Supplementary Information

Integrating Explainable Artificial Intelligence into Histopathological Risk Assessment: A Scoping Review and Meta-analysis

Bandar Alshreef; Yousif A. Kariri

Included in this upload-ready combined PDF
Supplementary File 1. PubMed search strategy
Supplementary File 5. Quality assessment tool used for study appraisal

Prepared as one combined supplementary document for Springer Nature / npj Digital Medicine submission.

Supplementary File 1. PubMed search strategy

Database searched: PubMed. Search window: database inception to November 2025. The syntax below was designed to capture studies on explainable AI in oncologic histopathology and was adapted as needed for Google Scholar, Scopus, and Web of Science.

Search string

```
(("explainable artificial intelligence"[Title/Abstract] OR "explainable AI"[Title/Abstract] OR
XAI[Title/Abstract] OR explainab*[Title/Abstract] OR interpretab*[Title/Abstract] OR "saliency
map*"[Title/Abstract] OR "class activation mapping"[Title/Abstract] OR CAM[Title/Abstract] OR
"Grad-CAM"[Title/Abstract] OR "gradient-weighted class activation mapping"[Title/Abstract] OR
"layer-wise relevance propagation"[Title/Abstract] OR LRP[Title/Abstract] OR SHAP[Title/Abstract] OR
LIME[Title/Abstract] OR "attention mechanism*"[Title/Abstract] OR "concept whitening"[Title/Abstract])
AND
(histopatholog*[Title/Abstract] OR patholog*[Title/Abstract] OR histolog*[Title/Abstract] OR "digital
pathology"[Title/Abstract] OR "whole slide image*"[Title/Abstract] OR "whole-slide
image*"[Title/Abstract] OR WSI[Title/Abstract])
AND
(cancer*[Title/Abstract] OR neoplasm*[Title/Abstract] OR tumor*[Title/Abstract] OR
tumour*[Title/Abstract] OR oncolog*[Title/Abstract] OR carcinoma*[Title/Abstract] OR
melanoma*[Title/Abstract] OR lymphoma*[Title/Abstract] OR adenocarcinoma*[Title/Abstract])
AND
("risk assessment"[Title/Abstract] OR prognos*[Title/Abstract] OR grading[Title/Abstract] OR
staging[Title/Abstract] OR stratif*[Title/Abstract] OR classif*[Title/Abstract] OR
diagnos*[Title/Abstract] OR detect*[Title/Abstract]))
AND english[Language]
```

Notes

• The review also searched Google Scholar, Scopus, and Web of Science; equivalent search terms were adapted to the syntax of each platform.

• Eligibility was restricted to peer-reviewed English-language primary research in oncologic histopathology applying XAI methods to risk-relevant tasks such as diagnosis, grading, prognosis, stratification, staging, or related classification tasks.

• Conference abstracts, editorials, reviews, and non-oncology studies were excluded during screening.

Supplementary File 5. Quality assessment tool used for study appraisal

The included studies were appraised using a structured diagnostic-AI quality assessment framework tailored to explainable AI in histopathology. Each item was reviewed as **Yes**, **No**, or **Unclear**. Overall study appraisal was assigned holistically as **low**, **moderate**, or **high** risk of bias, with particular attention to patient selection, independent validation, and transparency of train/test splitting.

Key concerns contributing to higher risk of bias in this review included non-sequential case selection, absence of an independent validation dataset, and unclear description of model splitting or aggregation from slide-level to patient-level analysis.

Domain	Assessment item	Response options
1. Study design and patient selection	Were the study population and oncologic histopathology task clearly defined?	Yes / No / Unclear
1. Study design and patient selection	Were inclusion and exclusion criteria reported?	Yes / No / Unclear
1. Study design and patient selection	Were cases consecutive, representative, or otherwise selected in a way that minimized selection bias?	Yes / No / Unclear
2. Reference standard and outcomes	Was the reference standard (e.g., expert pathologist annotation, molecular label, or clinical outcome) clearly described?	Yes / No / Unclear
2. Reference standard and outcomes	Was the prediction target clinically relevant to diagnosis, grading, prognosis, staging, or risk stratification?	Yes / No / Unclear
2. Reference standard and outcomes	Was the unit of analysis (tile, patch, slide, or patient) explicitly stated?	Yes / No / Unclear
3. Data handling and model development	Were the training, validation, and test splits described clearly?	Yes / No / Unclear
3. Data handling and model development	Was patient-level separation or leakage prevention addressed where applicable?	Yes / No / Unclear
3. Data handling and model development	Were the AI model architecture and XAI method sufficiently described for reproducibility?	Yes / No / Unclear
4. Validation and performance reporting	Was an internal validation strategy reported?	Yes / No / Unclear
4. Validation and performance reporting	Was an independent external or held-out validation dataset used?	Yes / No / Unclear
4. Validation and performance reporting	Were discrimination metrics (e.g., AUC) and sufficient performance details reported?	Yes / No / Unclear
4. Validation and performance reporting	Where slide-level outputs informed patient-level claims, was aggregation methodology described?	Yes / No / Unclear
5. Explainability and clinical relevance	Was the rationale for the selected XAI technique stated?	Yes / No / Unclear

Domain	Assessment item	Response options
5. Explainability and clinical relevance	Did the explanation outputs correspond to pathologist-relevant morphologic features or regions of interest?	Yes / No / Unclear
5. Explainability and clinical relevance	Were explanation quality, usability, or sanity checks evaluated?	Yes / No / Unclear
6. Reproducibility and governance	Were dataset provenance and preprocessing sufficiently reported?	Yes / No / Unclear
6. Reproducibility and governance	Did the authors discuss reproducibility, deployment risk, reporting guidance, or quality/risk frameworks (e.g., CONSORT-AI, TRIPOD-AI, ISO 15189, ISO 14971)?	Yes / No / Unclear

Overall appraisal categories reported in the manuscript were low risk of bias, moderate risk of bias, and high risk of bias. The final judgement was based on the pattern and seriousness of concerns across the domains above rather than on a single item alone.