Supplementary Tables

Table 1: Proposed evaluation domains for AI-defined molecular characteristic prediction from histopathology images. Adequacy of external validation was pragmatically defined as ≥100 cases, reflecting commonly used rule-of-thumb thresholds in prediction model research. However, formal methodological guidance emphasises that appropriate sample size should be determined based on the precision of performance estimates and outcome prevalence, rather than fixed numerical cut-offs.

Domain	Item	Explanation	Rationale
1. Clinical intent & molecular definition (CONSORT AI 2a, 6a; SPIRIT AI 2, 7; CLAIM 3, 4, 14–16)	Target	Biological endpoint predicted	Clinical relevance
	Reference test	Gold standard comparator	Important for validity
	Target definition type	Binary (yes/no), continuous score, or outcome	Interpretability/comparability
	Intended clinical role	Whether model is designed for triage, decision support, or replacement	Clinical context
	Timing of reference test	Whether labels reflect pre-treatment biology or post-hoc outcomes	Causal interpretation
2. Dataset composition & provenance (CONSORT AI 4a, 5; SPIRIT AI 10, 14; CLAIM 7–10, 19–21, 35–36)	# Whole slide images (WSI)	Scale of data used	Statistical power
	Histotype restriction	Whether restricted to HGSOC or mixed OC	Generalisability
	Training dataset(s)	Source cohorts (*e.g.*, TCGA, institutional)	Bias and representativeness
	Training country	Geographic origin	Equity and domain shift
	External testing dataset(s)	Independent cohorts used for testing	Key for translation
	External testing country	Assesses geographic generalizability	Generalisability
	Outcome linkage	Whether molecular predictions are linked to survival or treatment response	Clinical relevance
3. Pre-analytical & imaging characteristics (CONSORT AI 4b; SPIRIT AI 11; CLAIM 9, 10, 13)	Scanner	Type of slide scanner	Systematic variability
	Magnification	Scale of image analysis (*e.g.*, 10X, 20X)	Systematic variability
	Resolution	Pixel-level detail	Reproducibility
	Staining protocol/normalisation	Type of stain protocol; whether stain variability is corrected; critical for cross-site deployment	Systematic variability
	Multi-scanner evaluation	Whether the model is tested across different scanners	Robustness

Table 1: (continued from last page)

Domain	Item	Explanation	Rationale
	Laboratory heterogeneity	Whether authors recognise inter-lab variation	Systematic variability
	Tile/patch extraction strategy	Definition of tile size, sampling method, and extraction strategy used to generate model inputs from WSIs	Reproducibility
4. Model evaluation & performance (CONSORT AI 10, 12, 17; SPIRIT AI 12–15; CLAIM 28–33, 37–39)	Internal AUC	Model performance on training/validation	Model evaluation
	External AUC	Generalisation performance on independent cohort(s)	Generalisability
	Uncertainty reporting	Difference between internal and external AUC; Confidence intervals for AUC; proxy for overfitting	Generalisability
	External testing $\geq$100 patients	Minimum robustness threshold	Statistical power
	Calibration report	Whether probability outputs are reliable (not just discriminative)	Clinical reliability
5. Governance, deployment & implementation (CONSORT AI 5, 21, 22; SPIRIT AI 16, 18; CLAIM 11, 42–44)	Human-AI interaction	Whether clinician involvement is defined	Real-world implications
	Proposed deployment setting	Research-only vs clinical pathway (CE-mark direction)	Translational maturity
	Data governance/privacy	Ethical and regulatory considerations	Clinical implementation
	Training paradigm	Centralised vs federated training	Scalability and equity
	Open-source code	Availability of implementation	Reproducibility and adoption

Table 2: Characteristics of included studies evaluating AI-based prediction of HRD from H&E slides in ovarian cancer, including model architecture, intended clinical role, histotype restriction, clinical outcome linkage, and external validation set size (WSIs). HGSOC=High-grade serous ovarian cancer.

Study	**Year**	**Model**	**Intended clinical role**	**Histotype restriction**	**Clinical outcome linkage available**	**# WSI**
Lazard *et al.*	2022	Convolutional neural network (ResNet-18) pretrained using Momentum Contrast (MoCo) self-supervised learning, followed by a multilayer perceptron classifier	Not explicit	Mixed	No	90
Ahn *et al.*	2024	PathoRiCH pipeline: tumour segmentation, feature extraction using ResNet-18 pretrained with SimCLR self-supervised learning, and aggregation using dual stream multiple instance learning (DSMIL)	Decision-support	HGSOC	Yes	-
Bergstrom *et al.*	2024	DeepHRD model: ResNet-18 feature extractor with average pooling aggregation across image patches	Decision-support (implied)	Mixed	Yes	-
Loeffler *et al.*	2024	ResNet-50 feature extraction with random patch sampling and attention-based aggregation for slide-level prediction	Not explicit	Mixed	No	-
Marmé *et al.*	2025	Multi-stage pipeline: tissue annotation and tumour segmentation, ResNet-18 (Camelyon trained) feature extraction, patch merging module, transformer-based aggregation, and multilayer perceptron classifier	Decision-support	Ovarian (unspecified)	Yes	447
Wu *et al.*	2025	HRDPath model: ResNet-50 feature extraction using CLAM (clustering constrained attention multiple instance learning), followed by vision transformer (ViT) modelling and DSMIL aggregation at the patient level	Decision-support (implied)	HGSOC	No	81+198

Table 2: (continued from last page)

Study	**Year**	**Model**	**Intended clinical role**	**Histotype restriction**	**Clinical outcome linkage available**	**# WSI**
Zhang *et al.*	2025	Structured pipeline: tumour segmentation using U-Net++, cell segmentation using HoverNet, cell-level feature extraction, and classification using a random forest model	Not explicit	Ovarian (unspecified)	No	22
Frenel *et al.*	2026	DiaThera: tumour detection with Deep Histopath, Vision transformer (DINOv2) for feature extraction and MIL-attention aggregation	Decision-support	HGSOC	Yes	422+89
Hada *et al.*	2026	Handcrafted nuclei feature extraction followed by a conventional machine learning classifier	Not explicit	HGSOC	No	-
Uttarwar *et al.*	2026	TRINITY model: multimodal contrastive learning framework with whole slide image embeddings (H-Optimus model), followed by gradient boosting classification	Decision-support (implied)	Mixed	No	74
Yang *et al.*	2026	dPathHRD model: transformer-based architecture with attention-based multiple-instance learning for slide-level prediction	Decision support	HGSOC	Yes	62+156
Zafar *et al.*	2026	IHGAMP framework: pretrained vision transformer (foundation model) for feature extraction, followed by ridge regression classifier	Screening/triage (implied)	Mixed	Yes	-

Table 3: Extended reporting characteristics for AI-defined HRD prediction models. Reporting of pre-analytical, imaging, and validation characteristics across included studies. NR= not reported for specific technical parameters; * indicates values inferred or partially described in the original study.

Study	**Colour/stain normalisation**	**Multi-scanner**	**Lab heterogeneity**	**External testing ≥100**	**Calibration**	**Scanner**	**Magnification**	**Resolution**
Lazard *et al.* 2022	NR	No	No	No	No	Aperio*	20X*	$\approx 0.5\mu m$/px*
Ahn *et al.* 2024	NR	No	No	No	No	NR	NR	NR
Bergstrom *et al.* 2024	Yes	No	No	No	No	NR	NR	NR
Loeffler *et al.* 2024	Yes	No	No	No	No	NR	NR	NR
Marmé *et al.* 2025	NR	No	No	Yes	No	NR	NR	NR
Wu *et al.* 2025	Yes	Yes	No	Yes	No	NR + Aperio*	NR + 20X*	NR + $\approx 0.5\mu m$/px*
Zhang *et al.* 2025	NR	No	No	No	No	NR	NR	NR
Frenel *et al.* 2026	NR	Yes	No	Yes	No	NR + Aperio*	NR + 20X*	NR + $\approx 0.5\mu m$/px*
Hada *et al.* 2026	NR	No	No	No	No	NR	NR	NR
Uttarwar *et al.* 2026	No	No	No	No	No	NR	NR	NR
Yang *et al.* 2026	NR	No	No	Yes	No	NR	NR	NR
Zafar *et al.* 2026	NR	No	No	No	No	NR	NR	NR

Table 4: Characteristics of included studies evaluating AI-based prediction of BRCA1/2 status from H&E slides in ovarian cancer, including model architecture, intended clinical role, histotype restriction, clinical outcome linkage, and external validation set size (WSIs).

Study	**Year**	**Model**	**Intended clinical role**	**Histotype restriction**	**Clinical outcome linkage available**	**# WSI**
Zeng *et al.*	2021	Hand-engineered image features extracted using CellProfiler, followed by conventional machine learning classifiers (*e.g.*, random forest, gradient boosting models)	Prognostic/ decision-support	HGSOC	Yes	-
Nero *et al.*	2022	Pipeline using tumour region annotation, feature extraction with a convolutional neural network (ResNet 50), and aggregation using attention-based and clustering-based multiple instance learning (CLAM)	Not explicit	Epithelial OC	Yes -	
Bourgade *et al.*	2023	Tumour segmentation using EfficientNet-B4, followed by random patch sampling, feature extraction with ResNet-50 pretrained using Momentum Contrast (MoCo) self-supervised learning, and attention-based multiple instance learning for slide-level prediction	Decision-support (implied)	HGSOC	No	103
Ho *et al.*	2023	Tumour segmentation followed by feature extraction using ResNet-18 and aggregation using average pooling across image patches	Exploratory	HGSOC	No	-
Li *et al.*	2024	Multimodal pipeline: tumour annotation and attention-based multiple instance learning (MIAM), combined with cell-level features and clinical variables, followed by dimensionality reduction using principal component analysis (PCA) and classification using a random forest model	Decision-support (explicit)	Mixed	Yes	105+130

Table 4: (continued from last page)

Study	Year	Model	Intended clinical role	Histotype restriction	Clinical outcome linkage available	# WSI
Xu *et al.*	2026	Multi-magnification image sampling with attention mechanisms, feature extraction using ResNet-50, and integration of features across scales using cascade fusion	Screening/triage (implied)	Ovarian (unspecified)	No	-

Table 5: Extended reporting characteristics for AI-defined BRCA1/2 status prediction models. Reporting of pre-analytical, imaging, and validation characteristics across included studies. NR= not reported for specific technical parameters; * indicates values inferred or partially described in the original study.

Study	Colour/stain normalisation	Multi-scanner	Lab heterogeneity	External testing $\geq$100	Calibration	Scanner	Magnification	Resolution
Zeng *et al.* 2021	NR	No	No	No	No	NR	NR	NR
Nero *et al.* 2022	NR	No	No	No	No	NR	NR	NR
Bourgade *et al.* 2023	Yes	No	No	Yes	No	Aperio*	20X*	$\approx 0.5\mu m$/px*
Ho *et al.* 2023	NR	No	No	No	No	NR	NR	NR
Li *et al.* 2024	Yes	Yes	Yes	Yes	No	Aperio* + KFBio KF-PRO-005-HI	20X* + 40X	$\approx 0.5\mu m$/px*
Xu *et al.* 2026	NR	No	No	No	No	NR	NR	NR

Table 6: Traffic light assessment of published AI-defined HRD and BRCA1/2 status prediction models mapped against the five-domain evaluation framework. Green indicates fully reported items, amber partially reported or inferred items, and red items not reported in the source publication.

Target	Study	Year	Domain 1: Reference test	Domain 2: Dataset & validation	Domain 3: Pre-analytics	Domain 4: Model & performance	Domain 5: Governance & deployment
HRD	Lazard *et al.*	2022					
	Ahn *et al.*	2024					
	Bergstrom *et al.*	2024					
	Loeffler *et al.*	2024					
	Marmé *et al.*	2025					
	Wu *et al.*	2025					
	Zhang *et al.*	2025					
	Hada *et al.*	2026					
	Uttarwar *et al.*	2026					
	Yang *et al.*	2026					
	Zafar *et al.*	2026					
	Frenel *et al.*	2026					
BRCA	Zeng *et al.*	2021					
	Nero *et al.*	2022					
	Bourgade *et al.*	2023					
	Ho *et al.*	2023					
	Li *et al.*	2024					
	Xu *et al.*	2026					

Table 7: Traffic-light assessment summarising risk of bias across four QUADAS-2 domains: patient selection, index test, reference standard, and flow and timing. Each domain was rated as low risk (green), unclear risk (yellow), or high risk (red) based on predefined criteria from the QUADAS-AI-P framework.

Target	Study	Year	Patient selection	Image selection	Index test	Reference standard	Flow & timing	Overall
HRD	Lazard *et al.*	2022						
	Ahn *et al.*	2024						
	Bergstrom *et al.*	2024						
	Loeffler *et al.*	2024						
	Marmé *et al.*	2025						
	Wu *et al.*	2025						
	Zhang *et al.*	2025						
	Hada *et al.*	2026						
	Uttarwar *et al.*	2026						
	Yang *et al.*	2026						
	Zafar *et al.*	2026						
	Frenel *et al.*	2026						
BRCA	Zeng *et al.*	2021						
	Nero *et al.*	2022						
	Bourgade *et al.*	2023						
	Ho *et al.*	2023						
	Li *et al.*	2024						
	Xu *et al.*	2026						