An overview of artificial intelligence applications for next-generation gynaecological pathology

Corresponding author:

Sarah Bell BSc (Med Sci), MBChB(Hons), MSc, FRCPath
Consultant Gynaecological Pathologist
Queen Elizabeth University Hospital, Govan Road, Glasgow, G51 4TF.
Pathologist, Industrial Centre for Artificial Intelligence Research in Digital Diagnostics.
Sarah.Bell@ggc.scot.nhs.uk
07732407547

Conflicts: None declared

James D Blackwood BSc(Hons), CITP
Healthcare AI Lead, Scottish Government. Chief Technology Officer and NHS National Programme Manager, Industrial Centre for Artificial Intelligence Research in Digital Diagnostics. AI Strategy and Portfolio Lead, NHS Greater Glasgow and Clyde
Conflicts of interest: Managing Director, Aingealag Limited. Chief Technology and Operating Officer, Simplifai Systems Limited.

Christina Fell PhD, MSc, MEng
Data Scientist/Post-doctoral Researcher
School of Medicine, University of St Andrews, UK
Conflicts of interest: none declared.

Mahnaz Mohammadi PhD, IISc
Data Scientist/Post-doctoral Researcher
School of Computing Science, University of St Andrews, UK
Conflicts of interest: none declared.

David Morrison BSc, MRes
Data Scientist/Post-doctoral Researcher
School of Computing Science, University of St Andrews, UK
Conflicts of interest: none declared.

David Harris-Birtill PhD, MPhys
Senior Lecturer
School of Computing Science, University of St Andrews, UK
Conflicts of interest: none declared.

Gareth Bryson BSc(Hons), MBChB, FRCPath
Consultant Pathologist and Clinical Director for Laboratory Medicine, NHS Greater Glasgow and Clyde
Honorary Clinical Associate Professor, University of Glasgow
Lead Pathologist Industrial Centre for Artificial Intelligence Research in Digital Diagnostics.
Conflicts of interest: GB is a member of the MSD UK Biomarker Advisory Board.

ABSTRACT

With the drive to roll out digital pathology in the UK, implementation of Artificial intelligence (AI) tools for pathology is now a possibility, bringing with it the potential

to change how we work as a specialty. AI promises many benefits for working practices such as improved efficiency and consistency, financial and productivity gains and ultimately a better service for our patients. Gynaecological pathology is a diverse specialty with many potential avenues for algorithm development, yet there are relatively few nearing clinical validation compared to other pathology specialties. This article provides a summary of the current landscape of AI in pathology with a focus on applications in gynaecological pathology. We discuss the ways pathologists can be involved in algorithm development and draw on our significant experiences in a nationally- funded programme for AI development and research. Finally we look to what the future might hold.

KEYWORDS
Artificial intelligence, deep learning, whole slide images, digital pathology, computational pathology, gynaecological pathology.

Introduction

There has been a recent drive to implement digital pathology in departments across the UK and as a result Artificial Intelligence (AI) integration is now a reality with some departments already using AI products or taking part in clinical validation of “close to market” products. There has been a significant increase in pathology AI models over the past few years, illustrated by the increasing number of publications in AI research and a small number of products achieving the necessary regulatory approvals to enter clinical workflows for validation. The potential benefits of AI in pathology are wide and include improving efficiency, patient safety and standardisation. Despite a wide range of AI tools being developed in both the research and commercial environments there is still a lack of translation into real world clinical use and as a result, pathology is perhaps slightly behind other healthcare settings such as radiology. For example, currently there are only 4 U.S Food & Drug Administration (FDA) approved pathology algorithms for clinical use in America whereas there are nearly 400 products in radiology[1].

Gynaecological pathology is a broad sub-specialty with a mixture of high volume, low complexity specimens such as endometrial and cervical biopsies, high complexity cancer work and cervical screening specimens, providing the perfect base for development of a wide range of pathology focused algorithms. Pathologist involvement in algorithm model development and translation into clinical practice is essential in order to ensure that any algorithm in clinical use is robust, clinically applicable and scalable.

There has already been a detailed review on the basics of AI and machine learning in a previous Diagnostic Histopathology article[2] and these concepts are therefore not being described in detail here. Instead, we outline our experiences with algorithm development including discussion of the infrastructure required for both research and clinical implementation and how a pathologist can get involved in these processes. We also provide some of our key learning points in developing a research platform

for AI integration in our department and discuss some of the barriers to AI integration going forwards. This article also aims to provide an overview of the range of AI algorithms in pathology focusing primarily on applications that could be applied to gynaecological pathology or that have already been developed in this field. Finally we discuss some of the challenges for AI implementation and what the future might hold for gynaecological pathology.

AI applications in gynaecological pathology

Clinical deployment of AI in pathology is in its infancy, with very few products at clinical validation stage. This is most likely due to the variable roll out of digital pathology across the UK, which underpins deployment of algorithms into clinical workflows. Gynaecological pathology as a subspecialty also has relatively few algorithms in the research development stages, though more are beginning to appear in the literature and some of these are described below. Part of this may be due to the largely commercial drive of AI development with many companies initially focusing on other higher volume subspecialities. **Figure 1** provides a summary of some of the current tools in pathology which could potentially be applied to gynaecological specimens.

Laboratory quality control (QC) and workflow

Digitisation of glass slides requires specific laboratory slide preparation steps in order to ensure tissue is appropriately mounted for tissue detection and subsequent high resolution scanning. Slides with quality issues such as air bubbles, tissue folds, pen marks, lens contamination or image blurring (out of focus) result in delays as the slides may need rescanned and/or recut. One example of an AI solution is HistoQC[3], which is available open source but there are many other commercial products available, often as an integrated part of AI platforms. These products can automate the laborious process of manual slide QC following scanning and can be implemented during the scanning process to detect the above QC issues. Slides can be flagged for re-scanning before a case reaches the pathologist's worklist. QC products can also automate colour balancing between different H&E stains and pre-order ancillary tests on diagnostic slides. Potential applications in gyn path could be fungal stains in inflammatory vulval biopsies or immunohistochemistry for p53 and MMR in endometrial biopsies in cases with malignancy, thus reducing turnaround time for case reporting.

Workflow prioritisation

Workflow prioritisation is another area where AI could bring significant benefits to pathology departments in terms of triaging cases and distributing work to pathologists. Tools that can automatically triage cases according to clinical priority and ensure they are sent to the most appropriate pathologist's digital worklist expediently, are of potential benefit in terms of turnaround time and reducing harm. One example already being rolled out for validation is a dermatopathology clinical

prioritisation tool[4]. This type of 'classification' algorithm can review whole slide images (WSI) of skin lesions, classify these into subtypes such as melanoma, basal cell carcinoma and squamous cell carcinoma with high sensitivity. One potential application of this type of tool in gynaecological pathology would be in Cervical Screening Programme specimens where identifying high risk lesions such as invasive cancers or high grade dysplastic lesions such as CIN and CGIN could allow for prioritisation of reporting.

Large scale pre-screening tools are emerging for high volume, low complexity samples where the ability to automatically 'screen out' normal samples from clinical work flow would result in a significant reduction in the volume of cases for reporting. This concept formed the basis of some of our project work in endometrial biopsy reporting as over 97% of our biopsy cases are benign (local audit data). We are currently what feels like a very long way from automated sign out of reports by AI but an initial read of a case could build confidence in algorithm performance and potentially reduce both the time to report and the number of second opinions required. An example of a large scale pre-screening algorithm in development is the COBIX work by Prof Snead and team in Warwick[5] supported by PathLake. They have developed a weakly supervised pre-screening algorithm that can distinguish between normal and neoplastic colon biopsies with high validation accuracy and sensitivity. High sensitivity is important in screening algorithms in order to ensure that malignant cases are not missed. This initial pilot work has since resulted in an NIHR funded multicenter study across the UK.

Morphological assessments

An advertised benefit of AI tools in cancer assessment is improving consistency of pathological features which have notorious inter and intraobserver variability but are important for prognosis such as grading and histotyping. Current available products include invasive cancer detection algorithms, mitotic counts and quantification tools to calculate Ki67 index, some of which are already integrated into pathology management system platforms as ready to use software. The latter is described by Flach et al[6] who are evaluating a number of algorithms in their department in Utrecht, including integrated Ki67 calculation and a mitosis counter. These could have some, albeit limited, applications in tumours of gynaecological tract origin, such as smooth muscle or neuroendocrine tumours. As these are relatively uncommon, overall efficiency or financial gains from introducing these types of products are likely to be marginal.

Gynaecological specific algorithms are very much in the development phase at present though more are appearing in the research landscape. One area of particular interest in terms of 'grading' is cervical intraepithelial neoplasia. Poor pathologist interobserver variability in grading CIN is well known[7] and improvement of the consistency of grading would in turn allow for more accurate treatment and follow up. Cho et al[8] trained a supervised deep learning model to classify 1106 images as non neoplastic or CIN 1 to 3 with an accuracy of 89% which was

comparable to pathologist performance of up to 93%. Oliveria et al[9] also recently described a weakly supervised model trained and validated on 600 samples which can grade cervical LLETZ specimens into four categories (low grade squamous intraepithelial lesion (LSIL), high grade squamous intraepithelial lesion (HSIL), non neoplastic and non representative) with a reasonable accuracy of 63.75% and sensitivity of 68.67%. With many of these studies, much larger prospective datasets with linked follow up data are required. Weakly supervised learning is the preferred avenue to gather large amounts of validation data without the need for detailed, time consuming annotations. Methods of deep learning and assigning ground truth are discussed in the section below on algorithm development.

Tumour detection algorithms can identify invasive cancer, provide tumour quantification metrics, histological subtype and grading or presence of lymph node metastases purely from analysis of the WSI alone. These types of algorithms are particularly well developed in breast cancer and prostate cancer with some products in clinical validation stages. Detection of nodal metastases is one potential application in gynaecological pathology as a significant proportion of patients with endometrial, vulval and cervical cancers undergo pelvic or groin lymphadenectomy at primary surgery. Furthermore sentinel node assessment is costly and time consuming from a pathology perspective in terms of slide review, additional step H&E sections and IHC requirements. Breast nodal metastasis detection models have been shown to be comparable and in some cases more accurate than pathologist assessment alone[10]. There are no comparable algorithms for gynaecological cancers.

In terms of histotyping, some models can predict histological subtype in ovarian and endometrial cancer. Both of these tumour types are known to have problems in interobserver reliability even when ancillary tests are used. An example is a recent study by Farahani et al[11] in which they trained a model using over 900 H&E WSIs. This model was able to classify ovarian carcinomas into the five commonest subtypes (high grade serous, low grade serous, endometrioid, mucinous and clear cell carcinoma) purely on digital features with an overall concordance of 80.97% (Cohen's kappa 0.7547) in their external validation set. Interestingly they also assessed performance using multiple slides, which would be normal clinical practice, versus a single slide, the former of which showed a higher concordance rate (86.56% versus 81.38%) in their internal dataset. Furthermore the challenges with classification in these types of algorithm are often similar to the diagnostic challenges pathologists face in typing tumours. It is unlikely that any classification algorithm will function at 100% accuracy but could be employed as a diagnostic 'aid' or digital second opinion to improve consistency.

Immunohistochemistry quantification

As discussed above, a useful application of AI lies in areas of poor reproducibility or variation such as immunohistochemical scoring. A good example of this is Programmed Death Ligand 1 (PDL1) immunohistochemistry, an important test to

determine patient eligibility for checkpoint inhibitor therapies such as nivolumab in a range of tumour types including advanced cervical cancer, lung cancer and upper GI cancer[12]. There are different scoring methods (Tumour Proportion Score: TPS and Combined Positive Score: CPS), PDL1 assays (22C3, 28.8 etc) and cut offs which determine treatment indication. Furthermore interpretation of staining can be challenging and time consuming even in the most experienced hands, particularly with cases close to the cut off values. Therefore inter and intraobserver correlation in scoring is poor, as clearly illustrated in a very recent paper by Robert et al showing poor to fair correlation coefficients for pathologists scoring upper GI PDL1[13].

There are a number of companies with PDL1 AI products in validation stages or close to clinical deployment. For example, Baxi et al[14]developed a PDL1 evaluation model based on clinical trial datasets, which detected more positive cases than manual scoring and showed similar or more significant impact on trial outcome data. This model is now being developed commercially and assessed in a variety of real world settings for validation, with comparable performance to a consensus of experienced reporting pathologists regardless of clone. AI driven scoring for companion diagnostics for specific drugs has the potential to better identify which patients should receive treatments and perhaps more importantly who should not, which could have significant financial implications for the NHS.

Another area of potential application in gyn path is interpretation of ER expression in endometrial cancer, which is prognostic in some specific subtypes[15]. AI supported interpretation could allow for more robust scoring for correlation with survival data. Mismatch Repair (MMR) immunohistochemistry is also the NICE[16] recommended method for assessment of mismatch repair deficiency in endometrial carcinoma and IHC evaluation could potentially be performed by AI. However there are models which can predict molecular alterations based on digital imaging features alone that may supersede this, thus negating the need for IHC or molecular testing.

WSI as a predictive biomarker

A WSI can now be utilized as a digital biomarker for particular tumours and pre-neoplastic conditions, where specific features in the digital image correlate with survival and treatment response.

Using AI to predict prognosis and risk of progression in pre-neoplastic conditions is a potential novel area of development that would provide additional information over and above a standard pathology report. For example, hyperplasia with and without atypia is a challenging area of gynae pathology with poor reproducibility. Rewcastle et al[17] recently developed an AI tool (ENDOAPP) to assess features within endometrial biopsies with a diagnosis of hyperplasia that correlated with prognosis and were able to assign a low risk versus high risk of progression score with a reasonable level of performance (accuracy of 88-91%), comparable to other classification schemes of hyperplasia such as EIN and WHO 2020.

There are many studies assessing radiological imaging based chemotherapy response in high grade serous carcinoma (HGS) of tubo-ovarian origin, as around 15-20% of patients will be resistant to platinum based therapies. AI models are now being developed in pathology specimens that can predict chemo-responders purely from assessment of morphological features on WSI. This is again a potential novel area for enhancing pathology reporting, as prediction of outcome purely on pathologist assessment of HGS morphology is not reproducible. There is some evidence that HGS associated with BRCA mutation shows distinctive morphological patterns such as solid, pseudoendometrioid and transitional ('SET') morphologies. However these are not reproducible enough to introduce into routine reporting[18]. Despite the introduction of routine BRCA mutation and Homologous Recombination Deficiency (HRD) testing to identify patients who might respond to poly ADP ribose polymerase inhibitor (PARPi) therapy, prediction of response to treatment remains challenging. Laury et al[19] describe development of an AI morphological assessment model to predict platinum based chemotherapy response using weakly supervised learning methods with high sensitivity (72%) and overall accuracy of 82%. Prediction of outcome in a trial setting with robust follow up data and genomic results may provide the optimal environment for moving these studies towards clinical deployment.

Current risk prediction in endometrial cancer is based on traditional pathological features such as tumour type, grade, stage and LVSI and further refined by molecular subtyping into p53 mutant, Mismatch repair (MMR) deficient, POLE ultramutant and No Specific Molecular Profile (NSMP) groups[20]. There are some pathological features associated with MMR deficiency such as endometrioid subtype, tumour infiltrating lymphocytes and mucinous differentiation but these are not reliably reproducible at a microscopic level. However analysis of WSI has the potential to unlock additional morphological information which is not necessarily apparent to the naked eye of a pathologist. There have only been a couple of studies to date using AI to predict molecular features in endometrial cancer, the most recent of which was developed on over 2000 cases including PORTEC trial datasets, with their associated clinical and molecular data. In this study[21] it was possible to predict the four endometrial molecular subtypes without the need for specific molecular testing. Interestingly, there was some overlap in distinguishing MMRd and POLEmut groups, where there is known to be overlap in histological features such as immune infiltrates and solid growth patterns. Their image based molecular classification also showed similar stratification for outcome compared to the true molecular data in the PORTEC 3 trial cohort, illustrating that prediction of the four classes is possible using this algorithm on HE scanned images alone. The real world benefit for these applications could be results available in minutes, as an adjunct to pathological reporting. However at present molecular testing is still the gold standard and any AI product would require molecular confirmation.

One important aspect of any implementation of AI predictive models is the interplay with other branches of diagnostics providing similar information. In particular, genomics is a transforming field, with solid tumour testing moving to large panel sequencing or even whole genome sequencing. Therefore the real challenge for AI is moving from predictions on single markers to providing novel predictive

information quickly that can complement or supersede other tumour testing modalities.

Education and training
The benefits of digital pathology for education and training were discussed more extensively in a previous article and are not covered in detail here. However there is potential for AI to complement this, transforming how we learn, contributing to continued professional development and also training our junior colleagues in the digital age. Examples could include algorithms for automated annotations, automated tagging and logging of training cases and digitally identifying any gaps in knowledge for trainees. Automated Quality Assurance may also be a possibility with reported cases automatically returned to the workflow without the pathologist knowing until the end of the report – thus providing a real life simulation of reporting and feedback.

Pathologist involvement in algorithm development

Pathologist involvement in AI is essential in order to ensure robust, clinically applicable algorithms for future use. Pathologists can assist with research projects or via the commercial sector, either at the initial model development stage or later in clinical validation. Some considerations are discussed below.

1) *Algorithm model design* – pathologists are already reporting cases and have excellent knowledge of the clinical workflow and potential implications of a product. Therefore initial concepts of areas where AI could improve practice, reduce pathologist workload or completely replace existing mechanisms, are highly valuable.
2) *Education and training of development team* – Pathologists can assist in model development by providing training on clinical workflow, implications of results and basic pathology interpretation. This can improve data scientists' understanding of the problem and aid in finding potential solutions when there are errors or discordances in model development.
3) *Cohort curation* - Curation of balanced datasets with appropriate metadata and anonymised linked clinical data underpin the success of any new model. Pathologists can be involved in identifying appropriate cases either retro or prospectively. They may also take part in expert pathology review as part of a consensus group in order to determine the 'ground truth' or the most likely correct diagnosis which the algorithm outputs will be compared to. Sets are usually split evenly into cases for training the algorithm, testing the algorithm and final validation.
4) *Weakly supervised and fully supervised learning* – In fully supervised learning the algorithm is given the slide level output (overall diagnosis) and detailed annotations (drawings around areas of interest on the slide with labels). This process is generally used for classification type models. The annotation

process can require a significant amount of pathologist time depending on the detail required and annotation drawings can be variable between pathologists. In weakly supervised learning, only the slide level and the output is provided with no annotations but this still requires establishing the ground truth of the slide first (explained above).

5) *Clinical validation* – Following initial development and testing, algorithms should be validated on an external test set to ensure there is no bias towards the initial data the algorithm was trained on and that it is clinically usable in other settings. Pathologists could be involved in using the tools as part of normal clinical work flow in order to establish real world experience, safety and interoperability with other software products and platforms.

Creation of a Scottish 'living lab' for AI development

The work described in the next few sections is part of iCAIRD which is funded by Innovate UK on behalf of UK Research and Innovation (UKRI) - project number: 104690

The Industrial Centre for Artificial Intelligence Research in Digital Diagnostics (iCAIRD) is a joint national pathology and radiology centre of excellence based in Scotland. Working in partnership with iCAIRD, industry, eHealth and biorespository teams at NHSGGC, we've developed a new digital pathology 'living laboratory' (**Figure 2**). This includes a research image management system that is isolated from our live clinical systems, a dedicated scanning facility, data extraction and de-identification tools, and an AI integration and orchestration platform.

We can import de-identified data from our live pathology system or use research-only images that have been generated using our research scanning facility which can include linked data from the electronic patient record such as demographic information, diagnosis codes, medicines and medical history.

AI is thus integrated in a realistic, near-clinical setting, for validation and evaluation. This allows us to go beyond mere technical performance evaluation and additionally examine workflow, usability, performance and potential scalability, as well as the value of AI which is essential for any future clinical roll out.

This experience has provided us with several general learning points for successful AI research and evaluation in the NHS:

1) Data storage: appropriate and secure storage to meet NHS code of practice. We split our data into live and archive, on fast and slower storage infrastructure. Fast storage is expensive, and it wouldn't have been sustainable for us to simply expand what we had – the costs running to several million pounds. Whilst not strictly necessary, as our archive storage is on-premises, we felt the time was right to prepare for a future where we'd archive our data in the cloud, so splitting live and archive data made sense.

2) Isolated research environment: Having a one-way interface from the clinical to the research environment protects the integrity of the clinical system whilst allowing use of real-world data in the research environment.
3) Dedicated research scanning: This minimizes the impact on operational pathology infrastructure with capacity for clinical utilization in times of heavy clinical demand.
4) De-identification: Ensuring the privacy of clinical information in research is paramount when sharing digitized data. The de-identification tool has allowed us to securely extract and anonymise data at scale for use by academic and industrial research partners, whilst maintaining patient privacy.
5) Clinical simulation: Having a flexible method of integrating projects and products into the research environment, and returning results matched to the originating case, is critical for mimicking realistic clinical workflow.
6) Partnership: Our living laboratory is a collaborative effort. It is only possible because pathologists and the wider laboratory medicine team work closely with colleagues in IT, research, innovation and information governance, in addition to partners in academia and industry.

Development of endometrial and cervical cancer detection algorithms

One of the project areas supported by iCAIRD (Scotland) is a proof of concept study in developing endometrial and cervical biopsy triage algorithms. These could potentially help with workflow by identifying urgent cases from backlogs. A summary of the workstrands are given below:

Fully supervised learning in endometrial biopsies[22]

We developed an AI algorithm trained and evaluated on nearly 300 endometrial biopsies which can automatically sort WSI into one of three categories, "malignant", "other or benign" or "insufficient", allowing prioritisation of malignant slides within the pathology workload and therefore reduce the time to diagnosis for patients with cancer. An explanation of the pipeline used for the fully supervised work is given in **Figure 3.** The final model shows reasonable performance metrics, accurately classifying 90% of all slides correctly and 97% of the malignant slides correctly. The output also includes segmentations of areas predicted as malignant in the form of heatmaps which could be used to guide pathologist's attention. **Figure 4** shows an example case with our algorithm result.

Fully supervised learning in cervical cancer

Similarly, we developed an AI algorithm, trained and evaluated on nearly 3000 cervical biopsies which can automatically classify WSI into "normal/inflammatory",

“low grade CIN”, “high grade CIN” and “malignant”. Performance was promising with 93.40% malignant sensitivity in the final test set (paper submitted for publication). **Figure 5** shows an example of the algorithm in action.

Weakly supervised learning in endometrial cancer[23]

Fully supervised learning for WSI model development is problematic due to the requirement for costly and time-consuming manual pathologist annotation. Weakly supervised learning which utilises only slide-level labels during training is becoming more widespread for this reason. We developed such an algorithm to assess endometrial biopsies utilising only the WSI and pathologist slide level label (i.e. overall diagnosis). Performance was promising, showing an accuracy of 97.04% in the best model.

Challenges, opportunities and future directions for AI in diagnostic pathology

At present there is a small market for AI in pathology with developments driven in part by academia but largely by industry. Therefore products which are currently closest to clinical deployment are those with the most benefit in terms of scalability and financial gain such as in prostate cancer, lung cancer and breast cancer. At present, the UK is only just starting clinical validation studies in partnership with commercial vendors and some of the digital pathology and IT infrastructure required to support this will take time to be deployed. Furthermore, there is a huge amount of work required to develop the governance systems around AI deployment and monitoring, both at a local and national level. There are many challenges to how we implement AI clinically and incorporate it into reports as, much like a trained pathologist, no tool can provide 100% accuracy and are an aide to the diagnostic process rather than a definitive solution.

A recent position statement on AI from the Royal College of Pathologists[24] acknowledges the important role of pathologists in developing, evaluating and implementing AI tools. Guidelines are urgently needed on deployment, governance and standards in order to ensure safety and robustness of integration across multiple health boards and networks. They also acknowledge the importance of patient engagement in this process.

There are also many ethical challenges to consider which were described in detail in a review article by McKay et al[25] and include factors such as data protection and safeguarding, patient consent and environmental impact. Ultimately we must ensure that we support the development of safe and robust systems for clinical deployment that meet our obligations in providing high quality safe care for patients.

Despite these challenges, we feel that the breadth and complexity of gynaecological pathology brings with it a great opportunity to be involved in the direction of algorithm

development at an early stage. We have the chance to upskill as a profession and become comfortable with digital workflow including incorporation of AI tools. The future for AI development for gynaecological pathology will involve stepping beyond the current trend of replacing existing pathology tasks and identifying novel ways to provide additional prognostic and predictive information, resulting in significant benefits for patients. Furthermore, as more AI tools appear in other healthcare sectors, patients may also expect that the same is afforded to their pathology specimens.

Conclusions

Artificial intelligence has the potential to revolutionise how we work in pathology for the benefit of our patients and our own working practices. In this article we have provided an overview of the current landscape of algorithms in development and validation phases and how these are being applied in gyn pathology. We have discussed how pathologists can play an integral part in algorithm design and evaluation and have also discussed some of our experiences in rolling out the infrastructure required for AI integration in clinical services. The challenges going forwards initially are ensuring digital pathology rolls out equitably across the UK, engaging our supporting services such as IT, persuading management of financial gains as a result of roll out and developing standardized approaches to deployment. Gyn pathology AI development is in its infancy and as a result there are plenty of opportunities for engagement in development and validation. The future is digital and as a subspecialty we should embrace the changes to come with open arms.

Acknowledgements
The work described in the section Creation of a Scottish 'living lab' for AI development is part of iCAIRD which is funded by Innovate UK on behalf of UK Research and Innovation (UKRI) - project number: 104690.

References
1. AI/ML enabled Medical Devices (Internet). (Updated 05/10/22. Cited 20th March 2023). Available from: https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices.
2. Moxley- Wyles B, Colling R, Verrill C. Artificial Intelligence in Pathology – an overview. Diagnostic Histopathology 2022; 26 (11): 513-520.
3. Janowczyk A, Zuo R, Gilmore H et al. HistoQC: An open source Quality Control Tool for Digital Pathology Slides. JCO CLin Cancer Inform. 2019;3: 1-7. PMID: 30990737.
4. Sankarapandian S, Kohn S, Spurrier V et al. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops, 2021, pp. 629-638.
5. Bilal M, Tsang YM, Ali M et al AI based pre-screening of large bowel cancer via weakly supervised learning of colorectal biopsy histology images medRxiv 2022.02.28.22271565; doi: https://doi.org/10.1101/2022.02.28.22271565.

6. Flach R, Fransen N, Sonnen A et al. Implementation of Artificial Intelligence in Diagnostic Practice as a Next Step after Going Digital: The UMC Utrecht Perspective. Diagnostics 2022, 12, 1042. https://doi.org/ 10.3390/diagnostics12051042.
7. Stoler M & Schiffman M. Interobserver reproducibility of cervical cytologic and histologic interpretations: Realistic estimates from the ASCUS-LSIL triage study. *JAMA* **285**(11), 1500–1505. https:// doi. org/ 10. 1001/ jama. 285. 11. 1500 (2001).
8. Cho B, Kim J, Park J. et al. Automated Diagnosis of Cervical Intraepithelial Neoplasia in Histology Images via Deep Learning. *Diagnostics* **2022**, *12*, 548. https://doi.org/10.3390/diagnostics12020548.
9. Oliviera S, Montezuma D, Moreira A et al. A CAD system for automatic dysplasia grading on H&E whole slide images. Scientific Reports 2023 (13):3970. | https://doi.org/10.1038/s41598-023-30497-z.
10. Ehteshami Bejnordi B, Veta M, Johannes van Diest P et al. Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer. *JAMA.* 2017;318(22):2199–2210. doi:10.1001/jama.2017.14585.
11. Farahani H, Boschman J, Farnell D *et al.* Deep learning-based histotype diagnosis of ovarian carcinoma whole-slide pathology images. *Mod Pathol* **35**, 1983–1990 (2022). https://doi.org/10.1038/s41379-022-01146-z.
12. Janjigian YY, Shitara K, Moehler M, et al. First-line nivolumab plus chemotherapy versus chemotherapy alone for advanced gastric, gastro-oesophageal junction, and oesophageal adenocarcinoma (CheckMate 649): a randomised, open-label phase 3 trial. Lancet. 2021, 398:27-40.
13. Robert M, Rüschoff J, Jasani B et al. High Interobserver Variability among Pathologists Using Combined Positive Score to Evaluate PD-L1 Expression in Gastric, Gastroesophageal Junction and Esophageal Adenocarcinoma, Modern Pathology (2023), doi: https://doi.org/10.1016/ j.modpat.2023.100154.
14. Baxi V, Lee G, Duan C et al. Association of artificial intelligence-powered and manualquantification of programmed death-ligand 1 (PD-L1)expression with outcomes in patients treated withnivolumab ± ipilimumab. Modern Pathology 2022 35:1529-1539. https://doi.org/10.1038/s41379-022-01119-2.
15. Jamieson A, Huvila J, Chiu D et al. Grade and Estrogen Receptor Expression Identify a Subset of No Specific Molecular Profile Endometrial Carcinomas at a Very Low Risk of Disease-Specific Death. Modern Pathology 2023 25;36(4): 100085. doi: 10.1016/j.modpat.2022.100085. Online ahead of print.
16. Testing strategies for Lynch Syndrome in people with Endometrial Cancer. NICE Diagnostics Guidance [DG42]. Published 28 October 2020.
17. Rewcastle E, Gudlaugsson E, Lillesand M et al. Automated Prognostic Assessment of Endometrial Hyperplasia for Progression Risk Evaluation Using Artificial Intelligence. Modern Pathology 2023 36: 100116. https://doi.org/10.1016/j.modpat.2023.100116.
18. Soslow R, Han G, Park K et al. Morphologic patterns associated with BRCA1 and BRCA2 genotype in ovarian carcinoma. Modern Pathology 2012 25(4):625-636.

19. Laury A, Blom S, Ropponen T et al. Artificial Intelligence based image analysis can predict outcome in high grade serous carcinoma via histology alone. Scientific Reports. 2021 11:19165 https://doi.org/10.1038/s41598-021-98480-0.
20. Concin N, Matias-Guiu X, Vergote I et al. ESGO/ESTRO/ESP guidelines for the management of patients with endometrial carcinoma. Int J Gynecol Cancer. 2021 Jan;31(1):12-39. doi: 10.1136/ijgc-2020-002230. Epub 2020 Dec 18. PMID: 33397713.
21. Fremond S, Andani D, Wolf J et al Interpretable deep learning model to predict the molecular classification of endometrial cancer from haematoxylin and eosin stained whole slide images: a combined analysis of the PORTEC randomized trials and clinical cohorts. Lancet Digital Health 2023:5:e71-82.
22. Fell C, Mohammadi M, Morrison D et al . Detection of malignancy in whole slide images of endometrial cancer biopsies using artificial intelligence. PLoS One. 2023 Mar 8;18(3):e0282577. doi: 10.1371/journal.pone.0282577. PMID: 36888621; PMCID: PMC9994759.
23. Mohammadi M, Cooper J, Arandelović O, et al. Weakly supervised learning and interpretability for endometrial whole slide image diagnosis. Experimental Biology and Medicine. 2022;247(22):2025-2037. doi:10.1177/15353702221126560.
24. RCPath. Royal College of Pathologists position Statement on Digital Pathology and Artificial Intelligence (Internet). Feb 2023. Cited 20th March 2023. Available at: https://www.rcpath.org/profession/digital-pathology.html.
25. McKay F, Williams B, Prestwich G et al. The ethical challenges of artificial intelligence driven digital pathology. The Journal of Clinical Pathology: Clinical Research. 2022 8:209-216.

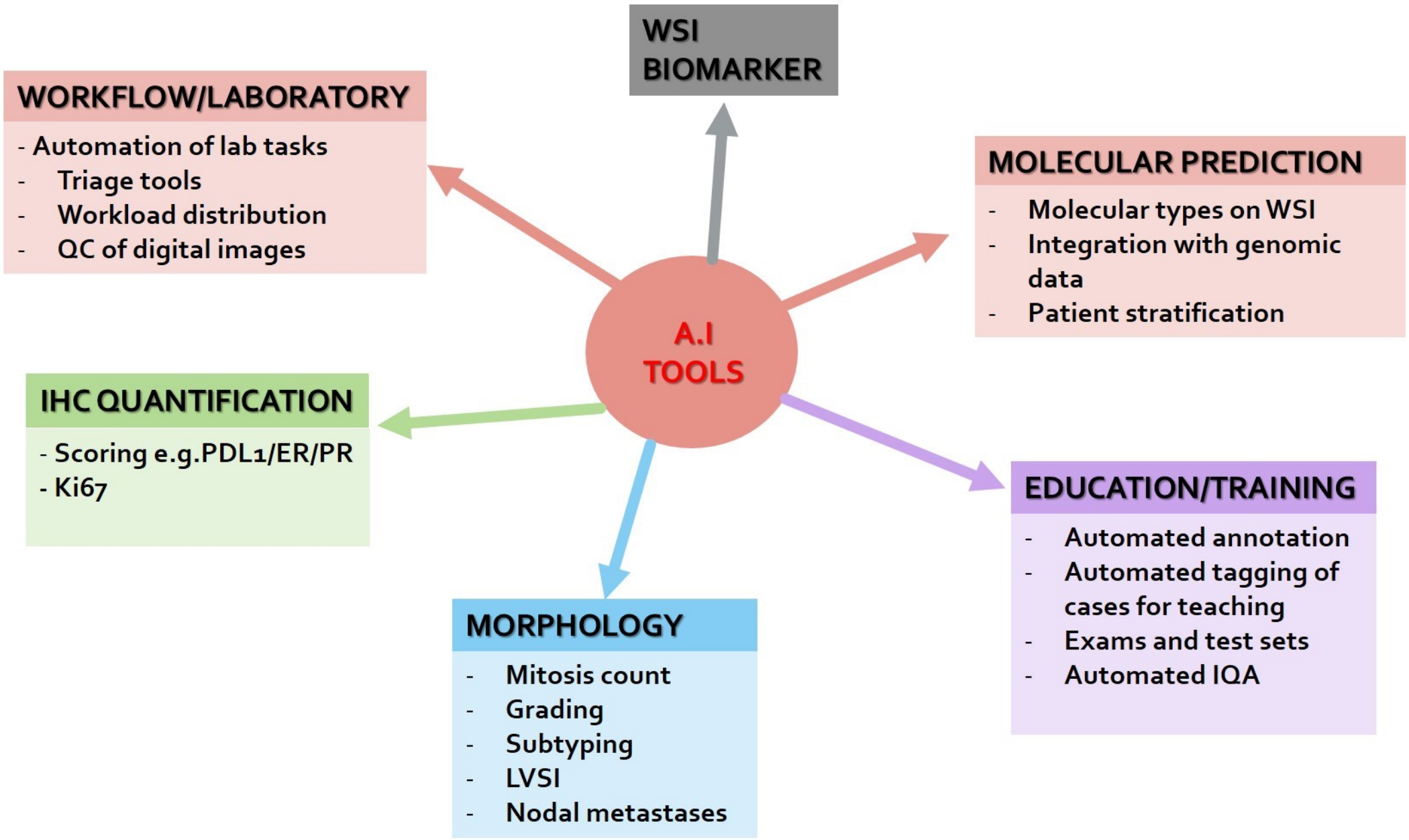
Figure 1: AI applications in pathology. This figure shows the range of current applications in pathology, many of which can also be applied to gynaecological pathology.

Figure 2: NHS Greater Glasgow and Clyde 'living laboratory': This shows our pathology digital architecture for integrating AI research and developments into our clinical workflow.

Figure 3: Endometrial processing pipeline[22]: The WSI is pre-processed so that only areas that contain tissue are used. WSIs are typically orders of magnitude larger than the input sizes of deep learning neural networks, therefore the WSI is split into many smaller patches. A CNN is trained to predict the probability of a patch being "malignant" or "other or benign". The predictions for all the patches from a slide are then reassembled into a heatmap. A slide model is then trained to give an overall prediction for the slide. The slide model categorises the slide as either malignant, benign or insufficient.

Figure 4 AI functionality – benign endometrium: This shows benign secretory endometrium on WSI (a). The algorithm heatmaps show strong benign prediction (b) with the corresponding tissue areas highlighted in white and no malignant predictions (c). The correct overall slide prediction is shown in (d) with benign tissue highlighted in red.

Figure 5: AI functionality - high grade CIN. Below shows a cervical biopsy WSI showing high grade CIN (a). The heatmaps show strong prediction for high grade CIN highlighted in the white areas (b) and normal/inflammatory tissue highlighted in white in image (c). The final overlay (d) clearly shows prediction of high grade CIN areas highlighted in blue.


WSI BIOMARKER
WORKFLOW/LABORATORY
- Automation of lab tasks
- Triage tools
- Workload distribution
- QC of digital images
MOLECULAR PREDICTION
- Molecular types on WSI
- Integration with genomic data
- Patient stratification
A.I TOOLS
IHC QUANTIFICATION
- Scoring e.g.PDL1/ER/PR
- Ki67
EDUCATION/TRAINING
- Automated annotation
- Automated tagging of cases for teaching
- Exams and test sets
- Automated IQA
MORPHOLOGY
- Mitosis count
- Grading
- Subtyping
- LVSI
- Nodal metastases

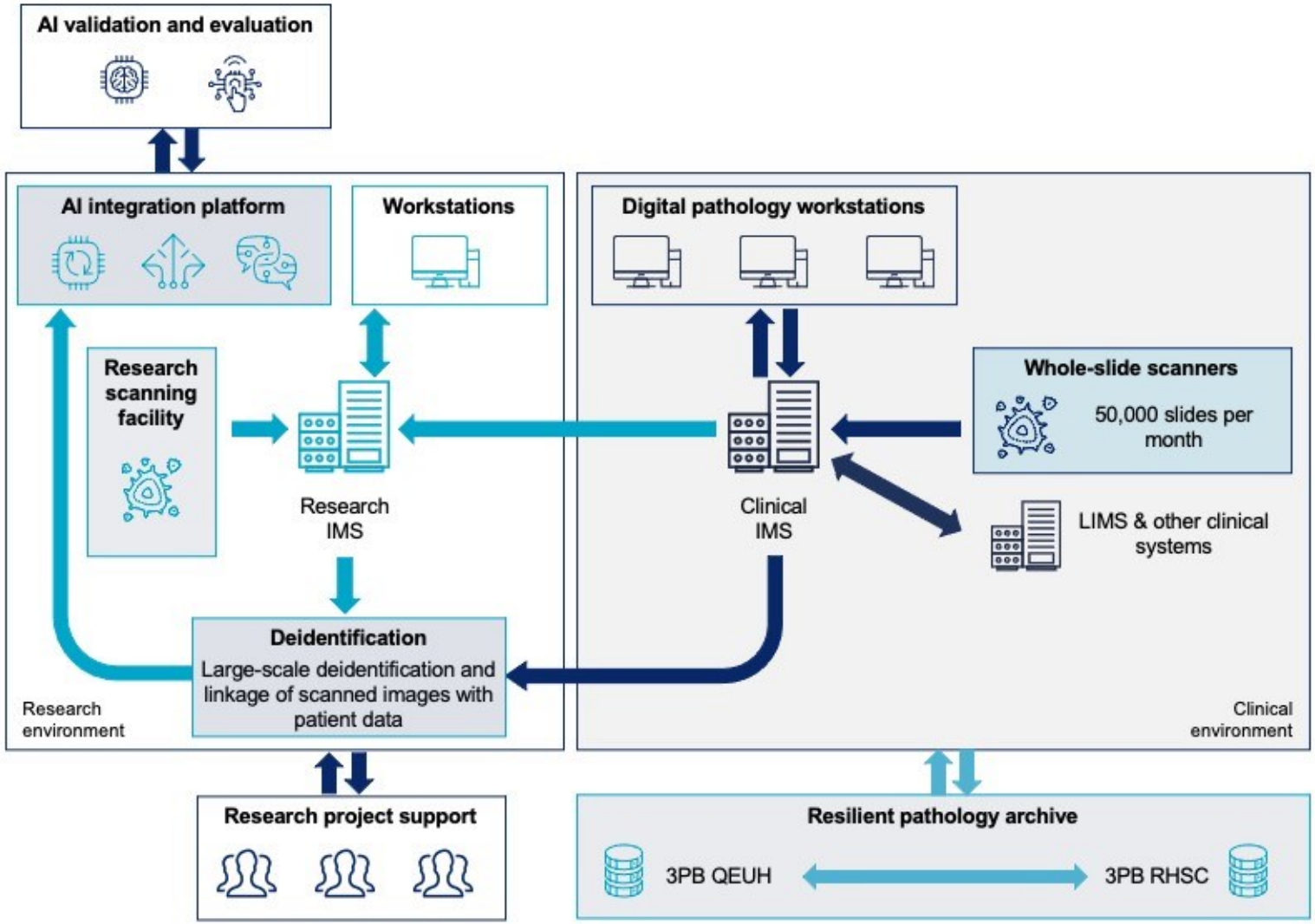

AI validation and evaluation
AI integration platform
Workstations
Digital pathology workstations
Research scanning facility
Research IMS
Clinical IMS
Whole-slide scanners
50,000 slides per month
LIMS & other clinical systems
Deidentification
Large-scale deidentification and linkage of scanned images with patient data
Research environment
Clinical environment
Research project support
Resilient pathology archive
3PB QEUH
3PB RHSC

Figure 2

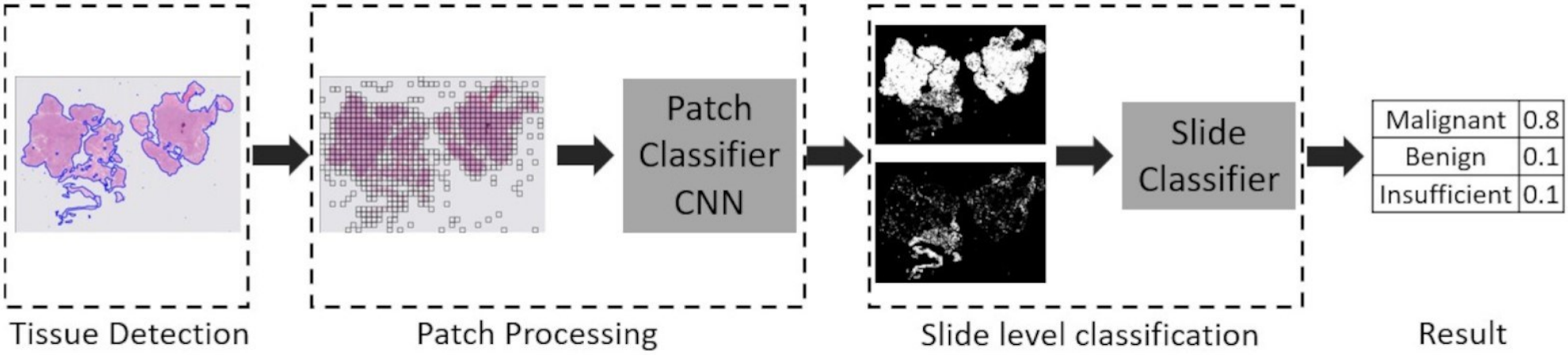

Patch
Classifier
CNN
Slide
Classifier
Malignant 0.8
Benign 0.1
Insufficient 0.1
Tissue Detection
Patch Processing
Slide level classification
Result

Figure 3

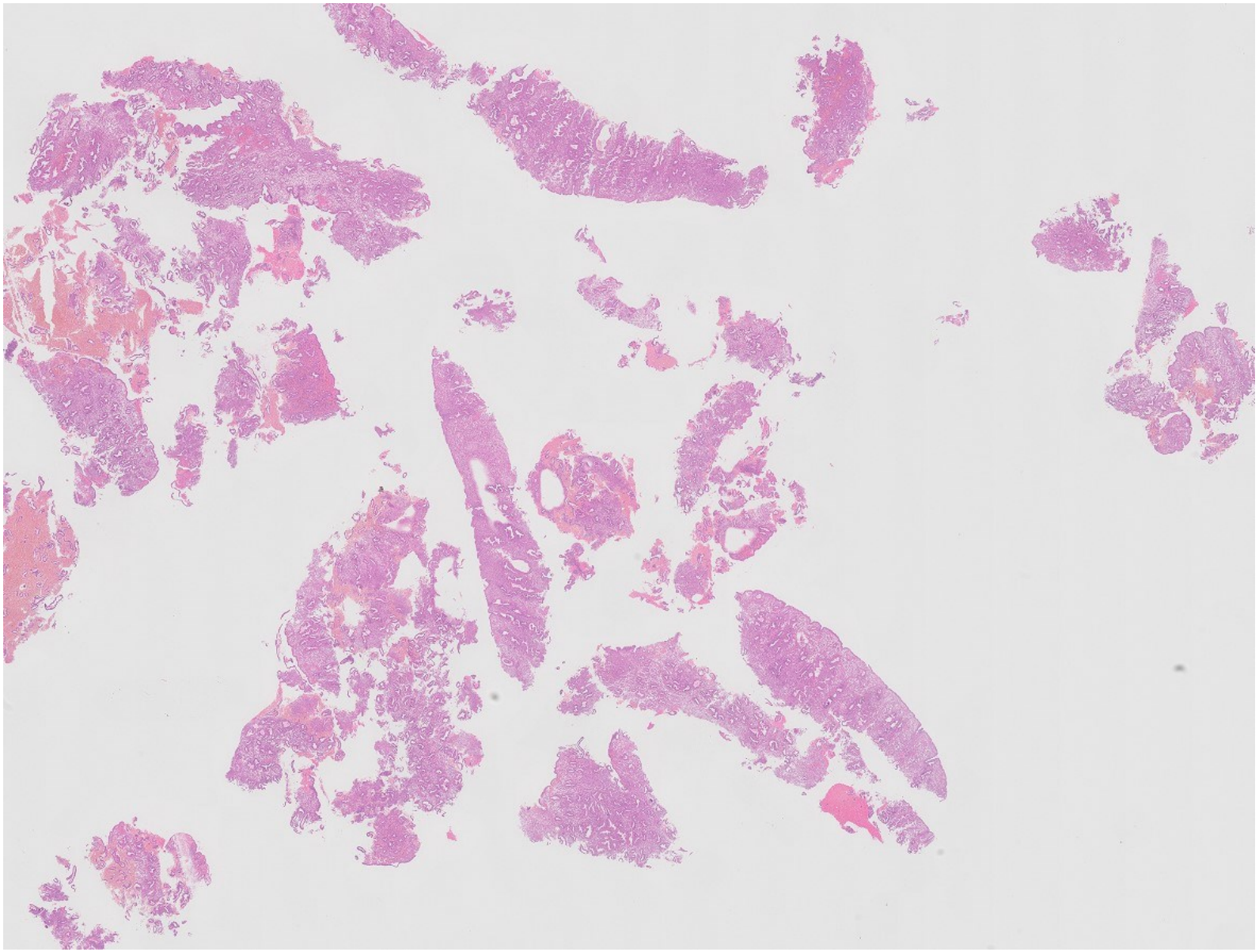


Figure 4a.

Figure 4b.

Figure 4c.

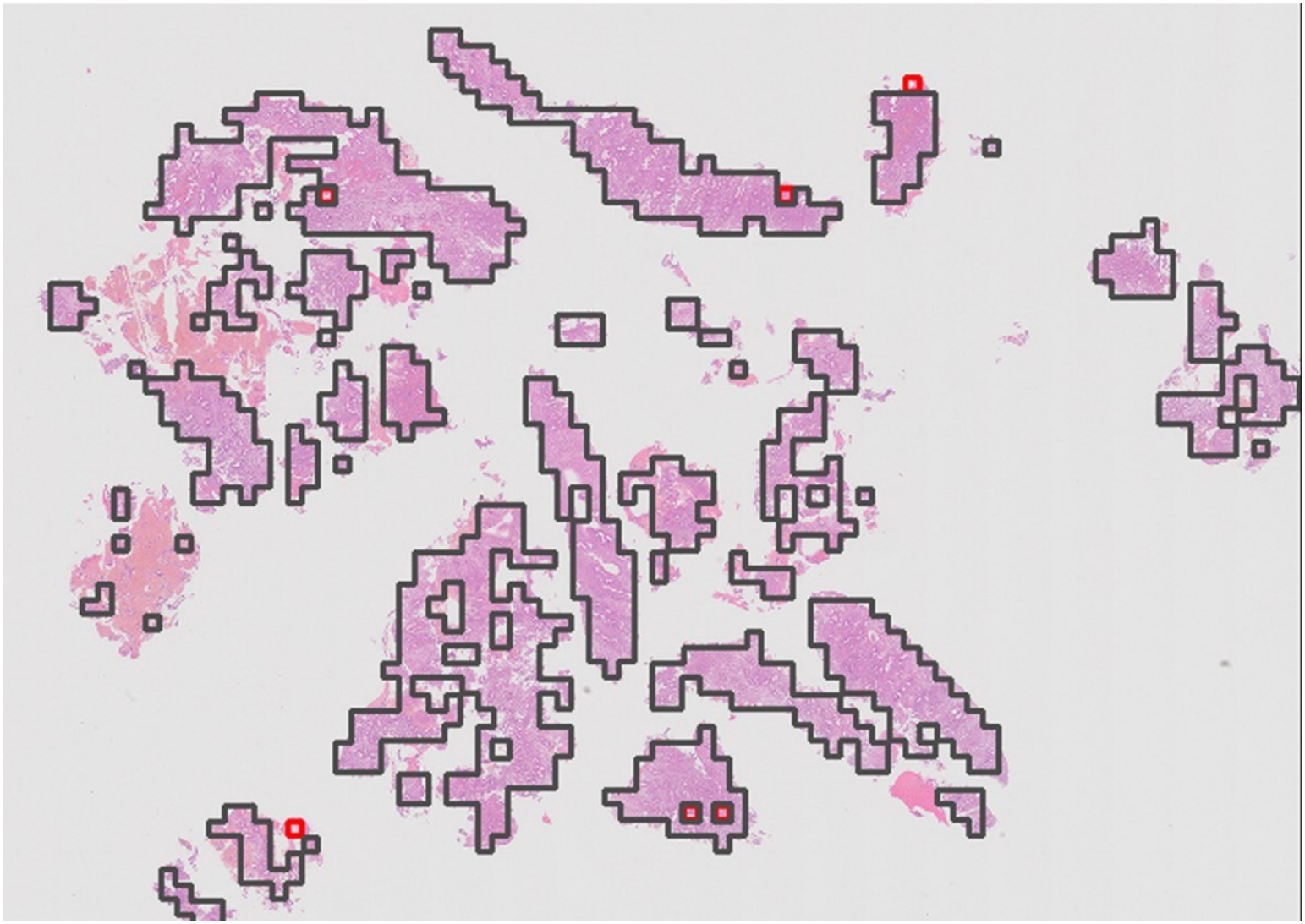


Figure 4d.

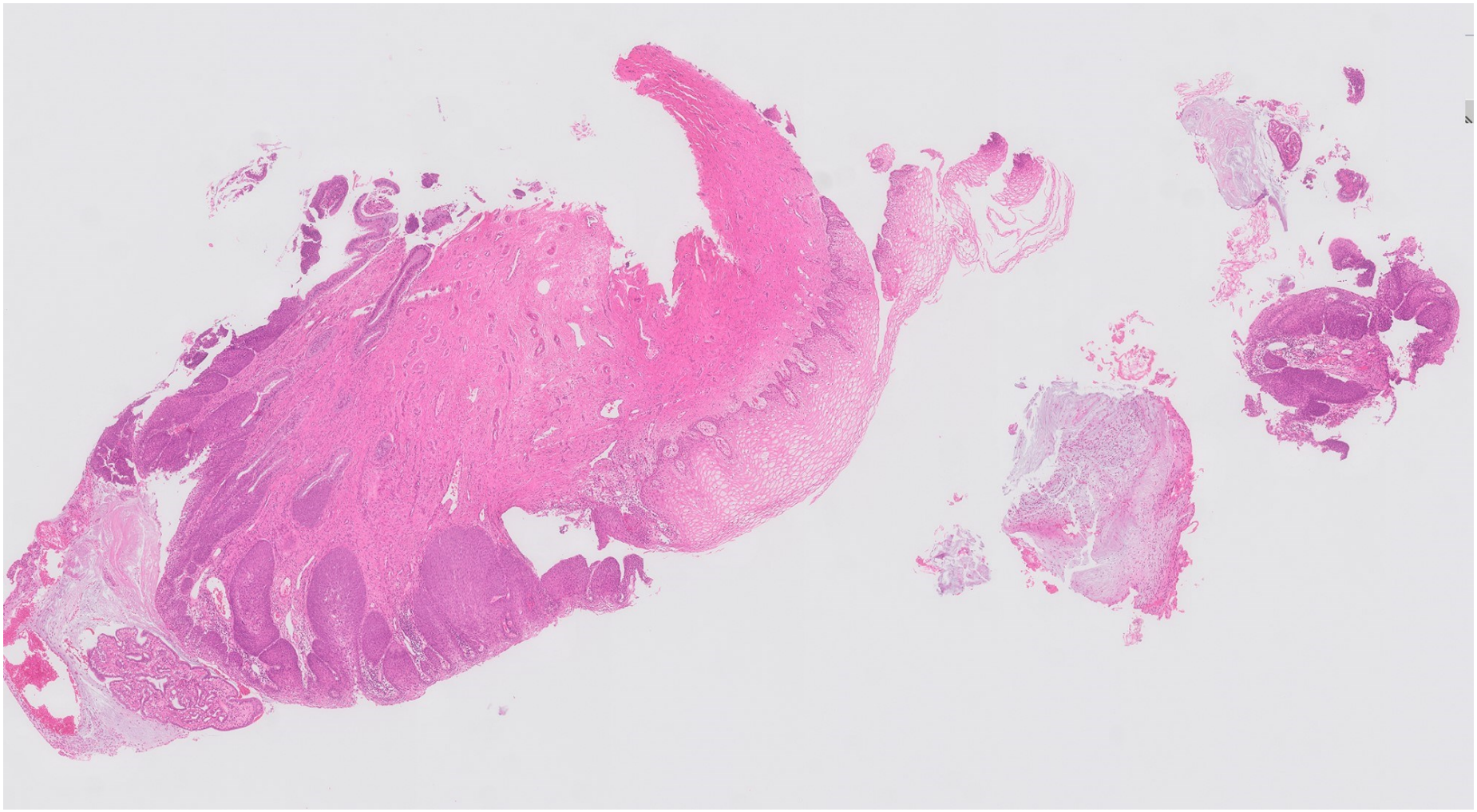


Figure 5a.

Figure 5b.

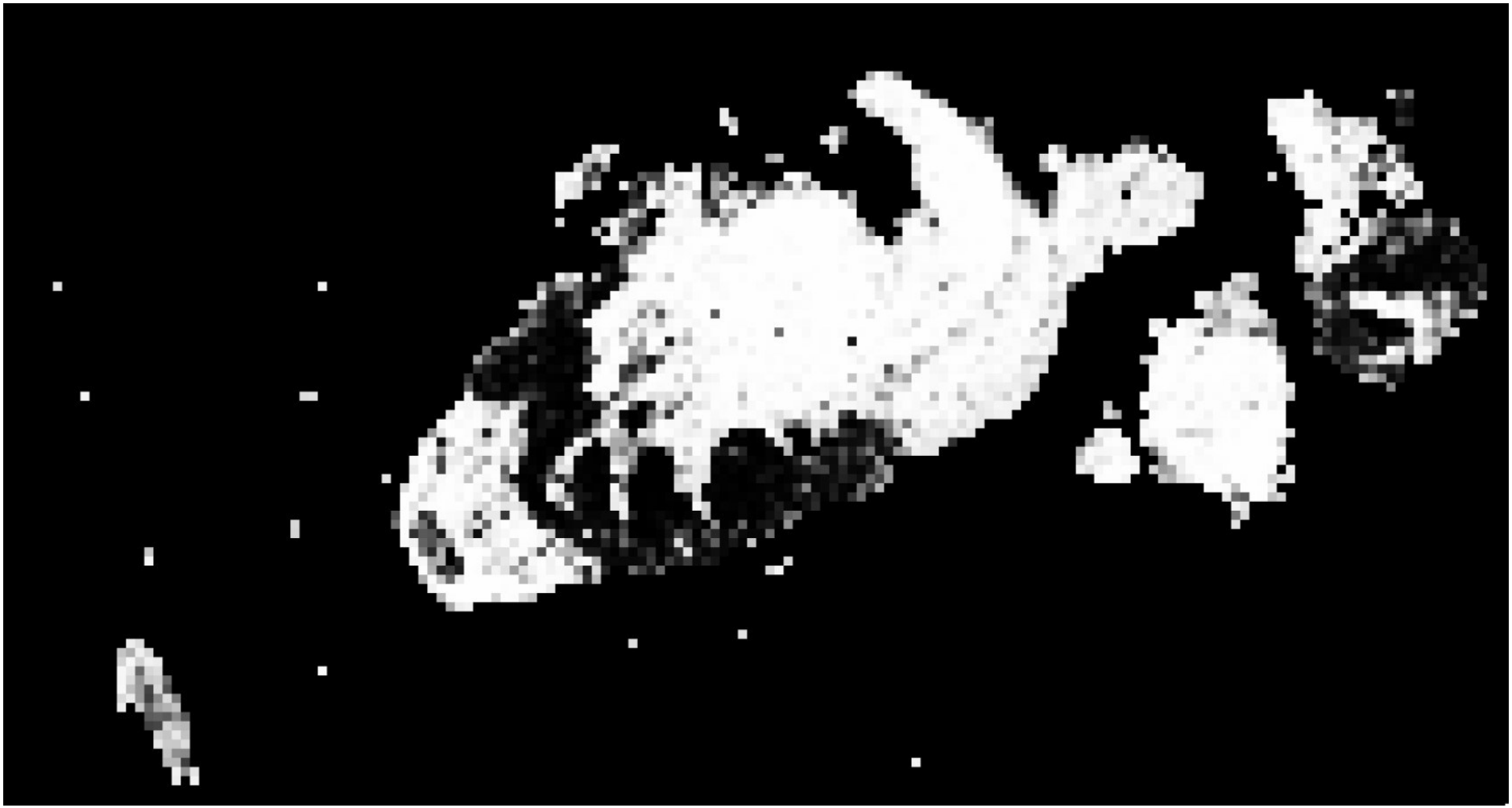


Figure 5c.

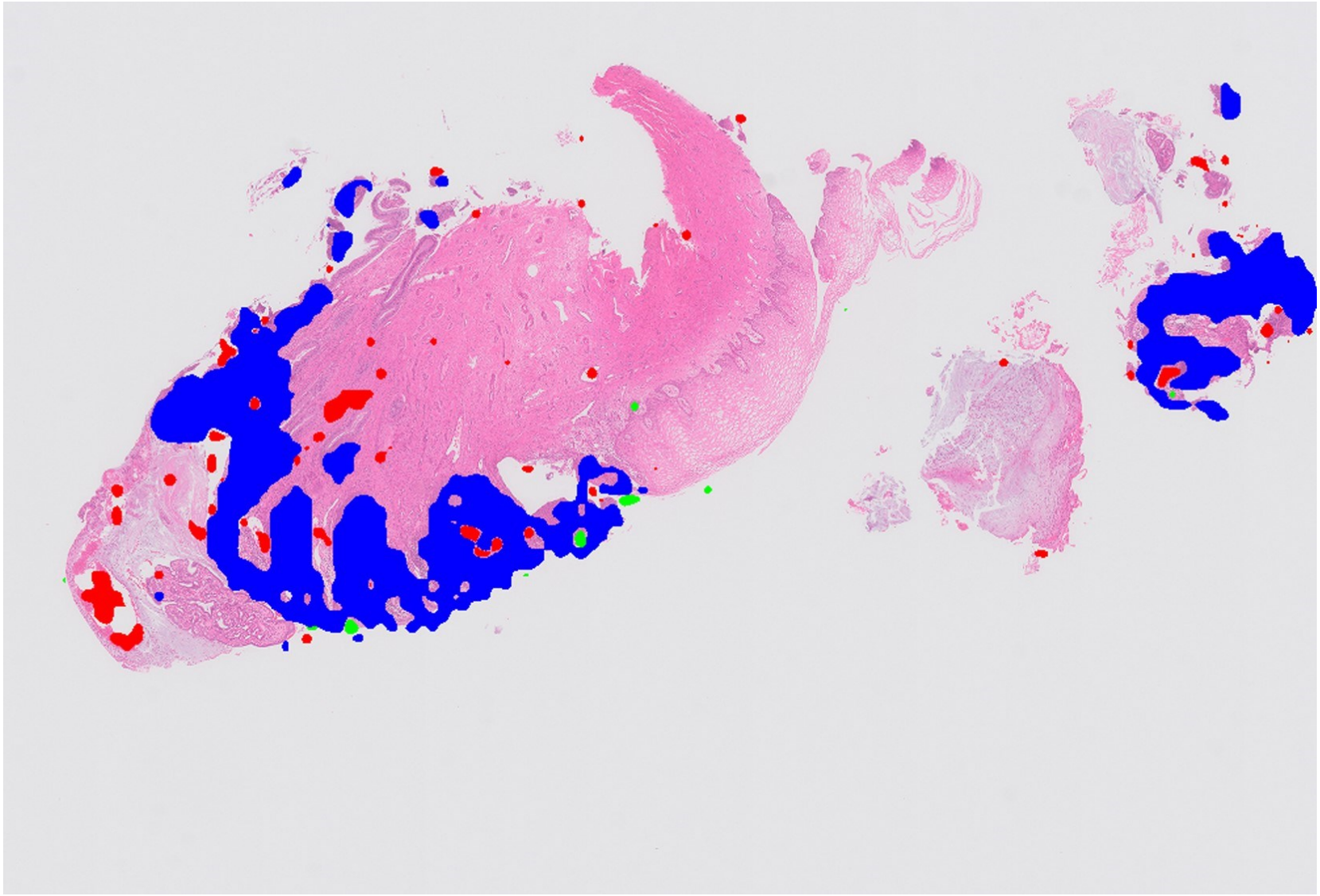


Figure 5d.