# Hierarchical Multi-Video Multiple Instance Learning: AI-Interpreted Point-of-Care Ultrasound for TB Detection

Trevor Brokowski[1], Mattia Nino Barbiere[2], Véronique Suttels[2], John A. Onofrey[1], and Mary-Anne Hartley[2]

[1] Yale University, New Haven, CT, USA
[2] Laboratory for Intelligent Global Health and Humanitarian Response Technologies (LiGHT), École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

**Abstract.** Tuberculosis (TB) remains the leading cause of death from an infectious agent, disproportionately affecting low-resource settings where diagnostic infrastructure is limited. Point-of-care lung ultrasound (POCUS) offers a portable, affordable modality for TB screening, but clinical diagnosis requires integrating heterogeneous evidence across multiple lung zones into a single patient-level decision, a challenge not addressed by existing single-video AI systems. We propose **Hierarchical Multi-Video Multiple Instance Learning (HMV-MIL)**, enabling patient-level TB detection directly from multi-video ultrasound acquisitions under weak supervision. HMV-MIL aggregates diagnostically informative information within and across videos to produce clinically meaningful patient-level predictions, while auxiliary pathology supervision aligns learned features with recognizable pulmonary findings. On a prospective cohort of 504 patients, HMV-MIL achieves 0.89 AUC, 0.82 sensitivity, and 0.75 specificity, outperforming baselines by 0.04-0.35 AUC and approaching WHO triage performance targets. These results demonstrate the feasibility of automated TB triage using frontline-acquired ultrasound data, supporting decentralized diagnostic assessment in settings where conventional testing remains inaccessible.

**Keywords:** Tuberculosis · Point-of-Care Ultrasound · Multiple Instance Learning · Weakly Supervised Learning

## 1 Introduction

Tuberculosis (TB) caused 1.23 million deaths in 2024, more than HIV/AIDS and malaria combined, with the burden concentrated in low- and middle-income countries [19,5]. Existing diagnostics present trade-offs: sputum microscopy detects only half of pulmonary TB cases [9]; NAATs are accurate but require stable infrastructure [7,3]; and chest radiography remains cost-prohibitive in primary-care settings [13]. Handheld point-of-care ultrasound (POCUS) probes offer portable, low-cost imaging by minimally trained operators [1,17], and early

AI studies have demonstrated LUS-TB screening using *expert-selected ultrasound images* meeting WHO sensitivity and specificity criteria [18]. However, most approaches analyze videos independently or rely on curated images whcih limits scalability and does not reflect frontline acquisition workflows[16,12,14]. Raw lung ultrasound (LUS) videos pose a harder problem: TB evidence is sparse, heterogeneous, and distributed across frames and bilateral thoracic zones, whereas existing LUS AI systems typically analyze single videos or manually selected keyframes.

We introduce Hierarchical Multi-Video Multiple Instance Learning (**HMV-MIL**), which frames patient-level TB detection as hierarchical multi-video MIL. Within each video, learned attention pooling identifies informative frames; across videos, gated MIL integrates site-level embeddings into a unified patient representation. Auxiliary pathology supervision using human-interpretable ultrasound findings regularizes the model toward recognizable pulmonary patterns. Validated on a prospective cohort of 504 patients acquired by minimally trained operators under a 14-site POCUS protocol, HMV-MIL demonstrates that raw multi-site LUS videos contain sufficient signal for automated TB triage and that hierarchical aggregation improves performance relative to conventional spatiotemporal architectures.

## 2 Related Work

AI-assisted TB screening has matured rapidly in chest radiography, with the WHO recommending computer-assisted diagnosis (CAD) since 2021 [19]. Systems such as qXR and CAD4TB can meet WHO triage benchmarks, however they depend imaging infrastructure often unavailable at primary-care levels in high-burden settings. Ultrasound offers an operationally aligned alternative: low-cost ( $2000), portable, minimal infrastructure required. Early studies demonstrate feasibility of AI-assisted LUS screening from individual frames [18] (AUROC 0.93); however, this approach requires trained sonographers to curate diagnostically informative frames prior to inference, presupposing the expert knowledge the system is meant to replace.

Existing multi-view AI systems, particularly in echocardiography, exploit standardized acquisition planes with stable anatomical meaning. TB lung POCUS differs: it is a weakly structured multi-site video problem with diagnostic evidence distributed across thoracic zones and frames. LUS deep learning has progressed from 2D CNNs to spatiotemporal and transformer architectures for COVID-19 classification [14], pneumonia detection [4], and B-line quantification [2,11,6]; however, these methods generally analyze single videos in isolation. Multiple-instance learning (MIL) provides a natural framework for weakly supervised clinical reasoning [8,10,15], but standard MIL assumes independent non-temporal instances and does not jointly model intra-video dynamics and inter-site anatomical aggregation.

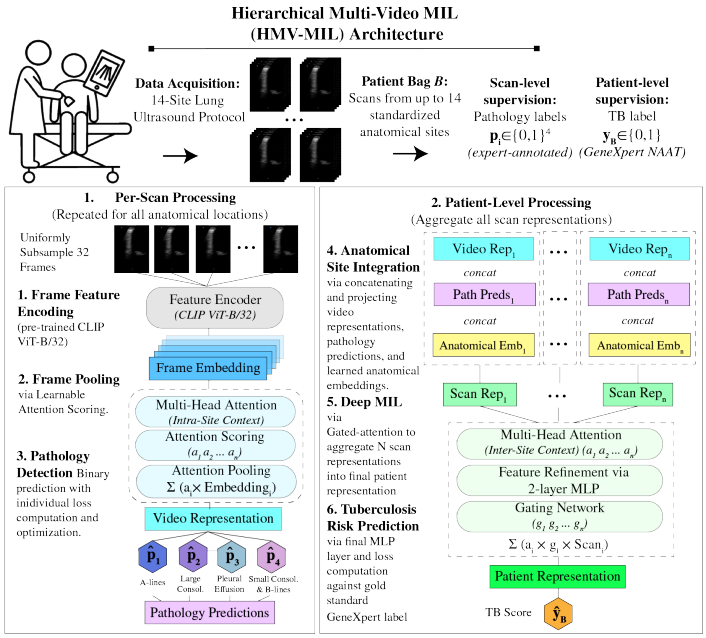


Fig. 1: **HMV-MIL Framework Overview.** Videos are subsampled and encoded with CLIP ViT-B/32, aggregated into video representations, and augmented with pathology predictions. Scan-level features and anatomical embeddings are combined using gated multi-head MIL attention to produce a patient representation for TB classification, trained with both patient-level and scan-level supervision.

## 3 Methodology

**Problem Formulation:** Each patient constitutes a bag $\mathcal{B} = \{V_1, \ldots, V_N\}$ of up to $N = 14$ videos from distinct anatomical sites. Each video $V_i \in \mathbb{R}^{T_i \times 3 \times 224 \times 224}$, where $T_i$ denotes the variable frame count for site $i$. Patient-level binary TB labels $y_{\mathcal{B}} \in \{0, 1\}$ are derived from GeneXpert MTB/RIF testing. As auxiliary supervision, each video $V_i$ carries expert-annotated binary pathology labels $p_i \in \{0, 1\}^4$ (A-lines, large consolidations, pleural effusion, B-lines/small consolidations), without direct TB indication. The model learns $f : \mathcal{B} \to y_{\mathcal{B}}$, requiring joint intra-video frame selection and inter-video site aggregation under anatomical heterogeneity and weak supervision.

### 3.1 HMV-MIL Architecture

Figure 1 illustrates HMV-MIL's architecture. (1) **Feature Extraction.** We use OpenAI CLIP (ViT-B/32) with pretrained weights and fine-tuned on COVIDxUS [4]. For TB adaptation, we freeze all transformer layers except the final encoder and

projection layers (15.2M trainable of 86.4M total). Each frame $F_t^i$ is independently encoded to $H_t^i \in \mathbb{R}^{768}$. (2) **Intra-Video Attention Pooling.** Frame embeddings are projected to latent representations $h_t^i \in \mathbb{R}^{512}$ via an encoder $f_{\text{enc}}$. Multi-head self-attention models temporal dependencies, producing contextualized embeddings $\tilde{h}_t^i$. A lightweight scoring network produces per-frame importance weights $\alpha_t = \text{softmax}(\text{MLP}_{\text{score}}(\tilde{h}_t^i))$, and the video representation is $z_{\text{video}} = \sum_t \alpha_t \tilde{h}_t^i$. This mechanism identifies diagnostically relevant frames while suppressing artifacts and redundant anatomy. (3) **Auxiliary Pathology Supervision.** For each $z_{\text{video}}$, four binary classifiers predict the presence of A-lines, large consolidations, pleural effusion, and B-lines/small consolidations. These heads regularize the visual encoder by aligning intermediate features with recognizable pulmonary patterns, and provide clinically interpretable per-site outputs for reliability auditing. (4) **Inter-Video MIL Aggregation.** Each site's representation $\tilde{H}_i \in \mathbb{R}^{512}$ is formed by concatenating $z_{\text{video}}$, the pathology prediction vector $p_i \in \mathbb{R}^4$, and a learned anatomical site embedding. A deep attention-based MIL module applies multi-head self-attention across sites, followed by MLP-with-residual transformation, and gated attention aggregation: $z_{\text{patient}} = \sum_{i=1}^{N} \frac{\alpha_i g_i}{\sum_j \alpha_j g_j} \tilde{H}_i' \odot \text{mask}_i$, where $\text{mask}_i$ handles missing sites. The patient embedding is classified as $\hat{y}_{\mathcal{B}} = \sigma(\text{MLP}(z_{\text{patient}}))$. To mitigate gradient interference, we sequentially update: (1) each pathology head independently; (2) the site integration, MIL, and TB classifier jointly; (3) the vision encoder and attention mechanism using the combined pathology + TB loss. Separate AdamW optimizers with hierarchy-scaled learning rates are used per group.

## 4 Experiments

### 4.1 Dataset and Protocol

We analyze a prospective cohort of 504 adults recruited at CNHU-PPC, Cotonou, Benin (October 2021-August 2023). Of these, 192 (38%) were GeneXpert-positive for TB (TB+) and 312 (62%) TB-. LUS was performed using Butterfly iQ 2.0 devices under a 14-site bilateral protocol, with 10–15s videos acquired at up to two depths per anatomical site, resulting in $\tilde{1}$4,258 videos and $\tilde{1}$.5M frames ($\tilde{2}$8 $\pm$ 6) videos/patient. Three trained sonographers annotated all analyzed videos for four pulmonary findings: A-lines (50%), B-lines/small consolidations (25%), large consolidations (20%), and pleural effusions (5%). All data were de-identified under Benin Ministry of Health and CNHU-PPC ethics approvals.

### 4.2 Evaluation and Implementation

We hold out a 20% test set (stratified by TB status, prior TB, and HIV status), with all data split at the patient level; the remaining 80% undergoes 5-fold cross-validation. Final results report mean performance from five models applied to the held-out test set. We compute mean and 95% CI for ROC-AUC and PR-AUC. The WHO screening threshold (90% sensitivity/$\geq$70% specificity) is applied by selecting the a threshold maximizing Youden's J on the validation folds,

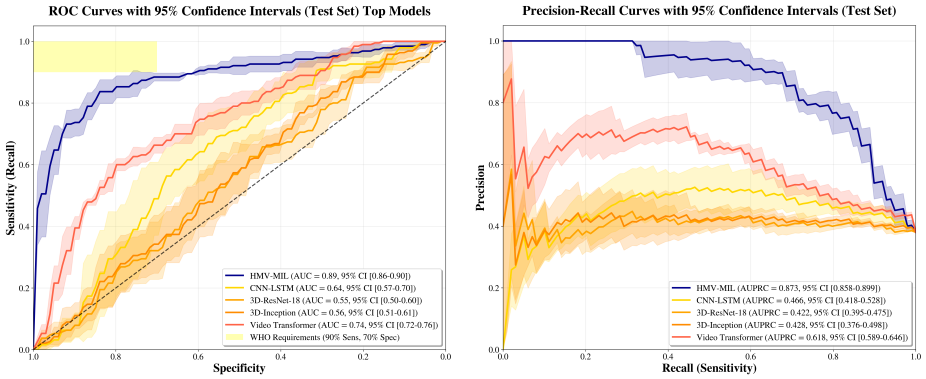


Fig. 2: ROC (left) and precision–recall (right) curves with 95% CIs. Yellow box: WHO screening targets ($\geq$90% sensitivity, $\geq$70% specificity).

Table 1: Patient-level TB detection. Mean $\pm$ SD across 5-fold cross-validation.

| Architecture | Params (M) | AUC | Sensitivity | Specificity |
|---|---|---|---|---|
| **HMV-MIL (Ours)** | 92.0 | $\mathbf{0.89 \pm 0.02}$ | $0.82 \pm 0.09$ | $\mathbf{0.75 \pm 0.25}$ |
| DINOv3-Bert | 418.6 | $0.85 \pm 0.02$ | $0.80 \pm 0.10$ | $0.67 \pm 0.21$ |
| CNN-LSTM (ResNet-18) | 25.7 | $0.64 \pm 0.05$ | $\mathbf{0.88 \pm 0.10}$ | $0.28 \pm 0.20$ |
| 3D ResNet-18 | 39.0 | $0.55 \pm 0.04$ | $0.59 \pm 0.26$ | $0.50 \pm 0.31$ |
| 3D-Inception | 35.6 | $0.56 \pm 0.04$ | $0.55 \pm 0.42$ | $0.53 \pm 0.37$ |
| Video Transformer | 94.5 | $0.74 \pm 0.02$ | $0.65 \pm 0.10$ | $0.71 \pm 0.11$ |

and then applied to the held-out test patients. Training used mixed-precision AMP, gradient accumulation (35 steps), AdamW ($\text{lr} = 2\times10^{-5}$), 200-epoch cosine schedule with positive-class weight 2.0, on 4×NVIDIA A100 GPUs with distributed data parallelism. All models were implemented in PyTorch (v2.2). We compare and ablate HMV-MIL across backbones: CNN-LSTM (ResNet-18), 3D ResNet-18, 3D-Inception (Kinetics-400 pretrained), Video Transformer (ViViT-B, Kinetics-400 pretrained), and DinoV3L (pretrained); and we assess the specific components of HMV-MIL (1) lightweight backbone (LeViT); (2) removal of LUS domain pre-training; (3) uniform frame subsampling vs. attention pooling; (4) frame pooling; (5) top-$k$ frame selection ($k \in \{1, 8\}$); and (6) removal of auxiliary pathology supervision. All implementations and baselines received the same number of video inputs per patient.

## 5 Results and Discussion

### 5.1 Architecture Comparison

Table 1 reports patient-level TB detection. **HMV-MIL achieves $\mathbf{0.89 \pm 0.02}$ AUC**, surpassing all conventional spatiotemporal baselines by 0.04-0.35 AUC:

Table 2: Ablation study. AUC mean ± SD across 5-fold cross-validation.

| Configuration | AUC |
|---|---|
| Full HMV-MIL | $\mathbf{0.89 \pm 0.02}$ |
| (1) Lightweight backbone (LeViT) | $0.53 \pm 0.05$ |
| (2) No LUS pre-training | $0.83 \pm 0.02$ |
| (3) Uniform frame sampling | $0.87 \pm 0.01$ |
| (4) Mean pooling | $0.88 \pm 0.02$ |
| (5) Top-$k$ selection ($k = 1$ / $k = 8$) | $0.87 \pm 0.03$ / $0.86 \pm 0.03$ |
| (6) No pathology supervision | $0.88 \pm 0.01$ |

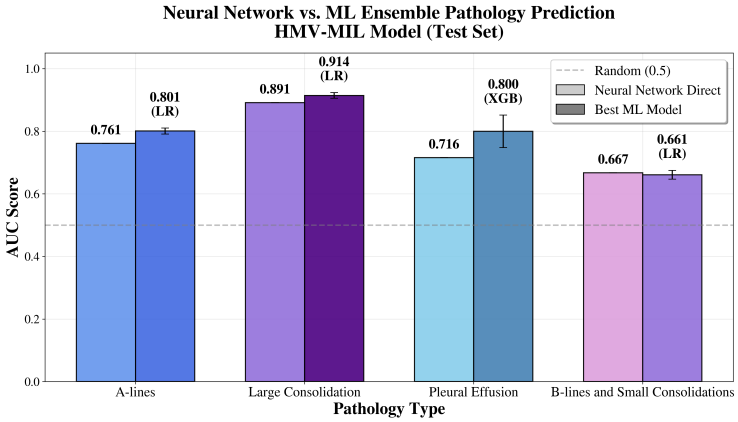


Fig. 3: Pathology classification: neural heads vs. ML ensembles (LR: logistic regression, XGB: XGBoost).

Video Transformer ($0.74 \pm 0.02$), CNN-LSTM ($0.64 \pm 0.05$), 3D ResNet-18 ($0.55 \pm 0.04$), and 3D-Inception ($0.56 \pm 0.04$). At the Youden-optimal operating point, HMV-MIL achieves 82% sensitivity and 75% specificity, approaching WHO triage targets (90%/70%). Fold-level AUC SD <0.02 demonstrates stable discrimination. Figure 2 shows ROC and precision–recall curves with 95% CIs.

DINOv3-BERT, despite having 418.6M parameters and being 4.5× larger than HMV-MIL, achieved an AUC of 0.85, which was 0.04 lower than HMV-MIL. This suggests that increasing backbone scale alone was insufficient in this setting, and is consistent with the hypothesis that TB lung ultrasound benefits from hierarchical aggregation. Conventional spatiotemporal architectures also underperformed, suggesting that dense frame-level encoding without selective temporal or site-level aggregation may be poorly matched to this task. One possible explanation is that these models process redundant frames, probe repositioning artifacts, normal anatomy, and motion noise together with sparse diagnostic findings, allowing weak TB-relevant signal to be diluted in global video representations.

### 5.2 Ablation Study

Table 2 isolates each component's contribution. **The domain-adapted CLIP backbone is the most critical element**: replacing it with LeViT drops AUC to $0.53 \pm 0.05$, and removing LUS pre-training yields $0.83 \pm 0.02$. This confirms that strong visual priors from domain-adapted pretraining underpin downstream temporal reasoning. In contrast, the frame selection strategy has modest impact: mean pooling ($0.88 \pm 0.02$), uniform sampling ($0.87 \pm 0.01$), $k = 1$ ($0.87 \pm 0.03$), and $k = 8$ ($0.86 \pm 0.03$) all fall within 0.01-0.02 AUC of the full model. Modest differences of this magnitude are expected near clinically relevant performance thresholds, where gains are increasingly constrained by noise and inter-patient heterogeneity; critically, all ablations degraded performance relative to the full model, confirming each component contributes. This suggests that diagnostically relevant information in LUS is concentrated in brief moments of optimal probe positioning, and that the attention mechanism successfully identifies these frames, mirroring expert sonographer workflow. Notably, $k = 1$ performance near-matches the full model, supporting a sparse-frame deployment strategy with reduced inference cost. Removing auxiliary pathology supervision yields $0.88 \pm 0.01$ AUC; while the AUC impact is modest, the primary value of pathology supervision is clinical interpretability, it produces per-site, human-readable pathology outputs enabling auditing and out-of-distribution flagging rather than serving primarily as an AUC lever.

### 5.3 Pathology Detection

Figure 3 summarizes auxiliary pathology head performance. Large consolidations achieve the highest discrimination (0.91 AUC, 0.60 AUPRC), followed by pleural effusion (0.80 AUC) and A-lines (0.80 AUC). B-lines/small consolidations remain most challenging (0.66 AUC) due to class overlap and lower prevalence. Figure 3 shows that ML ensembles (logistic regression, XGBoost) trained on pathology head outputs yield marginal improvements (0-0.08 AUC), particularly for low-prevalence findings, confirming that auxiliary heads capture complementary diagnostic cues. These pathology predictions also enable reliability auditing: e.g., consistent A-line predictions paired with a positive TB classification may flag out-of-distribution inputs.

### 5.4 Key Frame Analysis

Learned attention scores highlight diagnostically relevant frames, providing visual correspondence between model focus and clinical reasoning and supporting interpretable, auditable AI (Figure 4). This aligns with lung ultrasound practice, where diagnostically relevant information is often concentrated in brief moments of optimal probe positioning rather than distributed across extended sequences, mirroring expert sonographer workflow.

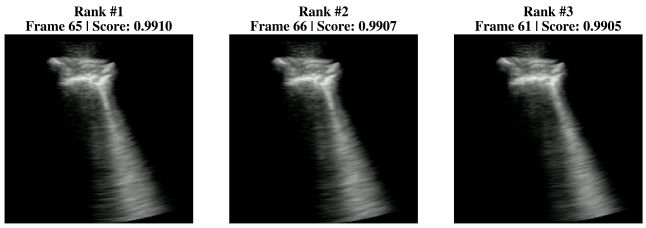


Fig. 4: Frames with highest attention weights. Visualization of the top-scoring frames ($s_t$) selected by the attention-pooling model at inference

### 5.5 Clinical Implications and Limitations

HMV-MIL is, to our knowledge, the first AI system capable of patient-level TB diagnosis from multi-video POCUS data, representing a meaningful step toward autonomous TB screening in settings where conventional diagnostics remain misaligned with the point of first healthcare contact. This matters at scale: over 3.8 million TB cases go undetected annually [19], disproportionately affecting communities served primarily by frontline health workers. Our work demonstrates that minimally trained healthcare workers, defined by less than 20 hours of training, can acquire data suitable for an AI-model to detect TB. Further our model can be deployed in a TF.js format at 344 MB suitable for consumer grade hardware.

At the Youden-optimal operating point, HMV-MIL achieves 82% sensitivity and 75% specificity. Table 3 reports the confusion matrix on the held-out test set ($n = 101$; 38 TB+, 63 TB−). PPV is 66% and NPV is 87%, reflecting the asymmetric class distribution and the model's stronger ability to rule out TB. For screening deployment where sensitivity is paramount, the operating threshold can be shifted along the ROC curve to achieve $\geq$90% sensitivity at reduced specificity. Acquisition quality affects model performance, and deployment should incorporate quality-control thresholds to flag low-quality inputs prior to inference.

Table 3: Confusion matrix on held-out test set at Youden-optimal threshold ($n = 101$).

| | **Predicted TB+** | **Predicted TB−** |
|---|---|---|
| **Actual TB+** ($n = 38$) | 31 (TP) | 7 (FN) |
| **Actual TB−** ($n = 63$) | 16 (FP) | 47 (TN) |

Limitations include single-site evaluation which constrains cross-population generalizability, and binary TB classification that does not account for co-infections. No publicly available LUS dataset for TB currently exists for external bench-

marking. An ongoing multi-country clinical trial will enable external validation, multi-device robustness assessment, real-time deployment optimization, and extension to other respiratory pathologies.

## 6 Conclusion

We presented HMV-MIL, a hierarchical multi-video MIL framework for patient-level TB detection from POCUS. By combining domain-adapted CLIP features, intra-video attention pooling, auxiliary pathology supervision, and inter-video gated MIL aggregation, HMV-MIL achieves 0.89 AUC on a 504-patient prospective cohort. This establishes the first AI system for patient-level TB diagnosis from multi-anatomical site ultrasound video data and a scalable foundation for equitable AI-guided screening in resource-limited settings. Code and implementation details are available here: https://github.com/tbrokowski/Ultr-AI-HMV-MIL.

**Acknowledgments.** We thank the patients whose participation made this work possible, and the clinical staff at CNHU-PPC, Cotonou, Benin for their support in data collection.

**Disclosure of Interests.** The authors have no competing interests to declare that are relevant to the content of this article.

## References


1. Abrokwa, S.K., Ruby, L.C., Heuvelings, C.C., Bélard, S.: Task shifting for point of care ultrasound in primary healthcare in low- and middle-income countries: a systematic review. EClinicalMedicine **45**, 101333 (2022). https://doi.org/10.1016/j.eclinm.2022.101333
2. Arntfield, R., VanBerlo, B., Alaifan, T., Phelps, N., White, M., Chaudhary, R., Ho, J., Wu, D.: Development of a convolutional neural network to differentiate among the etiology of similar appearing pathological b lines on lung ultrasound. BMJ Open **11**(3), e045120 (2021)
3. Aslan, G.: Diagnosing tuberculosis: traditional and new methods against an old enemy. Journal of Clinical Practice and Research **46**(5), 431 (2024)
4. Ebadi, A., Xi, P., MacLean, A., Tremblay, S., Kohli, S., Wong, A.: Covidx-us – an open-access benchmark dataset of ultrasound imaging data for ai-driven covid-19 analytics. arXiv preprint arXiv:2103.10003 (2021)
5. Goletti, D., Meintjes, G., Andrade, B.B., Zumla, A., Lee, S.S.: Insights from the 2024 who global tuberculosis report–more comprehensive action, innovation, and investments required for achieving who end tb goals. International Journal of Infectious Diseases **150** (2025)
6. Hong, D., Choi, H., Hong, W., et al.: Deep-learning model accurately classifies multi-label lung ultrasound findings. Scientific Reports **14**, 22228 (2024)
7. Horne, D.J., Kohli, M., Zifodya, J.S., Schiller, I., Dendukuri, N., Tollefson, D., Schumacher, S.G., Ochodo, E.A., Pai, M., Steingart, K.R.: Xpert mtb/rif and xpert mtb/rif ultra for pulmonary tuberculosis and rifampicin resistance in adults. Cochrane Database of Systematic Reviews **6**(6) (2019)
8. Ilse, M., Tomczak, J., Welling, M.: Attention-based deep multiple instance learning. arXiv preprint arXiv:1802.04712 (2018)
9. Kar, S., et al.: Same-day sputum microscopy: The road ahead in tuberculosis diagnosis. BMC Infectious Diseases **13**, 504 (2013). https://doi.org/10.1186/1471-2334-13-504
10. Lu, M.Y., Williamson, D.F., Chen, T.Y., Chen, R.J., Barbieri, M., Mahmood, F.: Data-efficient and weakly supervised computational pathology on whole-slide images. Nature Biomedical Engineering **5**(6), 555–570 (2021)
11. Lucassen, R.T., Jafari, M.H., Duggan, N.M., Jowkar, N., Mehrtash, A., Fischetti, C., Bernier, D., Prentice, K., Duhaime, E.P., Jin, M., et al.: Deep learning for detection and localization of b-lines in lung ultrasound. IEEE journal of biomedical and health informatics **27**(9), 4352–4361 (2023)
12. Ouyang, D., He, B., Ghorbani, A., Yuan, N., Ebinger, J., Langlotz, C.P., Heidenreich, P.A., Harrington, R.A., Liang, D.H., Ashley, E.A., et al.: Video-based ai for beat-to-beat assessment of cardiac function. Nature **580**(7802), 252–256 (2020)
13. Qin, Z.Z., Sander, M.S., Rai, B., Titahong, C.N., Sudrungrot, S., Laah, S.N., Adhikari, L.M., Carter, E.J., Puri, L., Codlin, A.J., et al.: Using artificial intelligence to read chest radiographs for tuberculosis detection: A multi-site evaluation of the diagnostic accuracy of three deep learning systems. Scientific reports **9**(1), 15739 (2019)
14. Roy, S., Menapace, W., Oei, S., Luijten, B., Fini, E., Saltori, C., Huijben, I., Chennakeshava, N., Mento, F., Sentelli, A., et al.: Deep learning for classification and localization of covid-19 markers in point-of-care lung ultrasound. IEEE Transactions on Medical Imaging **39**(8), 2676–2687 (2020)

15. Shao, Z., Bian, H., Chen, Y., Wang, Y., Zhang, J., Ji, X., Zhang, Y.: Transmil: Transformer based correlated multiple instance learning for whole slide image classification. In: Advances in Neural Information Processing Systems. vol. 34, pp. 2136–2147 (2021)
16. Shea, D.E., Kulhare, S., Millin, R., Laverriere, Z., Mehanian, C., Delahunt, B., Banik, D., Zheng, X., Zhu, M., Ji, Y., Ostbye, T., Mehanian, M.M.S., Uwajeh, A., Akinsete, A.M., Horning, M.P.: Deep learning video classification of lung ultrasound features associated with pneumonia. In: Proceedings of the CVPR 2023 Workshop on Deep Learning for Ultrasound Imaging and Applications (DL-UIA) (2023), open Access PDF
17. Stolz, L.A., Muruganandan, K., Bisanzo, M., Sebikali, M., Dreifuss, B., Hammerstedt, H., et al.: Point-of-care ultrasound education for non-physician clinicians in a resource-limited emergency department. Tropical Medicine & International Health **20**(8), 1067–1072 (2015). https://doi.org/10.1111/tmi.12511
18. Suttels, V., Brokowski, T., Wachinou, A.P., Wolleb, J., Hada, A.R., Du Toit, J.D., Fiogbe, A.A., Guendehou, B., Alovokpinhou, F., Sefou, F., et al.: Lung ultrasound for the detection of pulmonary tuberculosis using expert-and ai-guided interpretation: A prospective cohort study
19. World Health Organization: Global tuberculosis report 2024 (2024), https://iris.who.int/bitstream/handle/10665/379339/9789240101531-eng.pdf?sequence=1