# AI and Digital Pathology in Gynaecological Histopathology

BAGP e-Poster 2026

## A Rapid Scoping Review of Clinical Tasks, Performance and Readiness

Roshan Faiz | Bank FY3, Airedale General Hospital, West Yorkshire, UK | Corresponding author

**15** Primary studies included

**3** Gynaecological tissue sites

**5** Clinical task categories

**0.933** Prospective AUC (Sn 0.934 · Sp 0.837)

**10** Years reviewed 2016 – 2026

## BACKGROUND

Gynaecological cancers **(cervix, uterus/endometrium, ovary)** represent a major global cancer burden. Histopathological assessment of biopsy and resection specimens remains the definitive basis for diagnosis, subtype classification, and staging — directly informing surgical and systemic treatment decisions.

Whole-slide imaging (WSI) infrastructure has created the technical foundation for computational pathology at scale. AI — particularly deep learning — offers potential to standardise diagnosis, reduce inter-observer variability, improve throughput, and extract prognostic information directly from H&E morphology.

*No prior review had systematically mapped AI applied to whole-slide digital pathology within gynaecological histopathology, leaving clinicians without a clear evidence synthesis.*

## AIMS

To map primary studies applying AI to whole-slide digital pathology across gynaecological histopathology (cervix, uterus/endometrium, ovary); to characterise clinical tasks, dataset sizes, AI architectures, and validation designs; and to summarise headline performance metrics to assess clinical readiness.

## METHODS

| | |
|---|---|
| **Design** | Rapid scoping review. Primary studies systematically identified, screened, and charted. Evidence narratively synthesised by clinical task category and assessed for readiness level rather than pooled statistically. |
| **Search** | PubMed and Europe PMC searched from 1 January 2016 to 11 February 2026. English-language publications only. Embase and Scopus were not searched — a recognised limitation. |
| **Inclusion** | Primary studies applying AI to whole-slide images (H&E ± IHC) from cervix, uterus/endometrium, or ovary. Required a diagnostic, subtype, molecular, or prognostic task and ≥1 quantitative metric (AUC, sensitivity, specificity, C-index). |
| **Exclusion** | Reviews/editorials; cytology-only studies; radiology or omics studies without WSI; purely technical papers without a gynaecological diagnostic or prognostic task. |
| **Extraction** | Clinical site, stain type, AI architecture (patch CNN / MIL / transformer), dataset size, validation design (internal / external / prospective), and headline performance metrics. |
| **Synthesis** | Evidence mapped by AI task category. Narrative synthesis of clinical readiness aligned to RCPath guidance on AI in pathology. No formal meta-analysis due to metric heterogeneity across studies. |

## STUDY SELECTION

1. **Database search** — PubMed & Europe PMC | 1 Jan 2016 - 11 Feb 2026
2. **Title & abstract screen** — Gynaecological site + WSI + AI task
   - x Non-gynaecological; no WSI; reviews; non-English
3. **Full-text review** — PMC full texts prioritised
   - x Reviews/editorials; cytology-only; no quantitative metric
4. **15 studies included** — Cervix n=4 | Endometrium n=8 | Ovary n=3

## REFERENCES


1. BAGP. Poster Competition Guidelines. 2026.
2. Wang W et al. PMC9668742 — Clinical-grade EC detection.
3. Zheng Y et al. PMC11034655 — Cervical WSI classifier.
4. Dawood Z et al. PMC8356709 — DeepCIN CIN grading.
5. Xu J et al. PMC10354251 — CST p16 robustness.
6. Williams M et al. PMC12691344 — EagleEye CIN2+.
7. Verweij SL et al. PMC9689570 — Uterine origin classifier.
8. Wang Y et al. PMC9994759 — Endometrial biopsy triage.
9. Guo Z et al. PMC8484685 — Panoptes CNN.
10. Volinsky-Fremond et al. PMID:36496303 — im4MEC.
11. Volinsky-Fremond et al. PMID:38789645 — HECTOR.
12. Saillard C et al. PMC11137095 — MSI prediction.
13. Breen J et al. PMC12147852 — MMRNet dMMR.
14. Farinella R et al. PMID:36065012 — Ovarian histotype.
15. Tong L et al. PMC11782474 — Foundation MIL ovary.
16. Nobre L et al. PMC11102549 — PathoRiCH HGSOC.
17. RCPath. Position Statement on AI in Pathology. 2022.

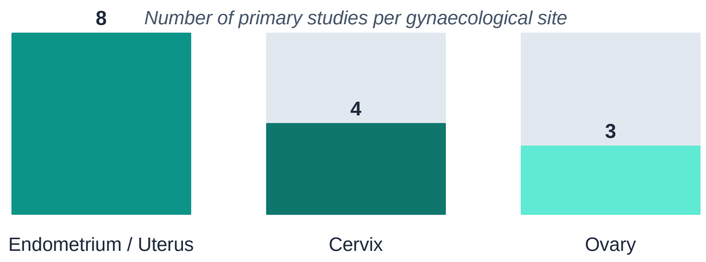

## RESULTS

**WSI-to-AI Clinical Pipeline — 5 task categories across all 15 studies**

① Tissue Biopsy — Cervix/Endo/Ovary → ② H&E / IHC Staining — ± Immunohistochem. → ③ Digital WSI Scan — Whole-slide image → ④ AI Model CNN·MIL·Transf — Patch→slide aggreg. → ⑤ Clinical Output — Triage·Class·Mol·Prog

## STUDIES BY TISSUE SITE (n=15)

*Number of primary studies per gynaecological site*




## CLINICAL READINESS MATRIX

*Evidence by task category and validation design (n=15 studies)*

| | Internal only | External multi-centre | Prospective ★ |
|---|---|---|---|
| **Triage / Detection** (Endometrium) | n=3 AUC up to 0.933 | — | n=1 ★ AUC 0.933 |
| **Classification / Subtyping** (All sites) | n=4 AUROC up to 0.977 | n=1 Bal. acc. 97% | — |
| **Molecular Surrogates** (Endometrium) | n=3 AUROC 0.827–0.969 | n=1 AUROC 0.790–0.863 | — |
| **Prognosis / Response** (Endo/Ovary) | n=1 C-index ~0.8 | n=1 C-index ~0.8 | — |

★ Only 1 of 15 studies achieved prospective clinical validation — endometrial cancer detection (AUC 0.933).

## ENDOMETRIAL & UTERINE FINDINGS (n=8)

**90%** Overall slide accuracy 2,909 slides · triage

**97%** Malignant-class accuracy Endometrial biopsy triage

**0.933** Prospective AUC Sn 0.934 · Sp 0.837

**Clinical-grade EC detection:** Prospectively validated in 317 consecutive cases; multi-centre AUC 0.971. Strongest translational evidence in this review.

**Molecular surrogates (Panoptes):** Histology AUROC 0.969; MSI-high AUROC 0.827; TP53 AUROC 0.873. Directly relevant to WHO 2020 / ProMisE classification frameworks.

**dMMR prediction (MMRNet):** Internal AUROC 0.897; external AUROCs 0.790 / 0.807 / 0.863 across 3 independent datasets — most methodologically robust molecular model.

**Molecular class (im4MEC):** Self-supervised attention model, 2,028 patients: macro-AUROC 0.874 (CV) / 0.876 (test). TCGA-style molecular class from H&E alone.

**Prognosis (HECTOR):** Multimodal deep learning for distant recurrence, 2,072 patients, 8 cohorts: C-index ~0.8 internal & external — but no prospective workflow evaluation exists.

**Subtype / origin classifier:** Uterine site-of-origin and subtype (TCGA/CPTAC datasets): AUROC up to 0.977.

## CERVICAL FINDINGS (n=4)

**93.3%** CIN2+ sensitivity EagleEye · 56/60 cases

**71.8%** CIN2+ specificity EagleEye · 28/39 cases

**CIN grading (DeepCIN):** Exact-class accuracy 85.0%, AUC 95.5; off-by-one accuracy 96.7% (150 WSIs + 947 labelled images). Single-scanner domain.

**Cervical WSI classifier:** Patch+XGBoost (1,738 train / 811 test WSIs): slide accuracy 71.53%; malignant sensitivity 93.4%.

**p16 IHC robustness (CST):** AUC 0.940–0.989 (CST) vs 0.905–0.986 (standard) across 6 test sets — demonstrates domain generalisation across scanner and staining variability.

## OVARIAN FINDINGS (n=3)

| Study | Task | Dataset | Key Metric |
|---|---|---|---|
| Farinella et al. | Histotype CNN | 948 train; 60 ext. | Concordance 80.97% (κ 0.75) |
| Tong et al. | Foundation MIL subtype | 1,864 + 2 ext. | Bal. acc. 89% int.; 97/74% ext. |
| Nobre et al. | PathoRiCH HGSOC response | 394 train; 420 ext. | AUC ~0.60 int. and ext. |

## VALIDATION LANDSCAPE (n=15 studies)

- **~12** **Internal validation only** — All three tissue sites
- **2** **External multi-centre** — Endometrium — HECTOR, MMRNet
- **1** **Prospective clinical** — Endometrial cancer detection only

## DISCUSSION

**Endometrial triage and detection show the clearest clinical readiness,** with the strongest evidence base and the only prospective clinical validation identified in this review.

**Cervical, ovarian, and molecular prediction applications remain promising,** but current studies rely more heavily on internal validation, smaller cohorts, and limited workflow-level evidence.

**The key gap between technical performance and clinical readiness** is the need for multi-centre external validation, prospective workflow evaluation, and robust governance and quality assurance before wider implementation.

## CONCLUSIONS

*AI in gynaecological digital pathology is closest to clinical deployment in endometrial triage and detection — but broader adoption remains limited by external validation, workflow evidence, and governance gaps.*

1. **Multi-centre external validation** Particularly for cervical biopsy decision support and molecular surrogate prediction — to establish robustness across scanner platforms and diverse populations.
2. **Prospective workflow impact evaluation** Currently absent for all task categories except endometrial cancer detection. Essential before clinical adoption.
3. **Governance and quality assurance** Rigorous pre-deployment validation and ongoing post-market performance monitoring — aligned to RCPath and NHS digital pathology governance frameworks.

## LIMITATIONS

| | |
|---|---|
| **Methodology** | Rapid scoping review; not formally registered with a protocol. |
| **Search scope** | PubMed & Europe PMC only — Embase and Scopus not searched; English-language only. Publication and language bias cannot be excluded. |
| **Hit counts** | Exact database hit counts not captured within this interface without risking inaccurate reporting. |
| **Heterogeneity** | Variation in performance metrics and task definitions across studies precluded formal meta-analysis. |
| **Small cohorts** | Ovarian histotype external validation (n=60 patients) and cervical CIN2+ (n=99 biopsies) limit the strength of conclusions for these specific tasks. |

*Declaration of competing interests: All authors confirm this work is not affiliated with any commercial sponsor or external organisation that could affect independence or objectivity.*