

A Systematic Analysis of Discordant Diagnoses in Digital Pathology Compared With Light Microscopy

Bethany J. Williams, MB, BS, BSc; Philip DaCosta, MBBS, MRCS, LRCP, FRCPath; Edward Goacher, BSc; Darren Treanor, MB, BSc, PhD, FRCPath

• **Context.**—Relatively little is known about the significance and potential impact of glass-digital discordances, and this is likely to be of importance when considering digital pathology adoption.

Objective.—To apply evidence-based medicine to collect and analyze reported instances of glass-digital discordance from the whole slide imaging validation literature.

Design.—We used our prior systematic review protocol to identify studies assessing the concordance of light microscopy and whole slide imaging between 1999 and 2015. Data were extracted and analyzed by a team of histopathologists to classify the type, significance, and potential root cause of discordances.

Results.—Twenty-three studies were included, yielding 8069 instances of a glass diagnosis being compared with a digital diagnosis. From these 8069 comparisons, 335 instances of discordance (4%) were reported, in which

glass was the preferred diagnostic medium in 286 (85%), and digital in 44 (13%), with no consensus in 5 (2%). Twenty-eight discordances had the potential to cause moderate/severe patient harm. Of these, glass was the preferred diagnostic medium for 26 (93%). Of the 335 discordances, 109 (32%) involved the diagnosis or grading of dysplasia. For these cases, glass was the preferred diagnostic medium in 101 cases (93%), suggesting that diagnosis and grading of dysplasia may be a potential pitfall of digital diagnosis. In 32 of 335 cases (10%), discordance on digital was attributed to the inability to find a small diagnostic/prognostic object.

Conclusions.—Systematic analysis of concordance studies reveals specific areas that may be problematic on whole slide imaging. It is important that pathologists are aware of these areas to ensure patient safety.

(*Arch Pathol Lab Med.* 2017;141:1712–1718; doi: 10.5858/arpa.2016-0494-OA)

The capacity to digitally capture, view, analyze, store, share, and view whole slide pathology images has led to widespread adoption of digital pathology for education and research in the health care sector and institutes of

higher education.^{1–3} Whole slide imaging (WSI) systems are used increasingly in Europe and North America for secondary diagnosis (eg, for second opinions or frozen section cases), but their use for large-scale primary diagnosis remains limited to a few centers internationally, including projects in Sweden, the Netherlands, and Canada.^{4–6} Implementation barriers include the cost-benefit of digital pathology, current lack of regulatory approval from the US Food and Drug Administration, a paucity of evidence validating the diagnostic concordance of WSI and light microscopy, and low acceptability of digital pathology among pathologists.^{1,7–11} If digital pathology is to be implemented for primary diagnosis on a large scale, regulatory bodies, both international and national; departmental heads; and individual pathologists will have to be confident that a diagnosis made on a digital microscope is equivalent to a diagnosis made by the same pathologist on a light microscope.

A limited number of studies have compared the diagnostic concordance of WSI and traditional light microscopy. In 2016, Goacher et al¹² undertook a systematic review of these studies in which study quality was assessed. They identified 38 studies, and reported a mean diagnostic concordance of WSI and light microscopy, weighted by the number of cases per study, of 92.4%. Of the 30 studies quoting concordance as a percentage, 60% showed a concordance of 90% or greater, of which 10 showed a concordance of 95% or greater. There was a trend for increasing concordance in the

Accepted for publication February 22, 2017.

Published as an Early Online Release May 3, 2017.

From the Departments of Histopathology (Dr Williams) and Cellular Pathology (Dr Treanor), Leeds Teaching Hospitals NHS Trust UK, University of Leeds, Leeds, United Kingdom; the Department of Histopathology, Airedale NHS Foundation Trust, Keighley, United Kingdom (Dr DaCosta); the Faculty of Medicine and Health, University of Leeds, Leeds, United Kingdom (Mr Goacher); and the Department of Cellular Pathology, Linköping University, Linköping, Sweden (Dr Treanor).

Dr Treanor is on the advisory board of Sectra (Sectra AB, Linköping, Sweden) and Leica Biosystems/Aperio (Vista, California) as part of Leica Biosystems Imaging, Inc, and has a collaborative research project with FFEI in 2014–2015 (FFEI Ltd, Hemel Hempstead, United Kingdom) where technical staff were funded by them. Dr Treanor is a coinventor on a digital pathology patent that has been assigned to Roche-Ventana on behalf of his employer in 2015. Dr Treanor receives (or will receive) no personal remuneration for any. The other authors have no relevant financial interest in the products or companies described in this article.

The content of this paper was presented at a platform talk at Pathology Informatics 2016; May 25, 2016; Pittsburgh, Pennsylvania.

Reprints: Bethany J. Williams, MB, BS, BSc, Histopathology Department, St James University Hospital, Beckett St, Leeds LS9 7TF, United Kingdom (email: bethany.williams2@nhs.net).

Category	Description
A	Inadequate dissection, sampling, or macroscopic description
B	Discrepancy in microscopy <ol style="list-style-type: none"> 1. A diagnosis that one is surprised to see from any pathologist 2. A diagnosis that is clearly incorrect, but that one is not surprised to see a small percentage of pathologists suggesting 3. A diagnosis where interobserver variation is known to be large (eg, difficult diagnosis, difference between 2 tumor grades)
C	Discrepancy in clinical correlation
D	Failure to seek a second opinion in an obviously difficult case
E	Discrepancy in report (includes misidentification)

Category	Description
1	No impact on care
2	Minimal harm, no morbidity Delay in diagnosis or therapy only, <3 of 12, unnecessary noninvasive further diagnostic efforts, unnecessary therapy without morbidity
3	Minor harm, minor morbidity Delay in diagnosis or therapy only, >3 of 12, unnecessary invasive further diagnostic efforts, delay in therapy with minor morbidity
4	Moderate harm, moderate morbidity Due to delay in diagnosis, due to otherwise unnecessary diagnostic efforts, due to otherwise unnecessary therapeutic efforts
5	Major harm, major morbidity Loss of limb, organ, or function of organ system due to unnecessary diagnostic efforts; severe morbidity due to delayed or unnecessary therapeutic efforts; death

more recent studies. The review found evidence to support a high level of diagnostic concordance for WSI overall.

The conclusions of the systematic review can be interpreted as encouraging for a diagnostic department that is considering a primary diagnostic digital adoption; however, if we invert the statistic, 92.4% concordance equates to 7.6% discordance. It can be argued that discordances are more valuable than concordances in analyzing the potential patient safety impact of digital diagnosis, and an evaluation of the type, severity, unexpectedness, and root cause of these discordant diagnoses can allow us to explore safety aspects of digital pathology adoption and identify potential pitfalls—areas of digital diagnostic interpretation that may require more attention or practice before full digital adoption for primary diagnosis. Experience in analyzing error rates and types in telepathology, including that published by Dunn et al,¹³ has contributed greatly to our understanding of the limitations and strengths of this diagnostic medium, and gathering and analyzing similar data from WSI studies is likely to prove equally beneficial.

The primary aim of this review and analysis was to systematically examine the published literature on discordant pathologic diagnoses rendered by WSI compared with those rendered by light microscopy, and to identify areas that may be problematic to diagnose using digital microscopy.

MATERIALS AND METHODS

A systematic review protocol that had been used in a previous systematic review of WSI concordance was used. The review protocol is registered with the PROSPERO database (registration number CRD42015017859), and can be accessed online.¹⁴

Search Strategy

An electronic search was instigated on the databases Medline, Medline in Progress, EMBASE, and the Cochrane Library between 1999 and December 2015, using the previously published systematic review methodology.^{12,14} A search of clinicaltrials.gov (Bethesda, Maryland) was performed to identify any ongoing studies. Included studies underwent manual reference searching and citation tracking through PubMed and Google Scholar. Corresponding authors were contacted, where possible, to identify subsequent or ongoing research.

Paper Screening

Two reviewers independently subjected the abstracts of papers to the previously used systematic review screening algorithm.¹² In

cases of disagreement, a third independent reviewer was consulted. Full texts of all papers that fulfilled the initial screening algorithm were retrieved and reviewed. Only published journal articles were included in the review.

Data Extraction and Analysis

A standardized data extraction protocol was applied to all included studies. The lead researcher extracted pairs of discordant diagnoses (preferred diagnosis with discordant diagnosis) that were stored in a spreadsheet in which the source study and the method of diagnosis (glass or digital) used to render each diagnosis were concealed from the reviewers. A team of 3 discordance reviewers was assembled, all of whom were professional diagnostic pathologists, with 6, 18, and 34 years of pathology experience. The 3 discordance reviewers evaluated each diagnostic pair and assigned it a category based on the Royal College of Pathologists System of Categorization for Discrepancies.³⁸ In this system, discordances are assigned a letter code depending on the type of error (ie, errors in macroscopy, microscopy, clinical correlation, failing to seek a second opinion, misidentification). For this study, the B category, discrepancies in microscopy, was the most relevant. The B category errors were then stratified depending on how unexpected or understandable the error was (Table 1). Next, each reviewer assigned each discordant diagnostic pair a category corresponding to the potential for patient harm to be caused, from the Royal College of Pathologists guide to duty of care reviews (Table 2).³⁹ The spectrum of harm ranges from *no clinical impact, no harm*, which is categorized as 1, to *severe harm*, categorized as 5. Dimensions such as delay in diagnosis, unnecessary further diagnostic efforts, delays in therapy, unnecessary therapy, and resultant levels of morbidity or mortality were considered.

All discordances were reviewed independently by the 3 discordance reviewers. For the potential for harm categorizations, the Royal College categories 2 and 3 (minimal harm, no morbidity, and minor harm, minor morbidity) were merged into a single category of minimal/minor harm, and categories 4 and 5 (moderate harm, moderate morbidity, and major harm, major morbidity) were merged into a single category of moderate/major harm. Where reviewers disagreed on the categorization of a diagnostic discordance, cases were discussed and a consensus reached. Expert opinion was sought on the renal transplant biopsy data, as the review team did not feel they had sufficient subspecialty expertise in this area (Dr Carole Angel, MB, ChB, Leeds Teaching Hospitals NHS Trust, written communication, March 2016).

Table 3. Example of a Discordance Analysis for a Colonic Biopsy

Parameter	Example
True diagnosis	Focal active colitis
Discordant diagnosis	Normal colon
Preferred diagnostic modality	Glass
RCPATH expression of concern	B2 (clearly incorrect, but would expect a small proportion of pathologists to make the same error)
RCPATH potential for harm	3 (minor harm, minor morbidity)
Discordance type	Missed diagnosis
Diagnostic tasks/objects/features	Finding small object, neutrophils
Detail from paper	Granulocytes difficult to discern on digital. Improved at ×40.

Abbreviation: RCPATH, Royal College of Pathologists.

After all cases had been categorized, the lead researcher reunited the discordant pairs with source study data. For each individual discordance, the source paper was examined to extract data on the preferred diagnostic modality for that particular case; the type of diagnosis required; the type of discordance; any specific diagnostic tasks, objects, or features that would have enabled the pathologist to make the true diagnosis; and detail from the paper of any particular difficulties/observations encountered by the study pathologists. See Table 3 for an example of a discordance analysis.

RESULTS

Study Demographics

One thousand three hundred abstracts were checked and 39 full-text papers extracted. Of these, 23 contained detailed, extractable discordant diagnostic pair data.^{15–37} Publication dates ranged from 2006 through 2015, with the majority of studies published post-2010. These 23 papers included 8069 instances of a glass diagnosis and a digital diagnosis being compared. Out of these 8069 glass-digital read pairs, 335 instances of discordance were recorded, which represents approximately 4% of 8069 glass-digital comparisons.

The included studies used a range of scanners from 7 different vendors. Viewing hardware varied greatly both within and among studies. Many studies provided little information on viewing hardware/scanners or failed to standardize viewing hardware. The majority of studies scanned slides at a routine magnification of ×20, with more recent publications tending to use ×40, and some varying the scanning magnification depending on the type of case, for example, diagnostic specimens at ×40 and therapeutic specimens at ×20.

The majority of included studies scanned a mixture of cases from a number of histopathology subspecialties. Ten studies were a mix of 2 or more distinct organ systems, which we termed a case mix, with gastrointestinal and skin the most popular single pathology specialties examined. The majority of recorded discordances occurred in gastrointes-

tinal, skin, genitourinary, and gynecologic cases. Unfortunately, many of the source studies lacked a sufficiently detailed breakdown of case types included, but it is likely that cases from these organ systems are overrepresented in the source studies; they are certainly all high-throughput specialties.

Severity and Implications of Discordance

Of the 335 reported discordances, glass was the preferred diagnostic modality in 286 cases (85%). Interestingly, the digital diagnosis was preferred in 44 cases (13%), with an equivocal response in the remaining 5 (2%).

The largest specific category of discordance was missed diagnosis of malignant/dysplastic/atypical conditions, where malignant tissue was given a benign diagnosis. In these cases, glass was the preferred diagnostic modality in 66 of 77 cases (86%). There were also 25 cases where benign tissue was erroneously diagnosed as malignant/atypical. Here glass was the preferred diagnostic modality in 23 cases (92%) (Table 4). The second greatest discordance type (70 cases) was where a case was recognized as malignant/atypical but incorrectly typed or graded. Here again, glass was the preferred diagnostic modality in 67 cases (96%). Discrepancies in the diagnosis of inflammation were also common.

Most discordances (169 cases) fell into the category of B3, areas of appreciable diagnostic difficulty and recognized interobserver variation, such as the difference between 2 adjacent grades of a malignant condition.

In total 21 B1 diagnoses were recorded. These would be regarded as surprising errors using the Royal College of Pathologists' System of Categorization for Discrepancies.³⁸ One type E discordance was recorded—a misidentification error, where digital was the preferred diagnostic modality (Table 5).

The majority (242; 72%) of the 335 discordances reported had the potential to cause minimal or minor harm to

Table 4. Types of Discordance

	Glass Preferred	Digital Preferred	Total
Missed malignant/dysplastic/atypical diagnosis	66	11	77
Erroneous malignant diagnosis	23	2	25
Difference in malignant diagnosis	67	3	70
Missed inflammation	40	2	42
Erroneous diagnosis of inflammation	6	2	8
Difference in diagnosis of inflammation	15	2	17
Invasion missed	2	5	7
Erroneous invasion	4	0	4
Other	63	17	80
Total	286	44	330

Table 5. Types of Discordance

Category and Description	Glass Preferred	Digital Preferred	Total
B1 Surprising error	19	2	21
B2 Expect small number of pathologists to make this error	123	16	139
B3 Area of appreciable difficulty/interobserver variation	144	25	169
E Misidentification	0	1	1

patients. This represents 3.00% of all glass-digital comparisons (242 of 8069). Only 28 of 335 (8%) had the potential to cause moderate or major harm to patients. This represents 0.35% (28 of 8069) of all glass-digital comparisons. For these, glass was the preferred diagnosis in 26 (93%). Digital was preferred in 2 of 28 cases (7%) with the potential for moderate/major harm (Table 6). Table 7 shows specific instances of major/moderate harm recorded on diagnoses made using digital and conventional glass slides. Instances where the glass slide diagnosis was preferred included benign breast tissue erroneously reported as invasive carcinoma on the digital read and a benign lung biopsy erroneously diagnosed as non-small cell carcinoma. Examples where malignant diagnoses were missed on the digital read of a case included gastric adenocarcinoma called acute gastritis, metastatic melanoma missed in a lymph node, and chronic lymphocytic leukemia missed in a skin biopsy. Digital was the preferred diagnostic modality for 2 cases with the potential for moderate/major harm: a carcinoid tumor that was missed in the glass examination of an appendix, and benign breast tissue erroneously diagnosed as ductal carcinoma in situ on glass.

Types of Discordance

Dysplasia Diagnosis.—The included studies reported 108 of 335 discordances concerning the diagnosis of dysplasia, representing 32% of all reported discordances. These were predominantly cases from the upper gastrointestinal tract and the cervix. Dysplasia is an area of appreciable interobserver and intraobserver variation, but nonetheless, dysplasia discordances seemed to be particularly prevalent. The majority of discordances were instances of missed diagnosis, where dysplastic tissue was diagnosed as benign or reactive tissue. Fifty-one cases of this type were reported, and they represented 47% (108) of all dysplasia-related discordances.

Interestingly, where there were differences in grading, dysplastic lesions tended to be undercalled on the digital microscope (33 cases undercalled, 8 cases overcalled). There were also errors in the other direction, with erroneous dysplasia diagnosed in benign tissue (14 cases) and a smaller number of overcalls in grading (10 cases). Of all the discordant dysplasia diagnoses, glass diagnosis was preferred in 101 of 108 cases (94%). This indicates that diagnosis and grading of dysplasia may be a pitfall of digital diagnosis (Table 8).

Locating Small Diagnostic Objects/Features.—Another common diagnostic feature implicated in discordance is the ability to find or not find a small diagnostic/prognostic object. Thirty-nine discordances of this type were recorded, with glass the preferred diagnostic medium in 30 of 39 (82%). The majority of these discordances would be classified as B2 errors in microscopy, which one expects to see in a small proportion of cases as a matter of course. Three small object location discordances were classified as surprising errors based on the context. In total, 5 of the

small object location discordances could have resulted in moderate/major patient harm. The types of small object missed included a range of malignant and benign features. Perhaps the most concerning of these are small tumors, metastases, and microsattellites. The most common small objects missed were foci of inflammation, more specifically cryptitis in colon biopsies. The detection of microorganisms was also a theme raised in the literature (Table 9).

Specific Problematic Entities Reported in the Literature.—Granulocytes were mentioned in 27 of 335 instances of discordance (11%). Disparities in detection of granulocytes, particularly eosinophils, may relate to differences in color of the cells on digital and their refractile textures. Similar issues with detection of other eosinophilic, refractile objects (nucleated red blood cells and eosinophilic granular bodies) were reported.

Difficulties were also described identifying 2 entities commonly recognized on the grounds of subtle textural and tinctorial qualities: blue mucin and amyloid. In the 2 reported cases of difficulty with amyloid,²² study participants were unable to detect the textural quality of amyloid on digital slides, which would have alerted them to examine the original glass slides with a polarizer (Table 10).

Misidentification Errors.—There was one confirmed misidentification error, where the reader providing a glass diagnosis viewed the wrong slides, and the digital read rendered the correct diagnosis.

DISCUSSION

The use of digital pathology in the clinic is increasing, with many departments piloting digital pathology in primary diagnostic settings. In light of this, guidance is needed regarding potential safety implications for patients. A number of validation studies have been reported in the literature, but as we discovered in our systematic review,¹² these vary greatly in terms of the number and types of participants and cases, the methodology, and the technologies examined. In the absence of a multicenter clinical trial, a systematic review remains the highest level of digital pathology concordance evidence available for those engaged in regulatory efforts.

Goacher et al¹² found a mean diagnostic concordance of WSI and light microscopy, weighted by the number of cases per study, of 92.4%. In this study, we aimed to complement this work with a systematic analysis of the discordant diagnoses reported in the validation literature, in the hope that this analysis would allow a more precise evaluation of primary digital diagnosis.

Table 6. Potential Harm of Discordant Diagnoses

	Glass Preferred	Digital Preferred	Total
No impact	51	9	60
Minimal/minor harm	209	33	242
Moderate/major harm	26	2	28

Table 7. Major/Moderate Harm Examples

Preferred Diagnostic Modality	Discordant Diagnosis	True Diagnosis
Glass	Invasive carcinoma with lobular features	Benign breast tissue
Glass	Non-small cell lung cancer	Chronic bronchitis
Glass	Acute gastritis	Adenocarcinoma
Glass	Lymph node, no abnormality detected	Metastatic melanoma
Glass	Dermatitis	Skin infiltration with chronic lymphocytic leukemia
Digital	Acute appendicitis	Goblet cell carcinoid tumor
Digital	Ductal carcinoma in situ	Fibrocystic change

Dysplasia

The diagnosis and grading of dysplasia is implicated as a possible pitfall of digital diagnosis. Most papers emphasize blurring of nuclear detail on digital scans, and implicate poor focus, exacerbated by compression artifact and limited dynamic range. These explanations focus on high-power diagnosis, but we should also consider low power. In cervical biopsies, dysplasia is often focal, and if focal abnormality is not picked up on the low-power assessment of the epithelium, confirmatory nuclear detail cannot be appreciated at high power. We may also need to consider the effect of scanning magnification and viewing hardware quality.

What potential strategies do we have to mitigate the risks of diagnosing dysplasia digitally? The first thing we could do is ensure pathologists are aware that dysplasia is a potential pitfall. Ordi et al³⁴ describe an increase in glass-digital concordance for cervical dysplasia as their study progressed, suggesting that there is a significant learning curve effect for digital dysplasia diagnosis. This is an area of diagnosis that may need a longer settling-in period before pathologists can confidently and safely sign out digital cases.

Pathologists working in relevant specialties might want to consider a self-validation procedure, with digital-glass reconciliation of dysplasia diagnoses while they establish satisfactory glass-digital concordance. Alternatively, there might be a role for optional or mandatory checks on glass following a digital assessment in particular scenarios—for example, diagnosing dysplasia in Barrett esophagus, a practice used in some digital pathology deployments (Anna Boden, MD, Linköping; David Snead, MB, BS, Coventry; verbal communications, January 2016). Some authors describe limited improvement in digital dysplasia diagnosis with slides scanned at ×40, so there may be justification for mandatory scanning of selected specimens at ×40 (eg, cervical biopsies, upper gastrointestinal biopsies). Unfortunately, there are insufficient data at present to judge whether scanning at ×20 versus ×40 has a significant impact on overall discordance rates, and this is an area that deserves more attention in future studies.

Locating Small Diagnostic Objects/Focal Diagnostic Features

Locating small diagnostic objects is highlighted as a potential problem on digital slide reads. Navigation is certainly implicated, both within and among slides, and the effects of display resolution and scanning magnification also warrant consideration. In many studies, authors explicitly state that pathologists found navigating cases cumbersome. Specific training in safe and efficient navigation strategies using digital software should be available to pathologists who are expected to use digital images clinically.

Appropriate use of whole slide and whole case thumbnails can aid navigation, and safety features such as indicator lights to warn pathologists of missed slides/regions could help. There may be a case for modifying workflows to incorporate a mandatory glass check, at least in the initial phases of digital deployment, for specimens such as sentinel lymph nodes, where detection of micrometastases should be optimal.

There is little evidence in the literature regarding minimum specification for viewing hardware or standardization of viewing hardware. Many authors found diagnostic biopsies, particularly where detection of inflammatory disease is important, were best scanned at ×40, with Snead et al³⁵ recommending ×60 in cases where detection of microorganisms is a priority.

Specific Objects/Features Causing Diagnostic Difficulty

Examination of the literature highlights a number of specific entities, including granulocytes, nucleated red blood cells, and amyloid, that were reported as having a different appearance on glass and digital slides. The importance or relevance of identifying these entities will vary among different subspecialties, and possibly among different pathologists, but it is important to mention areas where investigators have noticed an appreciable difference in appearance.

Specialty pathologists need to be aware of specialty specific diagnostic pitfalls and decide how important these features are to their own practice. Bauer and Slaw²³ report that scanning gastrointestinal biopsies at ×40 improved the ability of their pathologists to detect and correctly categorize granulocytes. Color calibration may potentially play a role.

Misidentification Errors

A single case of misidentification error was reported in the review source literature. In this case, the correct diagnosis was rendered on the digital slide, and the glass slide reviewer viewed the incorrect glass slide. It perhaps reminds us of the potential digital technology provides us to avoid the type of pathology error that should never occur: the misidentification of specimen, slides, or reports. The types of study design used in the source material for this analysis are unlikely to expose the full extent of misidentification errors,

Table 8. Dysplasia Diagnoses

	Glass Preferred	Digital Preferred	Total
Missed diagnosis	47	4	51
Grading undercall	33	0	33
Erroneous diagnosis	13	1	14
Grading overcall	8	2	10
Total	101	7	108

Category	Object	Glass Preferred	Digital Preferred
Neoplasia	Small primary tumor	1	2
	Lymph node metastasis	3	
	Tumor microsatellite	1	
	Focal tumor invasion	3	3
Inflammation	Foci of inflammatory activity	8	
	Granulomas	3	
Microorganisms	<i>Helicobacter pylori</i>	2	
	<i>Candida</i>	3	
Sparse cells	Reed-Sternberg cells	2	
Focal benign features		5	2
Focal immunopositivity		1	

which should be considered when evaluating the total impact of digital versus glass technology on diagnostic error rates.

CONCLUSIONS

Given the increasing trend toward using digital pathology for clinical diagnosis, including primary diagnosis, the need for evidence-based digital pathology guidelines and a systematic evaluation of the available evidence is paramount. In this analysis of 8069 comparisons of glass and digital diagnoses, we found 335 discordances. Of these, only 28 had the potential to cause moderate or major patient harm.

We identified a number of problem areas in digital diagnosis that warrant further exploration and explanation, namely the identification and grading of dysplasia, the location of small diagnostic objects and features, and the identification of certain specialty-specific diagnostic features.

This information can be used to inform safe departmental or institutional adoption of digital pathology and to help design systems and process to address these areas in the future. We believe that although digital deployment for primary clinical diagnosis is in its infancy, it is important to collect and share data on pitfalls and problem cases. To this end, it might be helpful to create a centralized database of cases, recorded in a standardized format.

Education and continuing professional development of pathologists on an individual level is vital to ensure a safe and responsible rollout of digital microscopy. Pathologists should be encouraged to gain confidence in risk-free or risk-mitigated diagnostic environments before adopting a 100% digital workflow.

The perceived success or failure of digital pathology in a specific laboratory will stand or fall based on the competency and confidence of individual pathologists, and it is therefore important that pathologists understand the strengths and limitations of the WSI systems. The studies

included in this analysis used a wide variety of different scanners, with different characteristics that could affect diagnostic interpretation of slides. In light of this, we believe it is important that diagnostic departments perform their own whole-system validations for WSI, to evaluate the strengths and weaknesses of the combination of hardware and software components they propose to use for primary diagnosis.

References

- Huisman A, Looijen A, van den Brink SM, van Diest PJ. Creation of a fully digital pathology slide archive by high-volume tissue slide scanning. *Hum Pathol*. 2010;41(5):751-757.
- Ayad E. Virtual microscopy beyond the pyramids, applications of WSI in Cairo University for E-education & telepathology. *Anal Cell Pathol*. 2011;34(4):195-196.
- Pantanowitz L, Szymas J, Yagi Y, Wilbur D. Whole slide imaging for educational purposes. *J Pathol Inform*. 2012;3:46.
- Thorstenon S, Molin J, Lundström C. Implementation of large-scale routine diagnostics using whole slide imaging in Sweden: digital pathology experiences 2006-2013. *J Pathol Inform*. 2014;5(1):14.
- Stathonikos N, Veta M, Huisman A, van Diest PJ. Going fully digital: perspective of a Dutch academic pathology lab. *J Pathol Inform*. 2013;4:15.
- Evans AJ, Chetty R, Clarke BA, et al. Primary frozen section diagnosis by robotic microscopy and virtual slide telepathology: the University Health Network experience. *Hum Pathol*. 2009;40:1070-1081.
- Romero Lauro G, Cable W, Lesniak A, et al. Digital pathology consultations—a new era in digital imaging, challenges and practical applications. *J Digit Imaging*. 2013;26(4):668-677.
- Brick KE, Sluzevich JC, Cappel MA, DiCaudo DJ, Comfere NI, Wieland CN. Comparison of virtual microscopy and glass slide microscopy among dermatology residents during a simulated in-training examination. *J Cutan Pathol*. 2013;40(9):807-811.
- Hanna MG, Pantanowitz L, Evans AJ. Overview of contemporary guidelines in digital pathology: what is available in 2015 and what still needs to be addressed? *J Clin Pathol*. 2015;68(7):499-505.
- Faison T. FDA regulation of whole slide imaging (WSI) devices: current thoughts. US Food and Drug Administration Web site. http://ftp.cdc.gov/pub/CLIAAC_meeting_presentations/pdf/Addenda/cliac0212/Tab_15_Faison_CLIAC_2012Feb14_Whole_Slide_Imaging.pdf. Accessed February 2, 2016.
- Digital Pathology Association. DPA Recommends whole slide imaging manufacturers submit de novo applications to the FDA for primary diagnosis in the United States. <http://www.prweb.com/releases/2016/01/prweb13163307.htm>. Accessed February 29, 2016.
- Goacher E, Randell R, Williams BJ, Treanor D. The diagnostic concordance of whole slide imaging and light microscopy: a systematic review. *Arch Pathol Lab Med*. 2016;141(1):151-161.
- Dunn BE, Choi H, Almagro UA, Recla DL, Krupinski EA, Weinstein RS. Routine surgical telepathology in the Department of Veterans Affairs: experience related improvements in pathologist performance in 2200 cases. *Telemed J*. 1999;5(4):323-337.
- Goacher E, Randell R, Williams BJ, Treanor D. The diagnostic accuracy of digital microscopy: a systematic review protocol. PROSPERO: International Prospective Register of Systematic Reviews CRD42015017859. http://www.crd.york.ac.uk/PROSPEROFILES/17859_PROTOCOL_20150313.pdf. Published 2015. Accessed January 2, 2016.
- Al Habeeb A, Evans A, Ghazarian D. Virtual microscopy using whole-slide imaging as an enabler for teledermatopathology: a paired consultant validation study. *J Pathol Inform*. 2012;3:2.
- Al-Janabi S, Huisman A, Vink A, et al. Whole slide images for primary diagnostics of gastrointestinal tract pathology: a feasibility study. *Hum Pathol*. 2012;43(5):702-707.

Object/Feature	No. of Cases
Neutrophils	19
Eosinophils	7
Mast cells	1
Nucleated red blood cells	8
Eosinophilic granular bodies	1
Amyloid	2
Blue mucin	1

17. Al-Janabi S, Huisman A, Vink A, et al. Whole slide images for primary diagnostics in dermatopathology: a feasibility study. *J Clin Pathol*. 2012;65(2):152–158.
18. Al-Janabi S, Huisman A, Willems SM, Van Diest PJ. Digital slide images for primary diagnostics in breast pathology: a feasibility study. *Hum Pathol*. 2012;43(12):2318–2325.
19. Al-Janabi S, Huisman A, Nikkels PG, ten Kate FJ, van Diest PJ. Whole slide images for primary diagnostics of paediatric pathology specimens: a feasibility study. *J Clin Pathol*. 2013;66(3):218–223.
20. Al-Janabi S, Huisman A, Jonges GN, ten Kate FJW, Goldschmeding R, van Diest PJ. Whole slide images for primary diagnostics of urinary system pathology: a feasibility study. *J Renal Inj Prev*. 2014;3(4):91–96.
21. Arnold MA, Chenever E, Baker PB et al. The College of American Pathologists guidelines for whole slide imaging validation are feasible for pediatric pathology: a pediatric pathology practice experience. *Pediatr Dev Pathol*. 2015;18(2):109–116.
22. Bauer TW, Schoenfeld L, Slaw RJ, Yerian L, Sun Z, Henricks WH. Validation of whole slide imaging for primary diagnosis in surgical pathology. *Arch Pathol Lab Med*. 2013;137(4):518–524.
23. Bauer TW, Slaw RJ. Validating whole-slide imaging for consultation diagnoses in surgical pathology. *Arch Pathol Lab Med*. 2014;138(11):1459–1465.
24. Buck TP, Dilorio R, Havrilla L, O'Neill DG. Validation of a whole slide imaging system for primary diagnosis in surgical pathology: a community hospital experience. *J Pathol Inform*. 2014;5(1):43.
25. Campbell WS, Lele SM, West WW, Lazenby AJ, Smith LM, Hinrichs SH. Concordance between whole-slide imaging and light microscopy for routine surgical pathology. *Hum Pathol*. 2012;43(10):1739–1744.
26. Campbell WS, Hinrichs SH, Lele SM, et al. Whole slide imaging diagnostic concordance with light microscopy for breast needle biopsies. *Hum Pathol*. 2014;45(8):1713–1721.
27. Fonyad L, Krenacs T, Nagy P, et al. Validation of diagnostic accuracy using digital slides in routine histopathology. *Diagn Pathol*. 2012;(7):35.
28. Gilbertson JR, Ho J, Anthony L, Jukic DM, Yagi Y, Parwani AV. Primary histologic diagnosis using automated whole slide imaging: a validation study. *BMC Clin Pathol*. 2006;6:4.
29. Gui D, Cortina G, Naini B, et al. Diagnosis of dysplasia in upper gastrointestinal tract biopsies through digital microscopy. *J Pathol Inform*. 2012;3:27.
30. Ho J, Parwani AV, Jukic DM, Yagi Y, Anthony L, Gilbertson JR. Use of whole slide imaging in surgical pathology quality assurance: design and pilot validation studies. *Hum Pathol*. 2006;37(3):322–331.
31. Houghton JP, Ervine AJ, Kenny SL, et al. Concordance between digital pathology and light microscopy in general surgical pathology: a pilot study of 100 cases. *J Clin Pathol*. 2014;67(12):1052–1055.
32. Jukic DM, Drogowski LM, Martina J, Parwani AV. Clinical examination and validation of primary diagnosis in anatomic pathology using whole slide digital images. *Arch Pathol Lab Med*. 2011;135(3):372–378.
33. Loughrey MB, Kelly PJ, Houghton OP, et al. Digital slide viewing for primary reporting in gastrointestinal pathology: a validation study. *Virchows Arch*. 2015;467(2):137–144.
34. Ordi J, Castillo P, Saco A, et al. Validation of whole slide imaging in the primary diagnosis of gynaecological pathology in a university hospital. *J Clin Pathol*. 2015;68(1):33–39.
35. Snead DR, Tsang YW, Meskiri A, et al. Validation of digital pathology imaging for primary histopathological diagnosis. *Histopathology*. 2016;68(7):1063–1072.
36. Velez N, Jukic D, Ho J. Evaluation of 2 whole-slide imaging applications in dermatopathology. *Hum Pathol*. 2008;39(9):1341–1349.
37. Wilbur DC, Madi K, Colvin RB, et al. Whole-slide imaging digital pathology as a platform for teleconsultation: a pilot study using paired subspecialist correlations. *Arch Pathol Lab Med*. 2009;133(12):1949–1953.
38. Royal College of Pathologists. Guide to conducting an investigative audit of cellular pathology practice. <https://www.rcpath.org/resourceLibrary/guide-to-investigative-audit.html>. Published March 2014. Accessed March 9, 2016.
39. Royal College of Pathologists. Guide to conducting a duty of care review. <https://www.rcpath.org/resourceLibrary/guide-to-investigative-audit.html>. Published March 2014. Accessed March 9, 2016.