasma cells, necrosis.
(B) Detection of Acid Fast Bacilli (AFB). Single AFB (small yellow circles) and clusters of AFB (large yellow circles).
(C) Detection of lymphocyte nuclei (green circles).

MODEL TRAINING

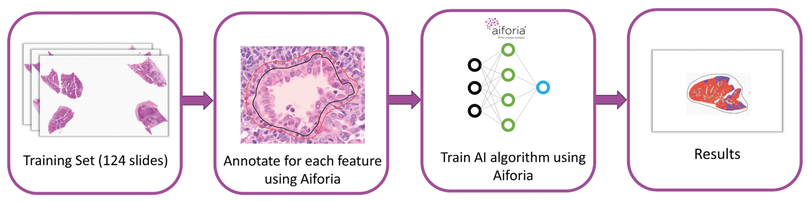


Figure 3. CNN training workflow. Supervised learning was performed using digitized images from 124 slides containing lung tissue sections from 250 mice. The model was developed following initial manual annotations of segmentation regions and objects, and then iteratively improved following review, addition/correction of annotations, and optimization of training parameters (not shown).

ACKNOWLEDGEMENTS

Individuals Areeha Batool, MS; Seung-Yi Lee, PhD; Anna Knuuttila, DVM, PhD; Sarah Kirschner-Kitz, Dr. Med.Vet, DECVP; Alireza Samiei, MD; Lindsey Smith, PhD and Mr. Tuomas Pitkanen are thanked for contributions to the AI model, and Veronika Ovsiannikova for figure design.

Services Biosafety services and facilities were provided by Tufts University's New England Regional Biosafety Laboratory. Histology services were provided by Tufts University's Comparative Pathology Core. Slide scanning services were provided by Vanderbilt University Medical Center's Digital Histology Shared Resource.

Funding NIH R21 AI115038; NIH R01 HL145411; American Lung Association Biomedical Research Grant RG-349504. Funders had no input for study design or interpretation.

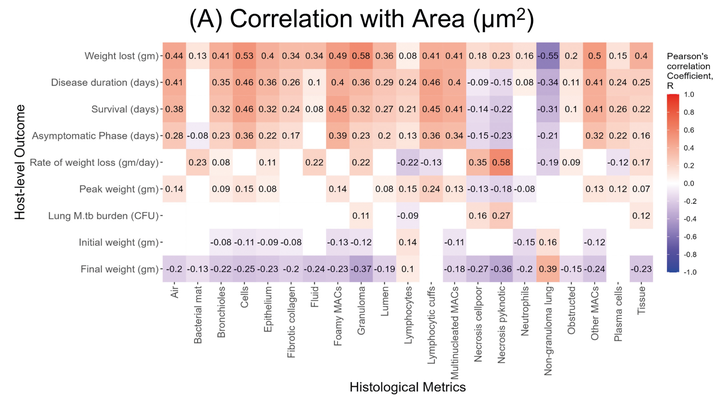
ETHICS STATEMENT

Tufts University's Institutional Animal Care and Use Committee approved experiments under protocol numbers G2012-53; G2015-33; G2018-33.

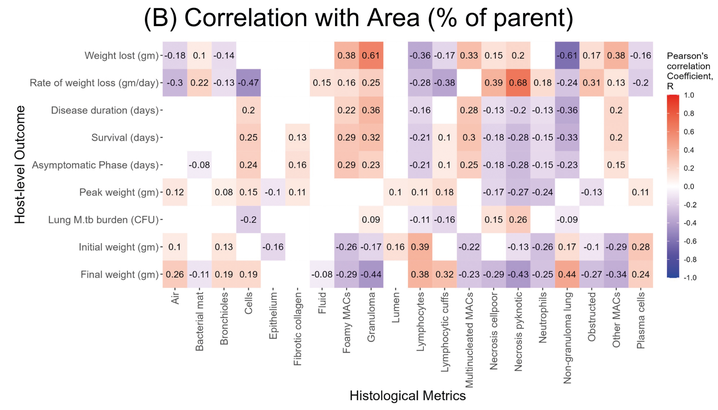
RESULTS

The average verification error rate for all layers, i.e., model compared to annotations in training regions, was 3.28%. The model quantified (i) area of all segmentation regions and percentage of parent area in Table 1; and (ii) numbers of immune cells and AFBs in Table 2. We used the data to identify significant correlations between granuloma features and host-level outcomes (Figure 4).

(A) Correlation with Area (µm²)



(B) Correlation with Area (% of parent)



(C) Correlation with Numbers of immune cells & AFB

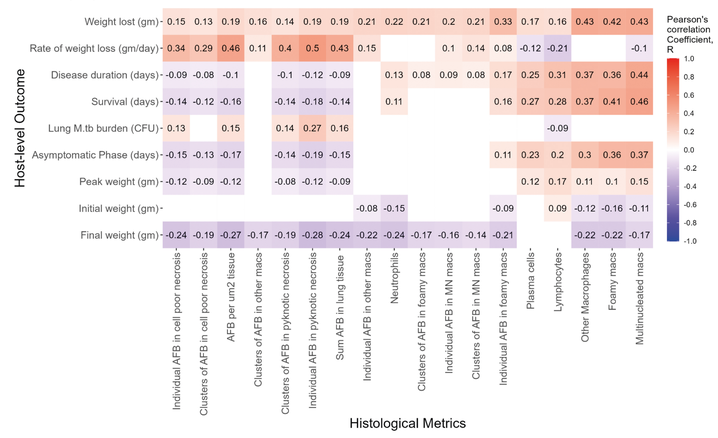


Figure 4. Correlation of AI model-quantified granuloma features (histological metrics) *vs* host-level outcomes. Significant correlations (q<0.05, adjusted for False Discovery Rate control level at 0.05) between absolute area (A), % of parent regions (B), and objects (C) are shown. Positive and negative correlations ranged from negligible to moderately positive. The strongest positive correlation was percent of granuloma area comprised of pyknotic nuclear debris *vs* rate of body weight lost, indicating that this type of necrosis is associated with rapid disease progression. The strongest negative correlation was between the percent of non-granulomatous lung tissue *vs* weight lost indicating that host ability to maintain normal lung tissue during infection is a sign of resistance to *M. tuberculosis*.

NEXT STEPS & CONCLUSIONS

This AI model is a work in progress. The next steps are to (i) perform validations to compare human-vs-model and human-vs-human on similar segmentation and object detection tasks; and (ii) critically examine the models' correlations and spatial metrics (not shown) for more insight. Overall, the AI-assisted, granuloma analysis provides a comprehensive, quantitative, and objective assessment of histopathological features that pathologists cannot perform manually. The data extracted will help to advance our understanding of tuberculosis pathogenesis; evaluate efficacy of novel vaccines and therapies; and in the future may aid in diagnosis of TB patients.

aiforia®
AI for image analysis

TEXAS BIOMEDICAL RESEARCH INSTITUTE