Practical guide to training and validation for primary diagnosis with digital pathology

Bethany Jill Williams [1,2] Darren Treanor[1,2]

[1]Department of Histopathology, Leeds Teaching Hospitals NHS Trust, Leeds, UK
[2]Faculty of Medicine and Health, University of Leeds, Leeds, UK

Correspondence to
Dr Bethany Jill Williams, Department of Histopathology, Leeds Teaching Hospitals NHS Trust, Leeds LS9 7TF, UK; bethany.williams2@nhs.net

Received 5 November 2019
Revised 12 November 2019
Accepted 13 November 2019
Published Online First 29 November 2019

ABSTRACT

Numerous clinical pathology departments are deploying or planning to deploy digital pathology systems for all or part of their diagnostic output. Digital pathology is an evolving technology, and it is important that departments uphold or improve on current standards. Leeds Teaching Hospitals NHS Trust has been scanning 100% of histology slides since September 2018. In this practical paper, we will share our approach to training and validation, which has been incorporated into the Royal College of Pathologists' guidance for digital pathology implementation. We will offer an overview of the Royal College endorsed training and validation protocol and the evidence base on which it is based. We will provide practical advice on implementation of the protocol and highlight areas of digital reporting that can prove difficult for the novice digital pathologist. In addition, we will share a detailed topographical list of types of diagnostic tasks and features which should form the basis of digital slide training sets.

INTRODUCTION

Developments in digital pathology technology and the recent Food and Drug Administration regulatory clearance of two whole slide imaging systems for the primary diagnosis of histopathological specimens have resulted in increasing levels of interest in clinical digital pathology deployment on a worldwide scale.[1 2] Digital pathology is still an evolving field, and relatively few clinical pathology departments are reporting high volumes of digital slides to date. In light of this, we feel it is important that more digitally mature departments share their knowledge and experience and highlight examples of best practice, particularly in the field of pathologist training.

When the team at Leeds Teaching Hospitals NHS Trust were developing our validation protocol, we wanted to draw on the existing evidence base for digital diagnosis. A systematic review of digital pathology accuracy synthesising data from 38 peer reviewed validation studies found a mean concordance of whole slide imaging diagnosis and conventional light microscopy diagnosis of 92.4%, compared with a concordance rate of 93.7% for repeat light microscopy review of cases.[3] Given the acknowledged interobserver and intraobserver variability in histopathological diagnosis, this statistic is very encouraging. A more recent review analysed in depth the small number of instances of digital:glass diagnostic discordance.[4] In this study, 8069 documented instances of digital slide and glass slide comparison were found, and among these, 335 instances of diagnostic discordance were recorded: 4% of all digital:glass comparisons. The majority of these discordances represented areas of appreciable diagnostic difficulty and recognised interobserver variation, such as the difference between two adjacent cancer grades. The largest single non-inferiority study of diagnostic discordance using whole slide imaging versus standard light microscopy, which included 1992 cases, found a major discordance rate with the reference standard diagnosis of 4.9% for WSI and 4.6% for standard light microscopy.[5]

THE TRAINING AND VALIDATION PROTOCOL

General principles and summary

Digital pathology remains a relatively novel technology, and while the literature suggests it is safe, there is limited experience of its use in clinical practice. In light of this, a cautious, safety focused approach is recommended by the Royal College of Pathologists,[6] where microscopes are still readily available for slide review where needed.

Any histopathology department will usually house a mixture of enthusiasts and sceptics, and pathologists are a heterogeneous population in terms of their background computer skills, attitude to technology and attitude to risk. A pathologist needs to reach a state where they are not just competent, but confident in their use of the digital pathology reporting system and the validity of their digital diagnosis. A number of approaches are possible, but a successful training and validation procedure should result in:

- Pathologists that are confident in their abilities and their limitations with digital diagnosis.
- Pathologists that are familiar with their hardware and software and can recognise and report performance issues.
- A department with a shared understanding and investment in their digital pathology system.
- A department that can develop bespoke ways of using digital to improve its outputs, workflows and working environment.

The College of American Pathologists guidelines advises that a minimum of 60 cases per use case should be viewed on digital and glass, with a washout period of at least 2 weeks between reads and diagnostic concordance rate observed.[7] This experimental validation design can help a department confirm that their digital pathology system produces diagnostic grade images, but does not offer the individual pathologist an opportunity to gain competence and confidence in digital reporting.

The Royal College of Pathologists recommends training and validation which reflects 'real world'

To cite: Williams BJ, Treanor D. *J Clin Pathol* 2020;**73**:418–422.

Table 1 Summary of Royal College of Pathologist's endorsed validation protocol for digital primary diagnosis

Phase	Overview
Training	One to one formalised training in digital microscope use Observed practice with feedback
Validation—training cases	Training set of approx. 20 challenging and informative cases relevant to the pathologists scope of work Participant views digital slides, makes notes on diagnosis and immediately checks corresponding glass slides, noting any difference in opinion Allows identification and mitigation of pitfalls
Validation—live reporting	All cases scanned prospectively Diagnosis made on digital slides with reconciliation with glass slides prior to sign-out Pathologist aims to complete approx. 2 months whole time equivalent workload in this way Difficulties reported and discussed Library of problematic cases assembled and viewed with group
Summary and recommendations	Validation document produced with each pathologist documenting concordance/discordance Recommendations made for scope of digital practice/further training

diagnostic, with the emphasis on individual professional development.[6] The Leeds validation protocol combines a brief period of hardware and software familiarisation, followed by focused training using cases relevant to the pathologists workload which test potential 'pitfalls' of digital diagnosis and a period of dual reporting, with initial digital assessment followed by a safety check on glass slides.[7] Table 1 summarises the phases of this validation protocol. Use of the protocol for the validation of a cohort of breast pathologists resulted in an observed clinical concordance rate of 98.8%.[8] Pathologists can train singly or in small cohorts, ideally grouped by subspecialty. Ideally, a departmental 'trainer' should oversee the validation of colleagues. This could be a consultant or suitably enthusiastic trainee. Alternatively, pathologists could self-train and self-validate, although discussion with peers is recommended where possible, as this facilitates sharing of and access to a wider range of 'difficult cases' and early discussion of departmental workflows.

Training phase

The aim of the training phase is to allow your pathologist to familiarise themselves with the hardware and software components of your departmental digital pathology system and provide feedback on the pathologist's use of that system to optimise their initial experience of digital reporting. An initial training package could include a group or individual teaching session based on a powerpoint presentation. This presentation should include the following:

- Description of the components of the departmental digital pathology system (scanners, image management software, reporting workstations including diagnostic screens, slide viewing software).
- Stepwise description of the validation/training protocol (outlined in table 1).
- Description of digital pathology workflows in the laboratory.
- Description and examples of common digital image artefacts/system performance issues and how to report these to appropriate team members.
- Commonly encountered areas of diagnostic difficulty on digital slides (these will be discussed later in this article).
- Contact details of key team members who can answer queries regarding digital pathology training, validation, scanning and so on.

At this stage, your pathologist should be given access to a digital copy of the training presentation, your standard operating procedure (SOP) for digital pathology validation, an SOP for digital reporting and a guide/instruction manual to using the digital pathology slide viewer.

After this, it is useful to have an individual session with the pathologist, in which trainer and pathologist open and view training cases. These should include larger, multislide cases which require navigation between slides. The trainer can observe the pathologists use of the mouse/other input device and offer suggestions for ergonomic and efficient navigation of slides and specimens. Basic features of the viewing software, including use of zoom and measurement and annotation tools should be demonstrated, until the pathologist is happy to open, navigate and assess cases without the assistance of the trainer.

Validation training cases

In this part of the validation, the pathologist views a set of pre-prepared educational cases, which are selected to reflect areas of expected diagnostic difficulty on digital and represent learning targets. The slide sets should be assembled from your own departmental archive, so they represent the histology and staining protocols from your own laboratory. Case sets should be assembled which reflect the practice of the individual pathologist—for instance, a breast pathologist should just view breast cases, someone that reports that lung and skin should view a mixture of both topographies. Care should be taken to include a range of tissue types, diagnoses and stains. Guidance regarding the choice of cases can be found in the 'potential pitfalls' section of this paper. It may be helpful to recruit a trainee pathologist to help assemble cases and create topographical training sets—these are also a fantastic resource for trainees to view.

We would suggest no more than 20 cases are assembled for each pathologist. 'Cases' can be a mixture of complete, multislide cases and single representative slides of particular entities. Inclusion of complete resection cases allows the pathologist to test their digital slide navigation skills and competence in use of digital measuring tools, while single slide cases can be used to demonstrate the digital appearance of particular diagnostic features (eg, amyloid, weddelite) and to assess their skills in digital dysplasia grading and mitotic scoring.

Once collected, the glass slides for the training cases should be scanned using the departmental scanning protocol. At Leeds, we recommend 40× equivalent magnification scanning for primary diagnostic work. The pathologist should be given access to the digital slides for the cases and the relevant clinical information pertaining to the case. The pathologist should view the digital slides for a case, record their diagnosis in a workbook and record their confidence in that diagnosis on a Likert scale of 1–7. They should then immediately consult the corresponding glass slides of the case and directly compare the glass slide and digital slide representation. This form of validation by direct

comparison allows the pathologist to appreciate subtle differences in the representation of the case on digital and glass slides and become confident in their interpretation of the digital slide. The pathologist should record any change in their assessment of the case after consulting the glass and again record their diagnostic confidence.

Once the pathologists have viewed all the cases, they can discuss these with the trainer and their colleagues and will hopefully have identified some key areas to concentrate on as they move on to the live reporting phase of the validation.

Validation—live reporting phase

In this phase of the validation, the pathologist is asked to make all their live diagnoses on digital slides, using their own workload. The pathologists make their diagnosis on the digital slides, but with immediate glass reconciliation prior to case sign-out. A whole time equivalent of 2 months allows the pathologist to view an appropriate breadth and depth of cases, including an appropriate mix of biopsies and resections. The length of time needed to gain confidence in digital reporting is likely to vary by pathologist, and some may take longer to navigate the learning curve than others. The pathologist should record all cases viewed and record any alterations made to diagnoses following glass slide review on a spreadsheet. The pathologist should be given regular opportunities to discuss discordant or difficult cases with the trainer/their peer group.

Discordant cases should be collected and used to create a library of 'difficult on digital' training cases, which can be used as a departmental resource for further training.

Validation summary and recommendations

When the pathologists have completed a suitable period of live reporting, their spreadsheet data should be reviewed and concordance and discordance statistics calculated and put into a report. Data reports should include:

- Record of all training meetings.
- Training set concordance rate as a %.
- Detailed description of discordances from the training set.
- Total number of cases viewed in the live reporting phase.
- Number and percentage of concordant cases.
- Detailed description of discordances from the live reporting phase.

Following review of the data, the pathologist and trainer should reach a mutual decision on the result of the validation procedure. There are three possible outcomes:

1. Fully validated for primary digital diagnosis in the specified diagnostic area.
2. Validated for primary digital diagnosis in the specified diagnostic area, with some exceptions.
3. Not validated for primary digital diagnosis in the specified diagnostic area at this time.

In the majority of cases, an outcome '2' will be the most appropriate designation. In this case, the pathologist and trainer should agree on the scope of digital practice and mandate glass slide checks for particular diagnostic scenarios/case types outside of the scope. For instance, if at the end of the validation procedure, the pathologists still lack confidence in mitotic scoring, they could agree to safety net glass slide reconciliation before sign-out for cases with borderline/critical mitotic count scores. As the pathologist gains experience postvalidation, the scope and exceptions can be reviewed and modified as appropriate.

TRAINING POINTS FOR PRIMARY DIGITAL DIAGNOSIS

Experience from Leeds Teaching Hospitals NHS Trust has identified a number of key areas where novice digital pathologists can experience difficulty. Diagnosis of all types of case is possible on the digital microscope, but confident and efficient sign out of all cases will take time and experience. 'Safety nets' such as the use of adjunct immunohistochemistry or glass slide deferral in particular circumstances or for particular types of case can be used and should not be viewed as 'failure' of the digital system. As pathologists' digital reporting experience grow, they will find that the proportion of cases they are comfortable to sign out increases. While relatively little is known about what the minimum specification should be for a digital pathology workstation for primary diagnosis, use of a quality, high resolution screens can improve pathologists' ability to assess some of the more challenging cases and features described below.

Detection of small diagnostic and prognostic objects

The smooth and efficient navigation of digital cases, both between slides in a multislide case and within a slide that requires a high magnification search can be problematic. The initial low magnification, whole slide image displayed on the computer screen can provide a fantastic 'spot diagnosis' of a predominantly architecture-based diagnosis, for example, adenomatous polyp, fibroadenoma, but it can also provide false reassurance. One of the most common diagnostic discordances that can occur when a novice starts digital diagnostic training is missing a small diagnostic or prognostic object.[4] Examples of this include missing a metastasis or micrometastasis in a sentinel lymph node case or failing to identify a single focus of cryptitis in a multislide colonic biopsy series.

It is vitally important that pathologists have sufficient time to adapt and develop their own navigation strategies on the digital microscope. The tried and tested 'lawnmower' technique to ensure complete high power coverage of a slide on the light microscope is difficult to replicate on the digital microscope. Judicious use of whole slide and whole case thumbnails can aid navigation of a digital case, and features such as indicators that warn pathologists of missed slides/regions of slides can help, particularly in the early stages of digital training.

Dysplasia

The diagnosis and grading of dysplasia on the digital microscope is a recurrent theme in the WSI discordance literature and is a potential pitfall for the new digital pathologist. There are two areas of concern here: diagnostic issues at 'low power' and 'high power'. Discordance can result from a failure to detect a focal region of dysplasia on the initial low power assessment of epithelium (eg, in a cervical biopsy). This type of problem is discussed above. The other issue implicated in the misdiagnosis/grading of digital dysplasia relates to the rendering of nuclear detail on digital scans, with some authors implicating poor focus, exacerbated by compression artefact and the limited dynamic range of the WSI. There is a definite learning curve for digital dysplasia assessment, and a validation procedure involving direct comparison of a pathologists digital and glass assessment of dysplasia cases can help the pathologist reconcile their digital and glass dysplasia identification and grading. Routine use of 40× scans for diagnostic biopsies and a high contrast, high resolution, medical grade display can also improve confidence in diagnosis of tricky or borderline cases.

Box 1 Items/features documented as having different appearance on glass slides and WSI

Item or feature
- Eosinophils.
- Neutrophils.
- Mast cells.
- Amyloid.
- Weddelite calcification.
- Mucin.

Mitotic figure counting

Accurate identification and counting of mitoses is another recurrent theme in the digital pathology discordance literature. In the absence of z-stacking, pathologists have to rely on an image captured at a single best plane of focus and cannot adjust this to focus through the depth of the nucleus for chromatin assessment. Similarly to assessment of dysplasia, there is a learning curve for digital mitotic counting. In cases of uncertainty, where the mitotic count on digital is at a critical cut-off level, which would affect overall grading and treatment for a patient, a confirmatory glass slide check should be encouraged. Mitotic counting is an area where artificial intelligence and computer assisted diagnosis could assist the digital pathologist in the near future.

Specific diagnostic items and features

Examination of the literature highlights a number of diagnostic/prognostic items and features which may have a subtly different appearance on a WSI (see box 1). Many of these items share common features: they are often eosinophilic, refractile entities. Other items of particular note include the weddelite form of calcification in breast biopsy specimens and amyloid. Both entities can be viewed on standard WSI images, but experience from validation studies suggests that there is a learning curve for confident recognition on the digital slide.

Potential pitfalls

Table 2 summarises some of the potential pitfalls of digital diagnosis in different diagnostic subspecialties, as evidenced by the validation literature and practical experience of validation. These potential pitfalls should form the basis of digital primary diagnostic training.

Continuing surveillance and audit

Following introduction of digital primary diagnosis, data should be collected routinely on:
- frequency and root cause of poor quality/out of focus/artefact containing WSI,
- frequency and details of instances when pathologists defer to glass slides.

WSI diagnosis can be audited in a similar way to existing departmental glass slide diagnostic audit, with a random sample representing a proportion of the diagnostic workload reviewed by a second pathologist.

CONCLUSION

We have presented a practical guide to advise clinical histopathology departments on how to train and validate their pathologists for primary digital diagnosis, summarised the key steps and considerations and provided a detailed list of evidence-based 'potential pitfalls' and training targets for digital reporting. Digital pathology technology and our

Table 2 Potential pitfalls of digital diagnosis

Histopathology subspecialty	Potential pitfalls
General	Identification and grading of dysplasia
	Identification of lymph node metastasis and micrometastasis
	Identification and quantification of mitotic figures
	Identification of granulation tissue
	Identification of micro-organisms
Breast	Identifying and grading of nuclear atypia
	Identifying microinvasion/lymphovascular space invasion
	Identification of lobular carcinoma
	Grading invasive cancers (mitotic count component)
	Identification of weddelite calcification
	Identification of sentinel node metastasis/micrometastasis
Skin and soft tissue	Identification and grading of squamous dysplasia
	Micro-organism detection
	Granulomatous inflammation
	Melanocytic lesions
	Granulocyte identification and differentiation
	Identification of sentinel node metastasis
	Identification of amyloid
	Identification of lymphoproliferative disease/malignancy
Endocrine	Identification of granulomata
	Identification of lymph node metastasis
	Identification of amyloid in medullary carcinoma of thyroid
	Classification of thyroid neoplasms—identification of cellular papillary features
	Identification of mitoses/atypical mitoses
Genitourinary	Identification and grading of urothelial dysplasia
	Identification of micro-organisms
	Identification of granulomatous inflammation
	Identification/classification of inflammatory cells (granulocyte typing)
	Identification of amyloid
	Identification of lymphoproliferative disease/malignancy
	Grading renal carcinoma (nuclear features)
Gastrointestinal	Identification and grading of oesophageal dysplasia
	Identification of focal activity in inflammatory bowel disease
	Identification of oesinophils in oesophageal biopsies
	Identification of granulomata
	Identification of micro-organisms—particularly *Helicobacter pylori*
Gynaecological	Identifying and grading cervical dysplasia
	Identifying metastasis/micrometastasis
	Assessing endometrial atypia
	Identifying mitotic figures (particularly in soft tissue uterine lesions)
	Identifying mucin
Head and neck	Identification and grading of squamous dysplasia
	Identification of micro-organisms, including fungal forms
	Identification of granulomata
	Identification and typing of inflammatory cells
Hepatobiliary/pancreatic	Interpretation of liver special stains
	Identification of dysplastic epithelium (especially gall bladder)
	Identification and typing of inflammatory cells
	Identification of granulomata
Cardiothoracic	Identification of dysplasia/malignancy in small biopsy specimens
	Identification of micro-organisms including mycobacteria
	Identification of granulomatous inflammation
	Identification of micrometastasis/malignant cells in EBUS (endobronchial ultrasound-guided transbronchial needle aspiration) specimens

Continued

Table 2 Continued

Histopathology subspecialty	Potential pitfalls
	Identification and classification of granulocytes in interstitial lung disease
Neuropathology	Identification of eosinophilic granular bodies
	Identification of necrosis
	Interpretation of nuclear detail
	Identification of mitotic figures
Placenta	Identification and classification of granulocytes
	Identification of nucleated red blood cells

appreciation of the scope and limitations of digital practice continue to evolve, and with this in mind, it is important that the pathology community continues to prioritise the quality and safety of our diagnosis with the introduction of new technologies and techniques.

Take home messages

- Pathologists benefit from a period of training and the opportunity to personally validate their use of the digital microscope.
- Digital pathology training and validation should reflect real world reporting environments as closely as possible while providing a safety net as the pathologist is relatively inexperienced in digital slide assessment.
- Training should encompass safe, comfortable use of digital pathology hardware and software.
- There are specific areas of digital diagnosis and slide assessment that should form the focus of training, including slide navigation, dysplasia assessment and mitotic count scoring.

Handling editor Runjan Chetty.

Contributors BJW and DT designed the validation protocol described in the manuscript. BJW drafted the article, with feedback and review from DT.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests Leeds Teaching Hospitals NHS Trust has a collaborative partnership with Leica Biosystems for a research digital pathology deployment. Both authors are part of the Northern Pathology Imaging Co-Operative.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; internally peer reviewed.

Data availability statement Data sharing not applicable as no datasets generated and/or analysed for this study. No data are available.

ORCID iD
Bethany Jill Williams http://orcid.org/0000-0002-6641-5503

REFERENCES

1 US Food and Drug Administration. FDA allows marketing of first whole slide imaging system for digital pathology. News release, April 12 2017.
2 FDA Reporter. Leica Biosystems: Receives FDA 510(k) clearance to market a digital pathology system for primary diagnosis. Press release May 31, 2019 2019.
3 Goacher E, Randell R, Williams B, *et al*. The diagnostic concordance of whole slide imaging and light microscopy: a systematic review. *Arch Pathol Lab Med* 2017;141:151–61.
4 Williams BJ, DaCosta P, Goacher E, *et al*. A systematic analysis of discordant diagnoses in digital pathology compared with light microscopy. *Arch Pathol Lab Med* 2017;141:1712–8.
5 Mukhopadhyay S, Feldman MD, Abels E, *et al*. Whole slide imaging versus microscopy for primary diagnosis in surgical pathology: a multicenter blinded randomized Noninferiority study of 1992 cases (pivotal study). *Am J Surg Pathol* 2018;42:39–52.
6 Royal College of Pathologists. Best practice recommendations for digital pathology, 2018. Available: https://www.rcpath.org/resourceLibrary/best-practicerecommendations-for-implementing-digital-pathology-pdf
7 Pantanowitz L, Sinard JH, Henricks WH, *et al*. Validating whole slide imaging for diagnostic purposes in pathology: guideline from the College of American pathologists pathology and laboratory quality center. *Arch Pathol Lab Med* 2013;137:1710–22.
8 Williams BJ, Hanby A, Millican-Slater R, *et al*. Digital pathology for the primary diagnosis of breast histopathological specimens: an innovative validation and concordance study on digital pathology validation and training. *Histopathology* 2018;72:662–71.