Getting Started with AI in Paediatric and Perinatal Pathology

Path-iQ Getting Started Series · No. 15 of 20 · A practical primer for the practising pathologist

July 2026

1. Why children are not small adults, algorithmically

Paediatric and perinatal pathology faces a specific and under-discussed AI problem: nearly every pathology foundation model and nearly every commercial algorithm has been trained on adult tissue. Adult tissue is what fills biobanks and hospital archives.

This matters more than it might appear. Paediatric tumours are biologically distinct entities, not paediatric versions of adult cancers. Normal paediatric tissue differs from adult tissue in cellularity, maturation and proliferative activity — normal paediatric bone marrow is hypercellular by adult standards, normal thymus is large and active, normal paediatric lymph nodes show florid reactive changes. A model calibrated on adults will systematically over-call abnormality in children.

The rule for this subspecialty is therefore unusually firm: assume no adult-trained tool is valid for paediatric use until demonstrated otherwise, and say so explicitly in any procurement conversation.

2. Where the field has moved fastest: paediatric neuro-oncology

The most substantial AI achievement in paediatric pathology comes from molecular classification. Sturm and colleagues demonstrated in *Nature Medicine* in 2023 that multiomic neuropathology improves diagnostic accuracy in paediatric neuro-oncology, and the prospective, population-based Molecular Neuropathology 2.0 study — over 1,200 newly diagnosed paediatric CNS tumour patients with blinded central neuropathological review alongside molecular testing — showed that combining DNA methylation profiling with targeted panel sequencing improved classification accuracy and resolved cases where histology was ambiguous.

The current Heidelberg methylation classifier, version 12.8, recognises 184 subclasses and demonstrated distinct survival curves across methylation subclasses in paediatric ependymoma and medulloblastoma cohorts. For paediatric neuro-oncology, machine-learning-based molecular classification is not emerging technology; it is standard of care in well-resourced settings.

3. The task map

Placental pathology. The largest-volume specimen in perinatal practice and the most under-resourced. Assessment of villous maturation, infarction, chorioamnionitis, maternal and fetal vascular malperfusion follows the structured Amsterdam consensus criteria — a defined framework that automation can target. Placental examination is frequently deferred or abbreviated due to workload; automated screening of placentas that would otherwise go unexamined is a genuine access argument rather than an efficiency one.

Small round blue cell tumours. Neuroblastoma, Ewing sarcoma, rhabdomyosarcoma, lymphoblastic lymphoma and Wilms tumour. Morphologically overlapping and distinguished by immunohistochemistry and molecular testing. Neuroblastoma specifically has an established histological classification incorporating stromal content, degree of differentiation and mitosis-karyorrhexis index — the last being a counting task with poor reproducibility and direct risk-stratification consequences.

Paediatric bone marrow. Discussed in the hematopathology primer; the haematogone-versus-blast problem is the defining paediatric-specific failure mode.

Perinatal autopsy. Declining consent rates and a shrinking workforce make this an area where decision support and standardised measurement have value, including automated organ weight interpretation against gestational-age norms.

Congenital and metabolic disease. Rare, morphologically subtle, and dependent on electron microscopy and biochemical correlation. Realistically a decision-support rather than automation target.

4. Your first 90 days

Weeks 1–3: start with the placenta. It is the highest-volume, most standardised, most under-examined specimen you handle, and the Amsterdam criteria give you an agreed scoring framework. Begin by auditing what proportion of placentas submitted actually receive full examination, and what your turnaround is.

Weeks 4–8: audit your molecular access for paediatric CNS and solid tumours. As in neuropathology, the highest-value intervention is usually expanding access to methylation classification and panel sequencing rather than deploying an image algorithm.

Weeks 9–12: for any tool under consideration, obtain the age distribution of the training data in writing. If the vendor cannot provide it, that is your answer.

5. Validation specifics

Age stratification must be explicit. Neonatal, infant, child and adolescent tissue differ. A tool validated in “paediatric” patients aged 12–17 tells you little about a neonate.

Gestational age governs interpretation of every perinatal specimen. Placental and fetal organ morphology is only interpretable against gestational-age-specific norms, and any model must incorporate this.

Autolysis is routine in perinatal specimens. Intrauterine fetal death specimens show maceration that is expected, not artefactual, and models trained on fresh surgical tissue will not handle it.

Specimen size is frequently minute. Paediatric biopsies are small by necessity, and models trained on adult-sized specimens may have effective minimum tissue requirements that your material does not meet.

6. Failure modes to anticipate

Normal paediatric hypercellularity read as abnormal. The single most predictable failure. Applies to marrow, thymus, lymph node and liver.

Haematogones called blasts. Documented for both flow gating and morphology classifiers.

Nephrogenic rests and other developmental remnants. Normal or near-normal embryonal remnants that resemble tumour.

Extramedullary haematopoiesis. Normal in neonatal liver and other sites; misread as infiltrate by adult-trained models.

Rarity compounding. Paediatric tumours are rare, and paediatric tumour subtypes rarer still. Abstention behaviour is essential.

7. What to watch next

The most important development would be the deliberate construction of paediatric-specific training corpora and foundation models. This is a collective-action problem — no single centre has the volume — and it will require the kind of international consortium that built the paediatric neuro-oncology methylation reference sets. Watch also for placental AI reaching deployment, since it is the application where AI could plausibly extend examination to specimens that currently receive none, which is a different and larger claim than improving accuracy on specimens already examined.

Vendor due-diligence checklist

Use this before any procurement conversation. The first seven questions apply to every AI tool in pathology; the last three are specific to this subspecialty. A vendor unable or unwilling to answer them in writing has told you something important.

Applies to any pathology AI tool

What is the intended use statement, and what is explicitly excluded from it?

What is the regulatory status in my jurisdiction — cleared, CE-marked, research use only?

Can I see performance on an independent external cohort, not only internal validation?

What is the prevalence of the target condition in the validation cohort versus my practice?

What is the failure mode when the model is uncertain — does it abstain, or always answer?

What monitoring does the vendor provide for performance drift after deployment?

Who is liable if an algorithmic output contributes to a diagnostic error?

Specific to this subspecialty

What is the age distribution of the training population, in writing?

Does the model incorporate gestational-age-specific norms for perinatal specimens?

How does it handle autolysed and macerated tissue from intrauterine fetal death?

Ask for answers in writing, and keep them. When the tool underperforms in your laboratory — and at some point it will — the gap between what was claimed and what was validated is where the conversation starts.

8. Key references

#

Source

Note

1

Sturm et al., *Nature Medicine*, 2023;29:917–926

Multiomic neuropathology in paediatric neuro-oncology

2

Heidelberg CNS Tumour Methylation Classifier v12.8, *Cancer Cell*, 2025

184 subclasses; MNP 2.0 cohort >1,200 paediatric patients

3

Pickles et al., *Lancet Child Adolesc Health*, 2020;4:121–130

Methylation profiling for paediatric CNS tumour diagnosis

4

Amsterdam Placental Workshop Group consensus criteria

Standardised placental scoring framework

5

International Neuroblastoma Pathology Classification

Mitosis-karyorrhexis index and risk stratification

Glossary of terms used in this primer

Whole-slide image (WSI) — a digitised glass slide, typically gigapixel in scale, produced by a scanner and viewed on screen rather than under a microscope.

Foundation model — a large model pretrained on vast unlabelled image data to learn general-purpose visual features, then adapted to specific tasks with comparatively little labelled data.

Weakly supervised learning — training where only a slide-level label is available (for example “contains cancer”) rather than region-by-region annotation. Most clinical pathology AI is built this way, because exhaustive annotation is unaffordable.

AUROC — area under the receiver operating characteristic curve, a single figure summarising discrimination across all thresholds. 0.5 is chance, 1.0 is perfect. It says nothing about calibration or about performance at the threshold you would actually use.

Calibration — whether a model’s stated probability matches observed frequency. A well-calibrated model that says “80% confident” is right about 80% of the time. Calibration matters more than accuracy when a clinician acts on the number.

Abstention — the model declining to answer when confidence falls below a threshold, rather than returning its best guess. Essential wherever rare entities are in play.

Batch effect — systematic differences arising from scanner, stain lot, laboratory or site rather than from biology. The leading cause of models performing worse elsewhere than where they were built.

Automation bias — the human tendency to defer to a system that is usually correct, and so to stop checking independently. A property of the workflow, not of the algorithm.

Predetermined Change Control Plan (PCCP) — a regulator-agreed specification of changes a manufacturer may make to an authorised device without a new submission. Its existence acknowledges that these systems are not static.

Shadow deployment — running an algorithm in parallel with routine practice, without acting on its output, to characterise its behaviour on your own material before it touches a live workflow.

Path-iQ — Global Pathology AI Review. Prepared for practising pathologists beginning AI adoption.