**POSTER PRESENTER**

**POSTER #2**

**DEVELOPMENT AND VISUAL ASSESSMENT OF A DEEP LEARNING SYSTEM FOR AUTOMATED TUBERCULOSIS SCREENING**

Tae Kyung Kim, BA, Paul H Yi, MD, Gregory Hager, PhD, Cheng Ting Lin, MD.

*Radiology Artificial Intelligence Laboratory, Johns Hopkins University School of Medicine*

**BACKGROUND:** Tuberculosis (TB) is one of the top 10 causes of death worldwide, with particularly high disease burden in China, which reports 1 million new cases of TB annually. Screening efforts amongst Chinese patients have been limited partially due to barriers to healthcare access, which could be addressed via automated chest radiograph (CXR) screening methods utilizing deep learning. The purpose of this study was to develop and test the performance of a deep convolutional neural network (DCNN) for the automated detection of tuberculosis on CXRs.

**METHODS:** We obtained 10,996 CXRs performed as part of regular clinical practice at the National Institutes of Health (NIH) via the NIH ChestX-Ray 14 database. A fellowship-trained cardiothoracic radiologist individually reviewed and annotated each CXR for findings suggestive of TB or no findings suggestive of TB. We trained the ResNet-50 DCNN (pretrained on ImageNet) using this dataset and validated & tested the DCNN using 800 CXRs obtained for TB screening in Montgomery County, Maryland (USA) and Shenzhen, China, respectively. Receiver operating characteristic (ROC) curve with area under the curve (AUC) were generated to evaluate the DCNNs’ performance. Visual validation of the algorithm was performed by creating heatmaps using Class Activation Mappings (CAM) as demonstrated by Zhou et al [*IEEE Conf. Comput. Vis. Pattern Recognit.* 2921–2929 (2015)].

**RESULTS:** Our highest-performing DCNN achieved AUC of 0.90 for detection of TB in patients with clinical & pathologically-confirmed TB, with an optimal threshold value resulting in sensitivity of 90% and 65% specificity, compared to sensitivity and specificity of 78% and 51% reported in radiologists. Heatmaps revealed that the DCNN appropriately emphasized the same regions of interest as the human radiologist, such as cavitary and non-cavitary parenchymal nodules/masses and hilar lymphadenopathy.

**CONCLUSIONS****:** We developed a DCNN utilizing the NIH ChestX-ray14 database with promising sensitivity and specificity to be used as a TB screening algorithm. The DCNN was highly generalizable, achieving AUC of 0.9 when tested against new CXRs performed in Chinese patients. The DCNN further emphasized similar regions as the human radiologist in its’ decision-making. These findings suggest that DCNNs may serve as clinically relevant tools for diagnosing TB; prospective clinical validation is warranted to determine an optimal AUC threshold for the DCNN.

**CONTENT CATEGORY:** Translational Science

**KEYWORDS:** *Radiology, Artificial Intelligence, Deep Learning, Tuberculosis, Global Health*