National Institute of Environmental Health Sciences
Division of Translational Toxicology

Performance evaluation of standard vs automated AI based image analysis of hepatic necrosis in KO vs WT mice infected with *Listeria monocytogenes.*

***Charan Ganta [1, 2], *David Cunefare[1,3], Priyanka Thakur[1,3], Prashant Rai[4], Eli Ney[1], Michael Fessler[4] and Ronald Herbert[1]**

[1]Division of Translational Toxicology, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA, [2]Inotiv, Morrisville, NC, USA, [3]Charles River Laboratories, Durham, NC, USA, [4]Division of Intramural Research, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA. ****Equal Authors***

Ab

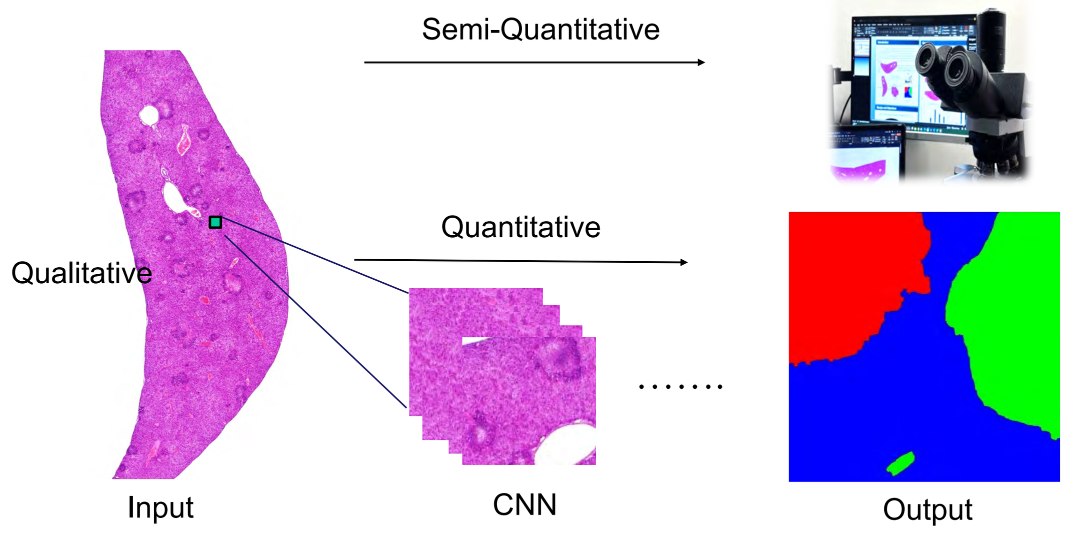
Introduction: Semiquantitative scoring of microscopic lesions is a gold standard practice in experimental/toxicologic pathology, however it is affected by intra- and inter-observer variability. With the emergence of artificial intelligence (AI) based image analysis, routine and time intensive evaluation could be automated and applied to repeat experimental studies. Here, we evaluated the performance of AI-based quantitative analysis with standard histopathological evaluation of *Listeria* induced hepatic necrosis/inflammation. **Experimental design:** Total 24 whole slide images (WSI) of the liver from 5 animal groups were used, of which 17 images were used for training and validation using an AI-based image analysis software. **Methods:** Two AI algorithms were trained, the first for tissue detection using a U-net convolutional neural network (CNN) at 4X resolution followed by a second using a DeepLab CNN at 20X magnification to detect necrosis/inflammation. The segmentations from these algorithms were used to determine the total tissue area and area of necrosis/inflammation for each image. The AI algorithms were validated against the diagnoses of two expert pathologists over 24 equally sized Regions of Interest (ROIs) evenly spread across the 8 validation images. **Results:** Compared to AI, the pathologist who was not involved in training had an average Dice's coefficient of 0.82 (r^2=95.7%); and the two pathologists compared to one another had an average Dice's coefficient of 0.84 (r^2=99.5%); across all images. **Conclusion:** There was a high correlation between the AI and pathologist -generated annotations for detecting liver necrosis/inflammation. **Impact statement:** This suggests, AI algorithm application can be applied to future studies with hepatic necrosis/inflammation.

Introduction

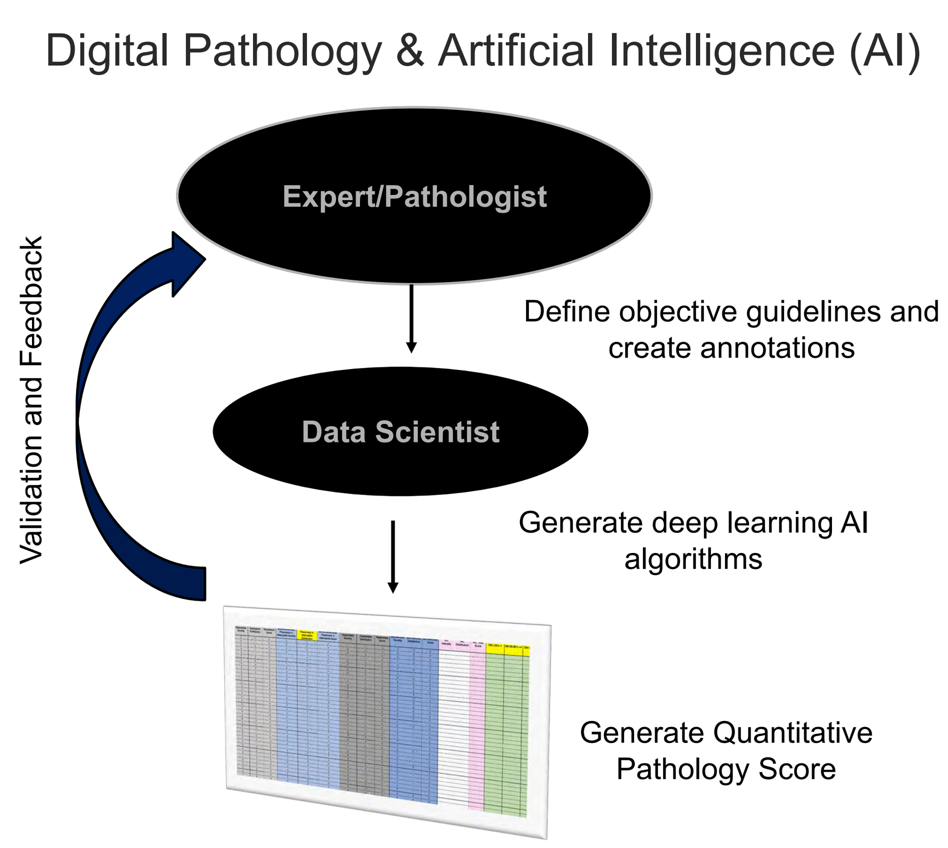
Whole slide imaging of histology slides using high throughput and high-resolution scanning platforms has enabled pathologists to increase the efficiency and accuracy of histopathological analysis.

In this study we evaluated performance of semiquantitative (manual) compared to quantitative (automated using AI based image analysis) scoring of *Listeria* induced hepatic necrosis/inflammation in WT and KO mice.

We used deep learning with convolutional neural networks (CNN) to directly measure the amount of necrosis or inflammation present in the tissue and validated the performance of the algorithm against that of expert pathologists.



Desig d Objectives

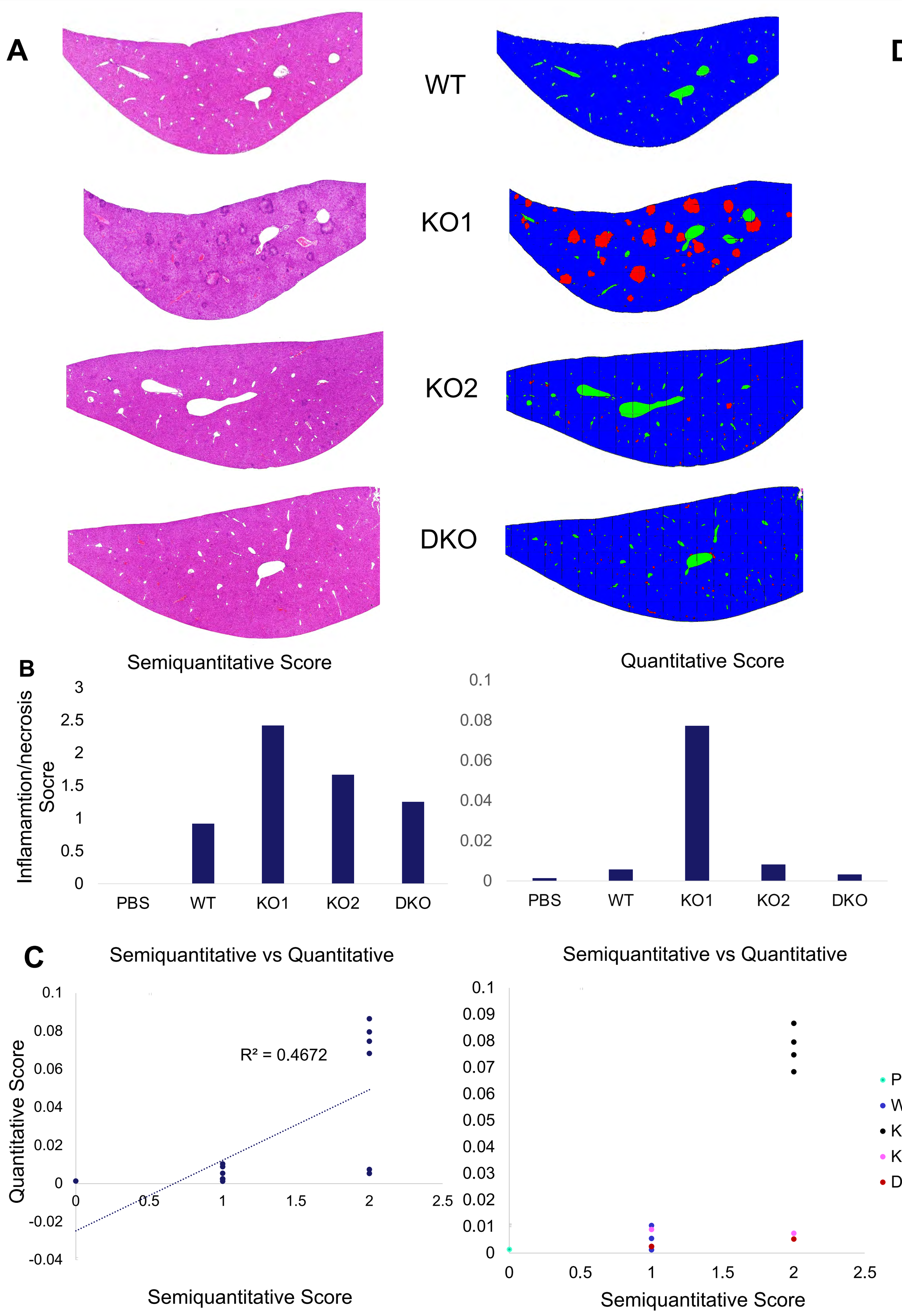


- Bias control
- Quantitative Scoring
- High throughput analysis
- Decreased turnaround time

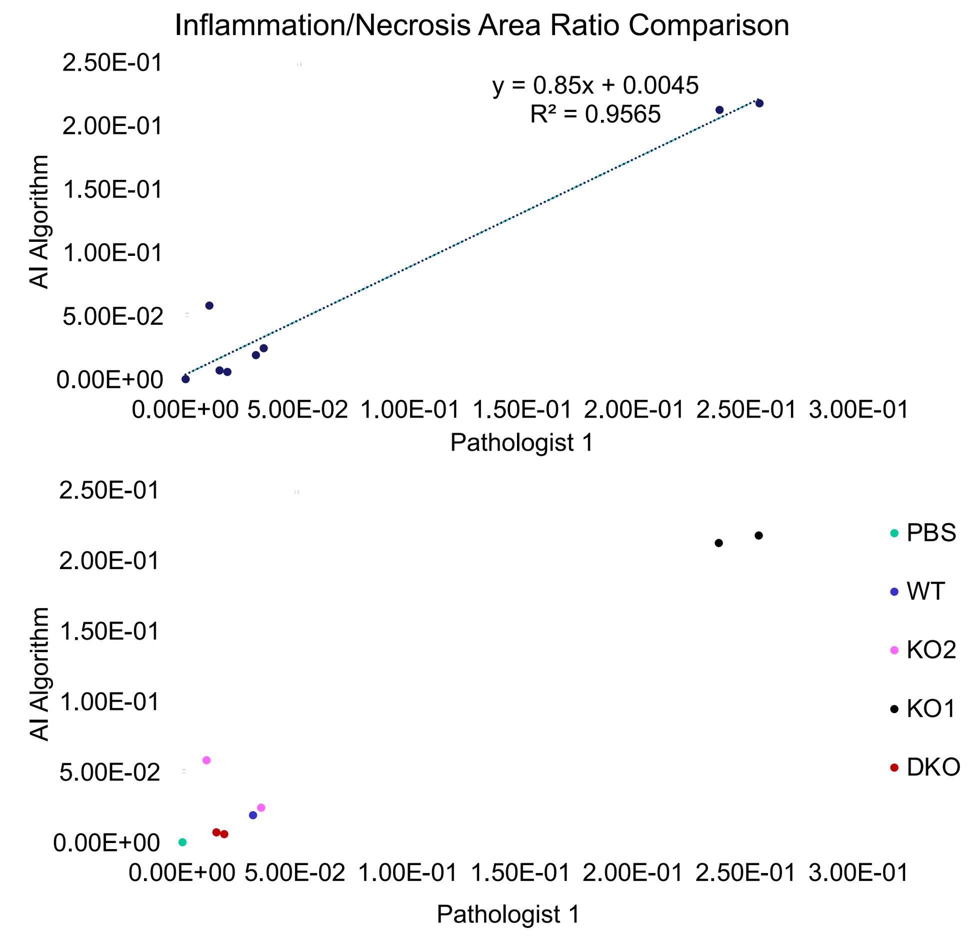
Meth d

- Total 24 whole slide images (WSI) from 5 different animal groups were used in this study.
- Image analysis was done using the Image Analysis Software platform. Two images from each group and one uninfected control (9 total) where used to train the AI applications.
- A U-net CNN was trained to detect tissue from background at 4X magnification.
- On the detected tissue a DeepLab CNN was trained to classify normal tissue, areas of necrosis/inflammation, and vasculature at 20X magnification.
- Quantified average area of necrosis/inflammation normalized by tissue area.
- Validation done on 8 images with no overlap with the training set, within 3 equal sized regions of interest per image. Two expert pathologists (pathologist 1/pathologist 2) annotated areas of necrosis/inflammation in the regions of interest (ROIs).
- Compared semi-quantitative scoring by pathologist 1 to quantitative analysis by the AI.
- Calculated agreement between the AI vs pathologist 1. Additionally, compared detected areas of inflammation/necrosis between the AI and pathologist and between the two pathologists (pathologist 1/pathologist 2).

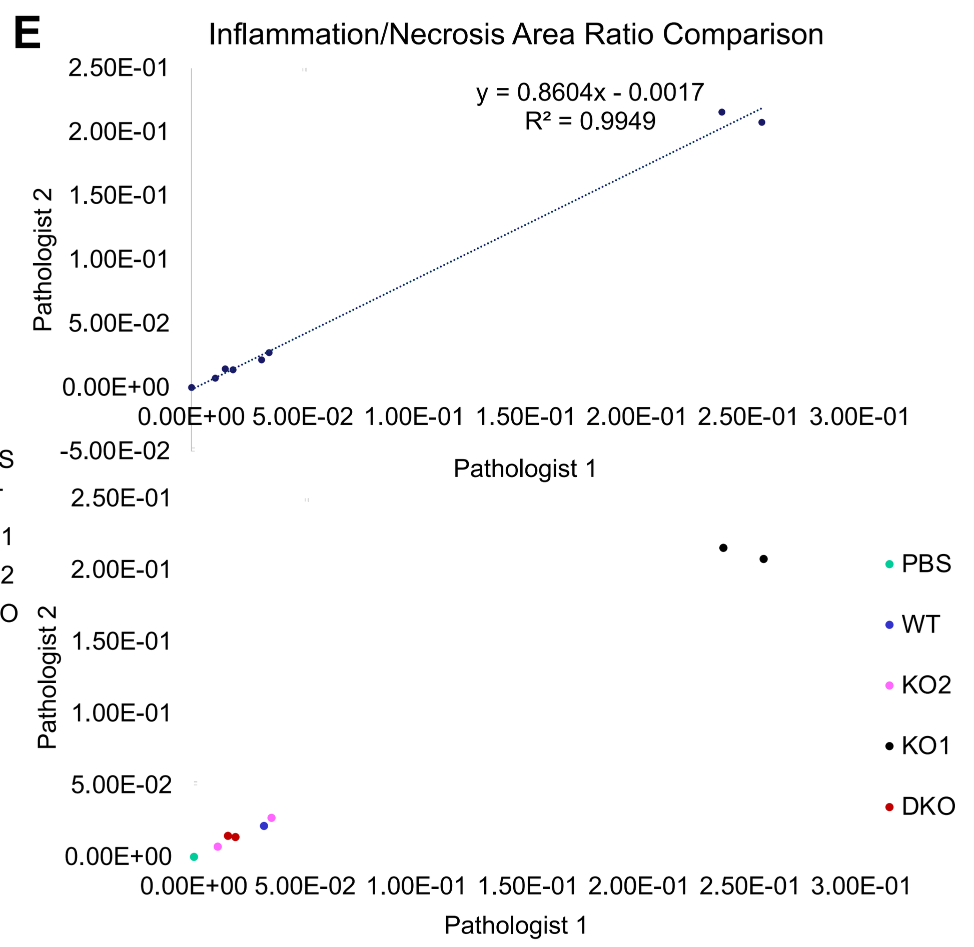
Result



Panel A, left Hematoxylin & Eosin (H&E)-stained images of liver sections from 4 groups used for scoring. Panel A, right - multicolor depiction of liver tissue representing various histological features recognized by AI algorithm for quantitative analysis: red - inflammation/necrosis; blue - normal tissue; green represents background. Panel B, left, - summary of semiquantitative scoring; and, Panel B, right - automated quantitative data summary. Panel C - correlation between semi-quantitative scoring by an expert pathologist and inflammation/necrosis area ratio automatically calculated by AI shown on the left, with group data shown on the right. WT=Wild type; KO= Knockout; Double KO.



Panel D showing correlation of inflammation/necrosis area ratios across validation data for annotations by pathologist 1 vs those generated by AI shown in the upper plot, with group data shown in the plot below. WT=Wild type; KO= Knockout; Double KO



Panel E showing correlation of inflammation/necrosis area ratios across validation data for two expert pathologists compared to one-another shown in the upper plot, with group data shown in the plot below. WT=Wild type; KO= Knockout; Double KO

Discussio

The coefficient of determination (r^2) between semi-quantitative (manual) and quantitative (AI automated) scoring is 46.7% suggesting the inherent difference in perception of the total area/lesion ratio between the two scoring methods, however there was agreement in the trend of lesion severity between different groups.

The reproducibility of AI based lesion detection between the training data set (pathologist 1) and testing data set (AI algorithm) showed an r^2 of 95.7%.

To confirm the accuracy of lesion detection (annotations) by pathologist 1, the lesions were independently annotated by pathologist 2 on the same training data set, the concordance in lesion detection between the pathologist 1 and pathologist 2 showed an r^2 of 99.5%.

Co clusio

There was a high degree of correlation between manual annotations generated quantitative data and automated AI generated quantitative data in the detection of hepatic necrosis/inflammation.

Automated quantitative evaluation of routine microscopic lesions is a powerful tool that can enhance the efficiency of histopathological analysis by providing additional support to the pathologist by shortening the turn around time when evaluating repeated and high throughput studies.

Established CNN algorithms (library) can be used and applied by investigating laboratory/pathologists for all future studies with minimal training.

Ref

Rai P, Sharpe M, Ganta CK, Baker PJ, Mayer-Barber KD, Fee BE, Taylor GA, Fessler MB. IRGM1 supports host defense against intracellular bacteria through suppression of type I interferon in mice. J Clin Invest. 2023 Nov 1;133(21):e171982.

Doğan RS, Yılmaz B. Histopathology image classification: highlighting the gap between manual analysis and AI automation. Front Oncol. 2024 Jan 17;13:1325271.

Rashidi HH, Tran NK, Betts EV, Howell LP, Green R. Artificial Intelligence and Machine Learning in Pathology: The Present Landscape of Supervised Methods. Acad Pathol. 2019 Sep 3;6:2374289519873088.

Zuraw A, Staup M, Klopfleisch R, Aeffner F, Brown D, Westerling-Bui T, Rudmann D. Developing a Qualification and Verification Strategy for Digital Tissue Image Analysis in Toxicological Pathology. Toxicol Pathol. 2021 Jun;49(4):773-783

Aeffner F, Sing T, Turner OC. Special Issue on Digital Pathology, Tissue Image Analysis, Artificial Intelligence, and Machine Learning: Approximation of the Effect of Novel Technologies on Toxicologic Pathology. Toxicol Pathol. 2021 Jun;49(4):705-708.

Zuraw A, Aeffner F. Whole-slide imaging, tissue image analysis, and artificial intelligence in veterinary pathology: An updated introduction and review. Vet Pathol. 2022 Jan;59(1):6-25.

Ackn wled

The authors would like to thank Eli Ney and Ashley Paragone from Image Analysis and Artificial Intelligence Core and Willie Cunningham, Quashana Brown and Renee Fisher form Histology and Necropsy Core, Comparative and Molecular Pathogenesis Branch (CMPB) at NIEHS for their support in making H&E slides and generating whole slide digital image scans for this study.