OPEN

Whole Slide Imaging Versus Microscopy for Primary Diagnosis in Surgical Pathology

A Multicenter Blinded Randomized Noninferiority Study of 1992 Cases (Pivotal Study)

Sanjay Mukhopadhyay, MD,* Michael D. Feldman, MD, PhD,† Esther Abels, MSc,‡
Raheela Ashfaq, MD, MS,§ Senda Beltaifa, MD,‖ Nicolas G. Cacciabeve, MD,‖
Helen P. Cathro, MBChB, MPH,¶ Liang Cheng, MD,# Kumarasen Cooper, MBChB, DPhil, FRCPath,†
Glenn E. Dickey, MD,‖ Ryan M. Gill, MD, PhD,** Robert P. Heaton, Jr, MD,‖ René Kerstens, MSc,‡
Guy M. Lindberg, MD,§ Reenu K. Malhotra, MD,§ James W. Mandell, MD, PhD,¶
Ellen D. Manlucu, MD,‖ Anne M. Mills, MD,¶ Stacey E. Mills, MD,¶
Christopher A. Moskaluk, MD, PhD,¶ Mischa Nelis, MSc,‡ Deepa T. Patil, MD,*
Christopher G. Przybycin, MD,* Jordan P. Reynolds, MD,* Brian P. Rubin, MD, PhD,*
Mohammad H. Saboorian, MD,§ Mauricio Salicru, MD,§ Mark A. Samols, MD, PhD,‖
Charles D. Sturgis, MD,* Kevin O. Turner, DO,§ Mark R. Wick, MD,¶ Ji Y. Yoon, MD,§
Po Zhao, MD, PhD,‖ and Clive R. Taylor, MD, PhD††

Abstract: Most prior studies of primary diagnosis in surgical pathology using whole slide imaging (WSI) versus microscopy have focused on specific organ systems or included relatively few cases.

From the *Department of Pathology, Cleveland Clinic, Cleveland, OH; †Department of Pathology, Hospital of the University of Pennsylvania, Philadelphia, PA; ‡Philips Digital Pathology Solutions, Best, The Netherlands; §Miraca Life Sciences, Irving, TX; ‖Advanced Pathology Associates, Silver Spring, MD; ¶Department of Pathology, University of Virginia, Charlottesville, VA; #Department of Pathology, Indiana University School of Medicine, Indianapolis, IA; **Department of Pathology, University of California, San Francisco; and ††Department of Pathology, Keck School of Medicine University of Southern California, Los Angeles, CA.

Correspondence: Clive R. Taylor, MD, PhD, Department of Pathology, Keck School of Medicine University of Southern California, Pathology, HMR 311, Los Angeles, CA 90033 (e-mail: clive.taylor@med.usc.edu).

The objective of this study was to demonstrate that WSI is noninferior to microscopy for primary diagnosis in surgical pathology. A blinded randomized noninferiority study was conducted across the entire range of surgical pathology cases (biopsies and resections, including hematoxylin and eosin, immunohistochemistry, and special stains) from 4 institutions using the original sign-out diagnosis (baseline diagnosis) as the reference standard. Cases were scanned, converted to WSI and randomized. Sixteen pathologists interpreted cases by microscopy or WSI, followed by a wash-out period of ≥ 4 weeks, after which cases were read by the same observers using the other modality. Major discordances were identified by an adjudication panel, and the differences between major discordance rates for both microscopy (against the reference standard) and WSI (against the reference standard) were calculated. A total of 1992 cases were included, resulting in 15,925 reads. The major discordance rate with the reference standard diagnosis was 4.9% for WSI and 4.6% for microscopy. The difference between major discordance rates for microscopy and WSI was 0.4% (95% confidence interval, −0.30% to 1.01%). The difference in major discordance rates for WSI and microscopy was highest in endocrine pathology (1.8%), neoplastic kidney pathology (1.5%), urinary bladder pathology (1.3%), and gynecologic pathology (1.2%). Detailed analysis of these cases revealed no instances where interpretation by WSI was consistently inaccurate compared with microscopy for multiple observers. We conclude that WSI is noninferior to microscopy for primary diagnosis in surgical pathology, including biopsies and resections stained with hematoxylin and eosin, immunohistochemistry and special stains. This conclusion is valid across a wide variety of organ systems and specimen types.

Key Words: whole slide imaging, microscopy, surgical pathology, digital imaging, pathology, primary diagnosis, noninferiority trial

(Am J Surg Pathol 2018;42:39–52)

Whole slide imaging (WSI), also known as digital pathology or virtual pathology, is a technology that involves high-speed, high-resolution digital acquisition of images representing entire stained tissue sections from glass slides in a format that allows them to be viewed by a pathologist on a computer monitor, where the image can be magnified and navigated spatially in much the same way as standard microscopy.[1] These images can be utilized for diagnosis by pathologists, creating a digital workflow that obviates the use of conventional bright field light microscopy (henceforth referred to as "microscopy"). Before substituting the time-honored, familiar and versatile microscope with WSI, several valid concerns need to be addressed. The most critical one is whether pathologic diagnoses rendered using WSI are comparable with (ie, noninferior to) diagnoses made by microscopy.[2]

Prior studies have compared WSI with microscopy in several different settings in diagnostic pathology, including frozen section diagnosis, consultation cases,[3,4] difficult surgical pathology cases,[5] and pathology of specific organ systems or subspecialties.[6–30] College of American Pathologists (CAP) guidelines published in 2013 recommend that each intended use be supported by a separate validation study.[1] One intended use in surgical pathology is primary diagnosis, defined as establishing a final diagnosis solely by review of digital images without the use of microscopy. Only a few validation studies have addressed the use of WSI for primary diagnosis in surgical pathology across a wide variety of specimen types from different organ systems.[31–42] Most of these have included relatively small numbers of cases, and many were not sufficiently powered to definitively demonstrate that WSI is noninferior to microscopy for this purpose. In some studies, issues with study design may have introduced the possibility of bias. Factors creating a risk of bias include patient selection, index test, reference standard, flow and timing, as recently described in detail by Goacher et al.[43] A well-known example is the wash-out period; studies with an inadequate wash-out period between microscopic and WSI reads introduce the possibility of recall bias. Studies that include readers in the adjudication process may introduce an element of bias in determining whether diagnoses made by microscopy and WSI are concordant. With this background in mind, the primary objective of our study was to demonstrate that WSI is noninferior to microscopy for primary diagnosis across the entire range of surgical pathology practice. Frozen section and cytology cases were excluded. Hematopathology cases were also omitted because of common use of higher magnification objectives.

METHODS

Over a 14-month period (July 2015 to September 2016), a blinded randomized noninferiority study comparing microscopy with WSI for primary diagnosis in surgical pathology was conducted at 4 institutions in the United States (2 academic centers and 2 commercial laboratories; the latter included an independent hospital-based pathology practice). Investigators from pathology departments at multiple other institutions were actively involved in planning, study design, execution, and data analysis. The study protocol was approved by Institutional Review Boards (IRBs) at all participating institutions.

Screening and Enrollment

Each participating institution was assigned a set of organ systems from which to enroll cases, for a total of 20 organ systems (Table 1). Only formalin-fixed paraffin-embedded surgical pathology cases were enrolled. Frozen sections and cases received in consultation were excluded. Target enrollment for each organ system and case type was predefined, based on discussions with the United States Food and Drug Administration (FDA), and were intended to reflect routine clinical practice while enriching for more difficult malignant cases. As an example, for colorectal cases, the enrollment target was 150 cases, including 50 benign/inflammatory biopsies, 50 biopsies of adenomas, 40 endoscopic biopsies of adenocarcinoma, and 10 adenocarcinoma resections. Cases were excluded if they met any of the following exclusion criteria: (1) slides for a case were not available at the site, (2) control slides for immunohistochemistry or special stains were not available, (3) slides selected did not match any subtype of the organ for which the case was selected, (4) clinical information available to the sign-out pathologist in the pathology requisition form could not be obtained, (5) selected slides contained indelible markings, (6) more than one case was selected for a patient, (7) the case consisted of frozen section slides only, or (8) the case consisted of gross specimens only. The most common reason for not including a screened case was that the target enrollment number for that specific diagnosis was met. For example, once the target of 120 consecutive benign core biopsies of prostate was reached, subsequent benign core biopsies were not enrolled. By this process, 12,338 cases were screened by 8 enrollment pathologists from 4 centers until the enrollment target of 2000 cases (3405 slides) was reached. These cases were submitted for scanning and subsequent review.

The inclusion criteria specified that the interval between accession of cases and selection into the study was to be at least 1 year. Cases were reviewed for enrollment in the study by 2 "enrollment pathologists" per institution. One of these individuals reviewed a list of consecutive pathology reports from organ systems assigned to that center and flagged cases for retrieval of glass slides. All glass slides for each case were reviewed (screened) by the enrollment pathologist. For biopsies, the enrollment pathologist selected key slides required for diagnosis, including hematoxylin and eosin and immunohistochemical stains. In addition, for resections, representative slides required for diagnosis and staging were selected, including negative or positive lymph nodes and margins. The second enrollment pathologist (validating enrollment pathologist) then reviewed all slides

TABLE 1. Cases Included in the Study (Full Analysis Set) by Organ System

Study Site	Organ System	No. Cases	Organ ($\pm$ Specimen Type)
1	Colorectal	150	Adenocarcinoma resection: 10
			Benign/inflammatory biopsy: 50
			Adenoma biopsy: 50
			Adenocarcinoma endoscopic biopsy: 40
	Urinary bladder	99	Carcinoma resection: 9
			Benign/inflammatory/non-neoplastic biopsy: 25
			Dysplasia biopsy: 25
			Noninvasive carcinoma (TUR or biopsy): 25
			Carcinoma (TUR or biopsy): 15
	Gynecologic	150	Endometrial biopsy/curetting: 50
			Cervical biopsy/curetting: 25
			Cervical cone/LEEP: 25
			Ovary benign neoplastic/non-neoplastic: 20
			Ovary malignant neoplastic: 30
	Liver/bile duct, neoplastic	49	Core biopsy: 39
			Wedge biopsy or resection: 10
	Brain	60	Non-neoplastic: 10
			Neoplastic biopsy: 25
			Neoplastic resection: 25
	Total (site 1)	508	
2	Prostate	299	Benign core biopsy: 120
			Benign resection: 29
			Adenocarcinoma biopsy: 120
			Adenocarcinoma resection: 30
	Lymph node	100	For presence/absence of metastasis: 75
			Non-neoplastic: 25
	Endocrine	100	Pancreas: 50
			Thyroid: 30
			Parathyroid: 10
			Adrenal: 10
	Kidney, neoplastic	50	Consecutive cases, all comers: 50
	Salivary	50	Consecutive cases, all comers: 50
	Total (site 2)	599	
3	Gastroesophageal junction	115	Rule out Barrett/dysplasia: 50
			Non-neoplastic/inflammatory biopsy: 65
	Skin	177	Non-neoplastic/inflammatory biopsy: 50
			Squamous/basal cell neoplasm, biopsy: 52
			Melanocytic lesion, biopsy: 75
	Hernial/peritoneal	7	Consecutive cases, all comers: 7
	Gallbladder	10	Consecutive cases, all comers: 10
	Appendix	10	Consecutive cases, all comers: 10
	Soft tissue tumors	21	Consecutive cases, all comers: 21
	Anus/perianal	50	Consecutive cases, all comers: 50
	Total (site 3)	390	
4	Breast	299	Benign/atypical core needle biopsy: 50
			Benign/atypical lumpectomy: 50
			In situ carcinoma core needle biopsy: 49
			In situ carcinoma lumpectomy: 50
			Invasive carcinoma core needle biopsy: 50
			Invasive carcinoma lumpectomy: 50
	Lung/bronchus/larynx/oral cavity/nasopharynx	95	Benign/inflammatory biopsy: 25
			Dysplasia biopsy: 22
			Carcinoma biopsy: 28
			Carcinoma resection: 20
	Stomach	99	Inflammatory including *Helicobacter pylori*: 50
			Polyps/neoplastic biopsy: 39
			Polyps/neoplastic resection: 10
	Total (site 4)	495	
	Total for all 4 sites (full analysis set)	1992	

TUR indicates transurethral resection.

selected by the first enrollment pathologist to ensure that diagnostic material reflecting the original diagnosis was present. The original diagnosis made in the course of routine patient care by the pathologist signing out the case (baseline diagnosis) was considered the reference standard.

Slide Scanning

A study coordinator compiled all cases selected by the enrollment pathologists and submitted them for digital scanning at participating sites using the Philips IntelliSite Pathology Solution (Philips, the Netherlands), which includes a scanner, an image management system and a display. A study technician was trained to scan slides using appropriate calibration and quality control measures. All slides were scanned as WSI for digital review using the Philips IntelliSite Pathology Solution.

Of the 2000 cases submitted for scanning, 8 (0.4%) were excluded for the following reasons: slide size did not meet scanner specifications (4 cases), no tissue was detected by scanner on any one of the slides selected for the case (2 cases), more than one case was selected for the patient (1 case), or slides were broken or damaged (1 case). This process yielded a "full analysis set" of 1992 cases (99.6% of enrolled cases).

Randomization

Original glass slides from all cases included in the study (full analysis set) were randomized and deidentified. Randomization was performed within an Electronic Data Capture (EDC) system provided by the manufacturer (eCaseLink Document Solutions Group, Malvern, PA). Original surgical pathology numbers were obscured and replaced by a study identifier (barcode label) by the study coordinator. Cases were then placed in random order and divided into batches of 20 cases, each of which contained a random mix of cases from various organ systems.

Interpretation of Microscopy and Whole Slide Images by Reading Pathologists

Randomized and deidentified slides from each case were presented for interpretation to 16 board-certified "reading pathologists" (4 at each center) different from the 8 enrollment pathologists whose role was described previously. Each reading pathologist followed standard training including self-familiarization with the WSI viewer. In order to represent the breadth of potential users of WSI, reading pathologists were selected to represent a variety of expertise, practice types (academic vs. nonacademic, generalists vs. subspecialists), subspecialty training and years of experience. Reading pathologists interpreted cases enrolled from their center only, blinded to the reference standard diagnosis.

All cases were interpreted by 2 modalities. The first (microscopy) involved viewing glass slides using a microscope, identical to the practice of routine surgical pathology. Each pathologist viewed glass slides in their office using their own microscope. The second method (WSI) involved viewing scanned digital images on a high-resolution monitor without the use of a microscope. Reading pathologists interpreted cases in batches of 20. After a batch of 20 cases was reviewed, the same pathologist was given a separate batch of 20 cases for review by the other modality. For example, a pathologist who interpreted cases 1 to 20 using microscopy might be assigned cases 71 to 90 for review using WSI, followed by cases 41 to 60

using a microscope, and so on. This process was repeated for all 16 reading pathologists until all assigned cases were viewed by each pathologist. After a wash-out period of at least 4 weeks, all cases were arranged in random order and interpreted a second time by the same reading pathologists using the other modality (ie, cases initially interpreted by microscopy were interpreted by WSI and vice versa). The wash-out period differed from case to case depending on its order in the randomly arranged cases. The mean wash-out period per pathologist ranged from 38.7 to 81.8 days. The minimum wash-out period was 27 days and the maximum was 143 days.

At least 2 workstations, each with a 27-inch monitor, were provided to each participating site and located in a room simulating a clinical practice environment. The diagnosis for each case was entered electronically into the EDC electronic database by each reading pathologist. Staging parameters on cases requiring staging were entered on paper using templates that incorporated key elements of CAP synoptic templates for each organ. The time that a pathologist either opened or closed a case in the EDC system was logged. Reading pathologists were allowed to freely consult textbooks and other literature online, whether using microscopy or WSI. Identical clinical information was provided to reading pathologists for both modalities. Information regarding prior diagnoses on the same patient was not provided. Reading pathologists were not allowed to request recuts or any additional special stains beyond those already provided, or to consult with other pathologists. The randomization process ensured that the order in which cases were presented to the reading pathologist for microscopic interpretation was different than the order for WSI interpretation.

Each diagnosis by a reading pathologist on a case (whether by WSI or microscopy) was termed a "read." As each case from any participating institution was interpreted twice by 4 reading pathologists, there were 8 "reads" per case, not including the original sign-out diagnosis.

Adjudication Phase

The diagnosis rendered by the original pathologist who signed out the case in the course of routine patient care using a microscope was considered the reference standard. A central panel of 3 "adjudication pathologists" independently determined the level of concordance between microscopic and WSI diagnoses and the reference standard. The adjudication panel did not include any of the enrollment pathologists or reading pathologists, and was selected from institutions different than the 4 centers that participated in enrollment and reading. Each adjudication pathologist had at least 10 years of relevant experience.

Two adjudication pathologists were provided a list of paired diagnoses, blinded to method of diagnosis (microscopy or WSI), reading pathologist and participating site/institution. Adjudication pathologists did not view glass slides for any case. Using an Adjudication Charter for each organ system, adjudication pathologists placed
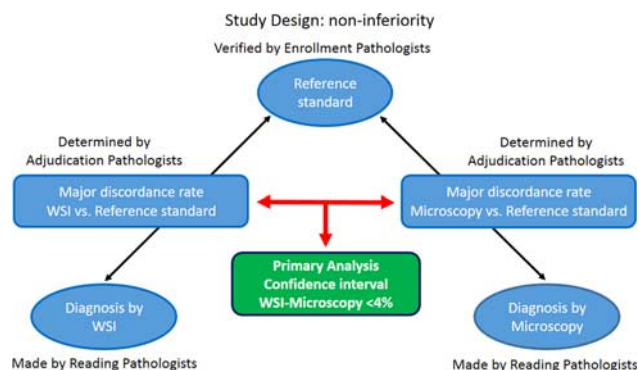
FIGURE 1. Study design.

each pair of diagnoses into one of 3 categories: concordant, minor discordant or major discordant. In keeping with widely accepted definitions, a major discordance was defined as a difference in diagnosis that would be associated with a difference in patient management.[32,44] In case of a disagreement between the 2 adjudication pathologists on the level of concordance between 2 diagnoses, the third adjudication pathologist served as a tie-breaker. The primary endpoint of the study was the difference between major discordance rates for microscopy and WSI by comparison with the reference standard. The study design is summarized in Figure 1.

RESULTS

A total of 1992 cases (3390 slides/3390 images) were included in the full analysis set, of which 923 slides (27%) were either immunohistochemical stains or special stains. The range of slides examined was 1 to 16 slides per case. Ten cases had 10 or more slides per case. Scanning performance is shown in Table 2. In the first scan of these 3390 slides, the Philips IntelliSite Pathology Solution was able to automatically detect an issue, such as no tissue or label detection, for 77 slides (2.3%). The images from 70 slides (2.1%) did not pass the image quality check by the scanning operator for slide-related issues such as prior ink markings, broken slides or debris on the slide. For 55 images (1.6%) the scanning technician identified an out of focus image (54 images, 1.6%) or missing tissue (1 image, 0.03%).

In the second scan (in cases where this was required), the Philips IntelliSite Pathology Solution was able to automatically detect an issue for 21 slides (0.6%). The images

TABLE 3. Major Discordance Rates by Organ System: Microscopy Versus Reference Standard and WSI Versus Reference Standard

Major Discordance Rate	Between Microscopy and Reference Standard (%)	Between WSI and Reference Standard (%)
<1%	Peritoneal (0)	Peritoneal (0)
	Gallbladder (0)	Gallbladder (0)
	Appendix (0)	Appendix (0)
	Soft tissue (0)	Soft tissue (0)
	Stomach (0.5)	Lymph node (0.3)
	Lymph node (0.8)	Stomach (0.8)
1%-4.9%	Colorectal (1)	Peri(anal) (1)
	Kidney neoplastic (1)	Colorectal (1.7)
	Gastroesophageal junction (1.3)	Salivary gland (2)
	Peri(anal) (2)	Gastroesophageal junction (2)
	Salivary gland (3)	Kidney neoplastic (2.5)
	Respiratory (4.2)	Respiratory (3.5)
	Breast (4.3)	Breast (4.2)
	Skin (4.7)	Liver/bile duct (4.6)
	Endocrine (4.7)	Skin (4.9)
≥5%	Gynecologic (5.2)	Brain (6.2)
	Liver/bile duct (5.6)	Gynecologic (6.3)
	Brain (5.8)	Endocrine (6.5)
	Bladder (6.1)	Bladder (7.3)
	Prostate (11.3)	Prostate (12)

from 7 slides (0.2%) did not pass the image quality check by the scanning operator for slide-related issues. For 22 images (0.6%), the scanning technician identified the image to be out of focus for 21 images and found "venetian blinds" at high magnification for 1 image. For clinical study operational reasons, slides were rescanned a maximum of 5 times before they were enrolled into the study.

Reading times were derived from available system data. Reading time was defined as the time it took a pathologist to open the case, read all available information, diagnose the case and enter the diagnosis in the system. Approximately 94% of reads were completed within 30 minutes of reading time. Reading times longer than 30 minutes were considered to not reflect the actual reading time since such instances generally occurred due to external factors; for example, the reader opened the case, was interrupted during the read, and forgot to close the case, resulting in an incorrect log. For exploratory analyses, and assuming that this would have the same effect on microscopic reads as on WSI reads, it was decided to include only times shorter than 30 minutes for the analysis. The mean reading time for microscopy was 78 seconds and the mean reading time for WSI was 84 seconds. The mean difference between the reading time for WSI and the reading time for microscopy was 6 seconds with a 95% confidence interval (CI) of (0.03-0.12). One site in the study performed a detailed analysis of reading times, which is being published in a separate manuscript.[45]

The number of cases by organ system, diagnosis and specimen type is shown in Table 1. For 1992 cases, a total of 15,936 reads (1992×8) was expected. However, 11 reads (7 by microscopy, 4 by WSI) were excluded as reading

TABLE 2. Scanning Performance

Scanning Attempt	Slides Scanned	WSI Passed	Reasons for Not Passing (% of Full Analysis Set)		
			Detection by PIPS	Slide-related	Image-related
1	3390	3179	2.3	2.1	1.6
2	202	152	0.6	0.2	0.6

PIPS indicates Philips IntelliSite Pathology Solution.

TABLE 4. Comparison of Major Discordance Rate for WSI and Major Discordance Rate for Microscopy by Organ System
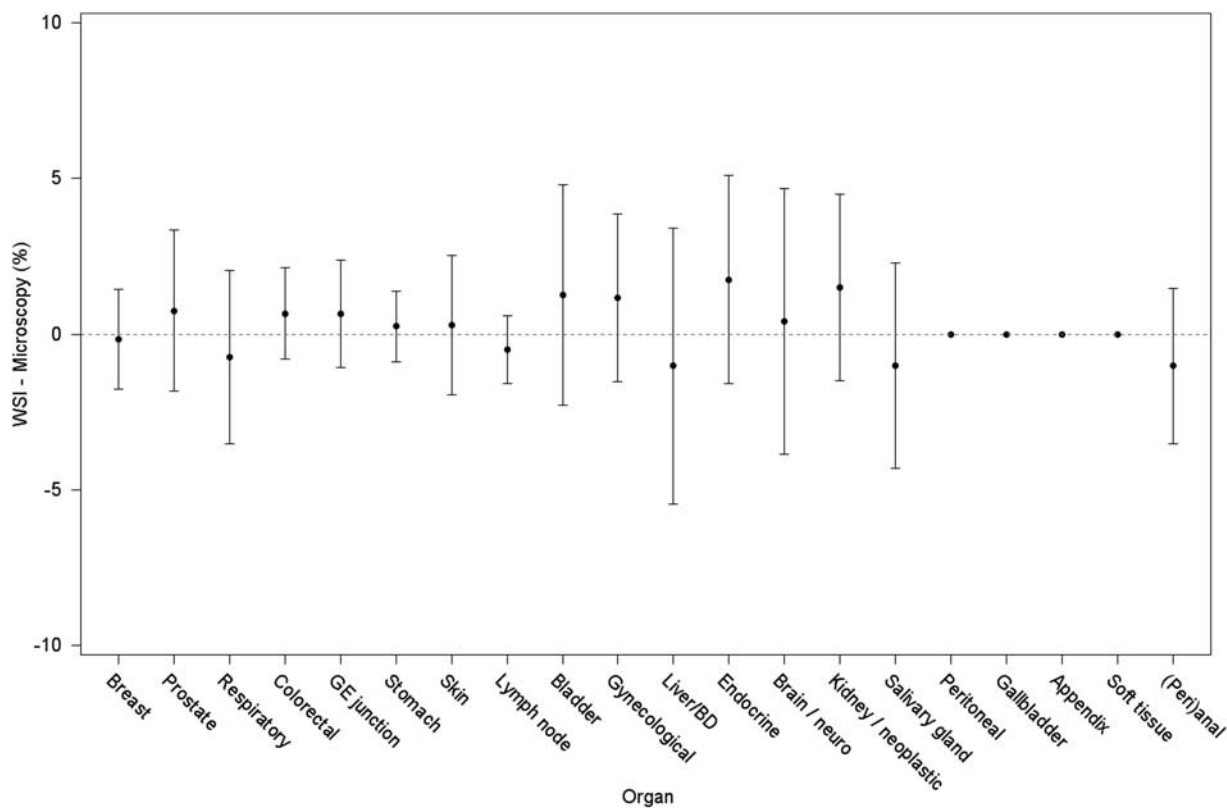
Difference in Major Discordance Rate	Major Discordance Rate for WSI Higher than Major Discordance Rate for Microscopy (Difference, in %)	Major Discordance Rate for WSI Lower than Major Discordance Rate for Microscopy (Difference, in %)	Major Discordance Rate for WSI Identical to Microscopy (Difference, in %)
<1%	Stomach (0.3) Skin (0.3) Brain (0.4) Colorectal (0.7) GE junction (0.7) Prostate (0.8)	Breast (0.2) Respiratory (0.5) Lymph nodes (0.7)	Peritoneal (0) Gallbladder (0) Appendix (0) Soft tissue (0)
≥1%	Gynecologic (1.2) Bladder (1.3) Kidney neoplastic (1.5) Endocrine (1.8)	Liver/bile duct (1) Salivary gland (1) Peri(anal) (1)	

Major discordance rate for WSI—rate of major discordance between WSI and reference standard.
Major discordance rate for microscopy—rate of major discordance between microscopy and reference standard.
GE indicates gastroesophageal.

pathologists selected "no diagnosis" for a variety of reasons, yielding a total of 15,925 reads (7961 by microscopy, 7964 by WSI).

The major discordance rate between microscopy and the reference standard was 4.6% (364/7961 reads) and the major discordance rate between WSI and the reference standard was 4.9% (393/7964 reads). The difference in major discordance rates for WSI and microscopy was 0.4%, with a derived 2-sided 95% CI of (−0.30% to 1.01%). As the upper limit of this CI was less than the

FIGURE 2. Difference in major discordance rates between WSI (WSI vs. reference standard) and microscopy (microscopy vs. reference standard) for each of 20 organ systems included in the study. The *y*-axis is the difference (in %) between the major discordance rate for WSI and the major discordance rate for microscopy. The *x*-axis lists the organ systems. For any given organ system, a WSI minus microscopy difference > 0 indicates that the overall major discordance rate for that organ system was greater for WSI than for microscopy. A negative percentage indicates that the major discordance rate for that organ system was greater for microscopy than for WSI. BD indicates bile duct; GE, gastroesophageal.

TABLE 5. Analysis of Organ Systems in Which Difference Between Major Discordance Rate for WSI and Major Discordance Rate for Microscopy was ≥1%

Type of Specimen	WSI: No Major Discordance* Microscopy: no Major Discordance (No. Reads)	WSI: Major Discordance Microscopy: no Major Discordance (No. Reads)	Microscopy: Major Discordance WSI: no Major Discordance (No. Reads)	WSI Major Discordance: Microscopy: Major Discordance (No. Reads)	Total Reads (Cases)†
Endocrine pathology	372	9	2	17	400 (100)
Neoplastic kidney pathology	194	4	1	1	200 (50)
Urinary bladder pathology	352	20	15	9	396 (99)
Gynecologic pathology	550	19	12	19	600 (150)

*Discordance refers to discordance with reference standard, that is, original sign-out diagnosis.
†Each case was read by 4 pathologists (reads = cases×4).

prespecified noninferiority threshold of 4%,[32] WSI was considered noninferior to microscopy, meeting the primary objective of the study.

Major Discordance Rates by Organ System: Microscopy Versus Reference Standard and WSI Versus Reference Standard

For each organ system, major discordance rates between microscopy and the reference standard, and between WSI and the reference standard are listed in Table 3. For cases from the peritoneum, gallbladder, appendix and soft tissue, there were no major discordances between either microscopy or WSI and the reference standard. For stomach and lymph node cases, discordance rates were very low (<1%) with both modalities. For most other organs systems/tissues, discordance rates between both modalities and the reference standard ranged from 1% to 4.9%. Major discordance rates between microscopy and the reference standard were highest (≥5%) for pathology of the brain, gynecologic tract, liver/bile ducts, urinary bladder, and prostate. These were very similar to the levels of discordance between WSI and the reference standard, with the exception of liver/bile duct cases, where major discordance rates for WSI were lower than microscopy. Of all organs/organ systems included in the study, prostate showed the highest major discordance rates, which were seen with microscopy (11.3%) as well as WSI (12%).

Overall, in 157/7596 reads (2%), there was a major discordance between WSI and the reference standard in cases where microscopy was concordant with the reference standard. In 127/7566 reads (1.6%), there was a major discordance between microscopy and the reference standard in cases where WSI was concordant with the reference standard.

Differences between major discordance rates for microscopy and major discordance rates for WSI by organ system are shown in Table 4 and depicted in Figure 2. For 4 organ systems, there was no difference between major discordance rates for the 2 modalities (peritoneum, gallbladder, appendix, soft tissue). WSI major discordances were slightly higher

(<1%) in stomach, skin, brain, colorectum, gastroesophageal junction, and prostate. The major discordance rate for WSI was ≥1% higher than the major discordance rate for microscopy in endocrine, neoplastic kidney, gynecologic, and urinary bladder pathology. These 4 organs/organ systems were selected for detailed analysis (see below). WSI major discordance rates were ≥1% lower than the major discordance rate for microscopy in liver/bile duct, salivary gland, and (peri)anal pathology. These organs/organ systems, where microscopy performed worse than WSI, were not subjected to additional analysis.

Endocrine Pathology: Detailed Analysis

This analysis was based on paired reads, that is one read by microscopy and one read by WSI for the same case by the same pathologist. Since each case was read twice by 4 pathologists, there were 4 paired reads per case. Of 400 paired reads on 100 cases in endocrine pathology, there were 9 reads in which WSI was judged to show a major discordance with the reference standard while the corresponding microscopic read was not (Table 5). Details of the diagnosis in these cases are provided in Table 6. Most of these occurrences (7) involved thyroid pathology. Six were caused by under-diagnosis and one by over-diagnosis of papillary thyroid carcinoma using WSI. There was only one occurrence each in adrenal pathology and pancreatic pathology. There were no cases in which 3 pathologists or all 4 pathologists made a major discordant diagnosis compared with the reference standard by WSI but a concordant (or minor discordant) diagnosis by microscopy. There was only 1 case in which 2/4 readers made a major discordant diagnosis by WSI but a concordant diagnosis by microscopy (case 0849, Table 6). The remaining occurrences were random errors in which a single pathologist (1/4) made a major discordant diagnosis by WSI but a concordant (or minor discordant) diagnosis by microscopy; in each of these instances, the remaining 3 pathologists made the same diagnosis by WSI and microscopy.

Neoplastic Kidney Pathology: Detailed Analysis

Of 200 paired reads from 50 cases in neoplastic kidney pathology, only 4 featured a major discordance

TABLE 6. Endocrine Pathology: Paired Reads in Which WSI Yielded a Major Discordance but Microscopy Showed no Major Discordance

Case Number	Reader*	Reference Standard†	Paired Reads of Same Case by Same Reader	
			Diagnosis by WSI (Major Discordance)	Diagnosis by Microscopy (no Major Discordance)‡
0849	E	Papillary thyroid carcinoma	Nodular hyperplasia with a dominant nodule	Papillary thyroid carcinoma
0849	H	Papillary thyroid carcinoma	Hyperplastic nodule	Papillary carcinoma of thyroid, follicular variant. Area of carcinoma surrounding peripheral rim of hyperplastic nodule
0856	H	Papillary thyroid carcinoma, follicular variant	Hyperplastic nodules with oncocytic metaplasia	Follicular carcinoma, oncocytic type (minor discordance)
0524	E	Papillary thyroid carcinoma, follicular variant	Multinodular hyperplasia	Papillary thyroid carcinoma
0521	E	Papillary thyroid carcinoma	Nodular hyperplasia	Papillary thyroid carcinoma
0514	E	Papillary thyroid carcinoma, classic type	Nodular hyperplasia and changes consistent with Graves disease	Follicular adenoma, would show to head and neck pathologist to exclude cribriform-morular variant of papillary thyroid carcinoma (minor discordance)
0512	H	Dominant hyperplastic nodule in a background of diffuse follicular hyperplasia with nodular changes consistent with history of Graves' disease. Chronic lymphocytic thyroiditis.	Papillary thyroid carcinoma, follicular variant	Hyperplastic thyroid nodule in a background of chronic thyroiditis
0663	F	Pheochromocytoma	Adrenocortical adenoma	Pheochromocytoma
0973	E	Well-differentiated neuroendocrine tumor of the pancreas, intermediate grade	Malignant diagnosis (yes), low grade	Malignant diagnosis (yes), intermediate grade

*The 4 reading pathologists at this site were designated E, F, G, and H.
†Original sign-out diagnosis.
‡Concordance or minor discordance compared with reference standard.

between WSI and the reference standard when microscopy was concordant (or minor discordant) with the reference standard (Table 7). There were no cases in which 3/4 or 4/4 readers made a discordant diagnosis compared with the reference standard by WSI but a concordant/minor discordant diagnosis by microscopy. There was only 1 case in which 2/4 readers made a discordant diagnosis by WSI but a concordant diagnosis by microscopy (case 1095, Table 7). The 2 other occurrences were random errors involving only a single pathologist.

Urinary Bladder Pathology: Detailed Analysis

There were 396 paired reads from 99 cases involving pathology of the urinary bladder, of which 20 featured a major discordance between WSI and the reference standard in the face of no major discrepancy between microscopy and the reference standard (Table 8). These involved interpretation of benign bladder biopsies in 5, carcinomas in biopsies or transurethral resections in 3, noninvasive carcinomas in biopsies or transurethral resections in 4, and carcinoma in a resected specimen in 1. There were no consistent problem areas where WSI caused diagnostic difficulties for all 4 readers (or even 3/4 readers). There was only 1 case in which the WSI diagnosis of 2 (of 4) readers was judged as a major discordance when the corresponding microscopic diagnosis was concordant or minor discordant (case 0276, Table 8).

Gynecologic Pathology: Detailed Analysis

Of 600 paired reads from 150 cases in gynecologic pathology, 19 paired reads involved a major discordance between WSI and the reference standard when microscopy was concordant or showed only a minor discordance (Table 9). Most involved endometrial biopsies (8), malignant diagnoses in the ovary (6), and cone biopsies or loop electrosurgical excision procedure excisions of the cervix (4). There were 3 cases in which 3 (of 4) pathologists made a major discordant diagnosis compared with the reference standard by WSI but a concordant or minor discordant diagnosis by microscopy (Table 9, cases 0062, 0361, 0418). In case 0062, which featured an ovarian tumor, 3 pathologists diagnosed carcinoma by microscopy while making a benign or less aggressive diagnosis on WSI. In case 0361 (endometrial biopsy), 3 pathologists made a more aggressive diagnosis on WSI and a benign diagnosis by microscopy. In case 0418, grading of dysplasia was more aggressive on microscopy than on WSI. As in the other organ systems where a detailed case-by-case analysis was performed, there were no consistent problem areas.

Overall, in the entire study set (1992 cases), there were only 3 cases (all in gynecologic pathology, discussed in the prior paragraph) where 3 of 4 pathologists made a major discordant diagnosis by WSI while making a concordant (or minor discordant) diagnosis by microscopy. There was not a single case in the study in which all 4 pathologists made a major discordant diagnosis by WSI

TABLE 7. Neoplastic Kidney Pathology: Paired Reads in Which WSI Yielded a Major Discordance but Microscopy Showed no Major Discordance

Case Number	Reader*	Reference Standard†	Paired Reads of Same Case by Same Reader	
			Diagnosis by WSI (Major Discordance)	**Diagnosis by Microscopy (no Major Discordance)‡**
1090	H	Renal cell carcinoma, papillary type	Metanephric adenoma. Would show to expert	Papillary carcinoma. Would do immunohistochemistry and show to expert
1095	F	Papillary urothelial carcinoma, high grade, with minute focus of early superficial invasion	High-grade papillary urothelial carcinoma, noninvasive	High-grade papillary urothelial carcinoma, noninvasive
1095	H	Papillary urothelial carcinoma, high grade, with minute focus of early superficial invasion	Urothelial carcinoma, low grade	Papillary urothelial carcinoma, high grade (minor discordance)
1102	H	Cystic nephroma	Multilocular cystic renal cell carcinoma	Multilocular cystic renal cell carcinoma (minor discordance)

*The 4 reading pathologists at this site were designated E, F, G, and H.
†Original sign-out diagnosis.
‡Concordance or minor discordance compared with reference standard.

while making a concordant (or minor discordant) diagnosis by microscopy.

DISCUSSION

The major question that validation studies of WSI seek to answer is whether a pathologist will make the same diagnosis on the same case using WSI as they would by microscopy. For this purpose, a WSI diagnosis that is "correct" is as satisfactory as a WSI diagnosis that is "incorrect," as long as the same diagnosis is made by microscopy. Reflecting this principle, the 2013 CAP guidelines state that "validation studies should establish diagnostic concordance between digital and glass slides for the same observer."[1] In keeping with these guidelines, our study was designed primarily to measure variability between the same pathologist(s) for the same case using 2 different modalities. To the best of our knowledge, this is the largest validation study performed in the United States comparing WSI and microscopy for primary diagnosis in surgical pathology. It is also the largest series worldwide in terms of number of reads, and the second-largest series worldwide in terms of cases.

In this study, several measures aimed at accurately assessing intraobserver variability and mitigating the risk of bias, including selection bias and recall bias.[43] These measures included selection of consecutive cases, inclusion of a validation pathologist to validate cases selected by the enrollment pathologist, randomization of reading order, division of cases evenly into batches, randomization of cases between reads, alternation of reading modalities by batch (ie, a batch of microscopy cases was followed by a batch of WSI cases on a different day), blinding of reading pathologists to the reference standard diagnosis, and adjudication of concordance by pathologists different from reading pathologists. Many of these measures were either not considered in prior studies or were not specified in published protocols. Table 10 lists the largest studies that have compared microscopy and WSI for primary diagnosis in surgical pathology using cases

from a variety of organ systems, with adequate reporting of major discrepancy rates.[31,32,39] A major difference between these studies and the current study is in the number of times a study case was interpreted (read) specifically for the study after the original sign-out. In 2 prior studies, each reading pathologist interpreted each case only once during the study (either by WSI or microscopy in Bauer et al[32]; by WSI only in Snead et al[31]), which was compared with the original sign-out diagnosis. In contrast, in the current study, each reading pathologist interpreted each case twice during the study after the original sign-out diagnosis. Hence, although the study by Snead and colleagues included a larger number of cases (3017 vs. 1992), the total number of reads performed during their study (excluding the original sign-out diagnosis) was lower (3017 vs. 15925). The study by Snead and colleagues was most similar to the current study in terms of scope and size, but the design of the 2 studies differed in the stringency of measures taken to reduce bias. For example, adjudication pathologists were different from the reading pathologists in the current study and were selected from institutions different from the reading pathologists, whereas reading pathologists (participating pathologists) were included in the adjudication panel (steering group) by Snead and colleagues. In both studies, however, the difference in major discrepancy rates for WSI and microscopy was reassuringly low (0.7% vs. 0.4%), supporting the contention that these methodologies are essentially equivalent for rendering a primary diagnosis in surgical pathology.

We were also able to report rates of interobserver variability (rate of major discordance between WSI and reference standard, or between microscopy and reference standard). In surgical pathology, interobserver variability is greatest in diagnostically challenging cases, and mainly serves to highlight known problem areas where agreement between observers is suboptimal, even among experts.[46–48] These problems are compounded when general surgical pathologists interpret cases that are difficult even for subspecialists, and when subspecialists interpret cases that they do not sign out in their highly subspecialized

TABLE 8. Urinary Bladder Pathology: Paired Reads in Which WSI Yielded a Major Discordance but Microscopy Showed no Major Discordance

			Paired Reads of Same Case by Same Reader	
Case Number	Reader*	Reference Standard†	Diagnosis by WSI (Major Discordance)	Diagnosis by Microscopy (no Major Discordance)‡
0202	A	No pathologic abnormality	Focal flat high-grade urothelial carcinoma in situ	Mild chronic cystitis with benign reparative epithelial atypia
0208	A	Papillary urothelial carcinoma, high grade, noninvasive	Papillary high-grade urothelial carcinoma invading at least into lamina propria	High-grade papillary urothelial carcinoma in situ
0210	D	High-grade papillary urothelial carcinoma. There is focal invasion of lamina propria	Noninvasive papillary urothelial carcinoma, low-grade	Papillary urothelial carcinoma, high-grade, without invasion (minor discordance)
0211	D	Papillary urothelial carcinoma, low-grade, noninvasive	Invasive papillary urothelial carcinoma, low-grade	Noninvasive papillary urothelial carcinoma, low-grade
0221	A	Bladder mucosa with chronic inflammation and fibrosis	Flat high-grade urothelial dysplasia	Chronic cystitis with urothelial hyperplasia
0227	A	Suspicious for keratinizing squamous cell carcinoma	Fibrosis; no evidence of malignancy	High-grade urothelial carcinoma (minor discordance)
0227	D	Suspicious for keratinizing squamous cell carcinoma	Minute fragment with mildly atypical squamous epithelium; favor squamous metaplasia	Minute fragment of tissue highly suspicious for invasive squamous cell carcinoma
0229	A	High-grade myoinvasive urothelial carcinoma	No histopathologic abnormality; no evidence of malignancy	High-grade urothelial carcinoma invading through bladder wall into perivesical soft tissue
0231	A	Acute and chronic cystitis with focal mucosal ulceration	Myoinvasive high-grade urothelial carcinoma	Mucosal ulceration and chronic cystitis with benign reparative epithelial atypia
0243	A	Mucosal ulceration with marked acute and chronic inflammation extending into muscularis propria	Myoinvasive high-grade urothelial carcinoma	Erosive cystitis with granulomatous features and benign reparative epithelial atypia
0245	D	Urothelial carcinoma in situ	Marked chronic cystitis	Carcinoma in situ
0246	A	Noninvasive high-grade urothelial carcinoma	Papillary high-grade urothelial carcinoma with invasion into lamina propria	Papillary high-grade urothelial carcinoma in situ
0250	A	Chronic cystitis with focal mucosa erosion	Flat high-grade urothelial carcinoma in situ	Erosive chronic cystitis with benign reparative epithelial atypia
0265	A	Noninvasive high-grade papillary urothelial carcinoma	High-grade urothelial carcinoma with invasion of lamina propria	Flat high-grade urothelial carcinoma in situ
0276	B	Focal high-grade urothelial dysplasia	No pathologic abnormality	No pathologic abnormality (minor discordance)
0276	D	Focal high-grade urothelial dysplasia	No pathologic abnormality	Atypical urothelium (minor discordance)
0280	B	Flat high-grade urothelial carcinoma in situ	Chronic inflammation	Urothelial carcinoma in situ
0281	B	Urothelial carcinoma in situ	Reactive changes and chronic inflammation	Urothelial carcinoma in situ
0501	B	Urothelial carcinoma in situ	Cystitis	Chronic cystitis (minor discordant)
0507	B	Urothelial carcinoma in situ	Chronic inflammation	Urothelial carcinoma in situ

*The 4 reading pathologists at this site were designated A, B, C, and D.
†Original sign-out diagnosis.
‡Concordance or minor discordance compared with reference standard.

practices, as for some pathologists in this study. It is important to emphasize that reading pathologists in this study were not permitted to use standard procedures that would be available in "real-life" settings (and were possibly available to the pathologist who originally signed out the case, creating the reference standard diagnostic benchmark), such as obtaining recuts or deeper levels, comparing the case with prior specimens, ordering additional special stains, showing difficult cases to colleagues or obtaining extradepartmental consultation.

Given the effort expended in recent years to validate WSI, its many potential benefits are worth reemphasizing.[43] WSI is already being used clinically in some centers for providing consultations on difficult cases to pathologists at remote locations, providing frozen section interpretations at distant sites, conducting slide conferences and tumor boards

with participants at off-site hospitals, performing proficiency testing/quality assurance, decreasing problems associated with retrieval of glass slides from physical storage sites for comparison to current cases, eliminating problems with loss of staining quality over time or loose cover slips, and using scanned images for semiquantitative image analysis (eg, HER2/neu, estrogen receptors, Ki-67). In the realm of education, the ability of WSI to be "in many places at once" obviates the need to physically transport glass slides, allows for greater flexibility in interacting locally with medical students, residents, fellows, and faculty, and facilitates educational uses such as multicenter conferences, teaching conferences at remote sites, and global pathology education.[49–52] Virtual atlases containing hundreds of educational digital images can be viewed or annotated any time and from anywhere. Links to WSI can be provided within journal articles,

TABLE 9. Gynecologic Pathology: Paired Reads in Which WSI Yielded a Major Discordance but Microscopy Showed no Major Discordance

Read Number/ Case Number	Reader*	Reference Standard†	Paired Reads of Same Case by Same Reader	
			Diagnosis by WSI (Major Discordance)	Diagnosis by Microscopy (No Major Discordance)‡
0048	A	Borderline mucinous tumor	Cystic corpus albicans	Borderline mucinous tumor
0049	B	Metastatic breast carcinoma	Granulosa tumor	Metastatic carcinoma consistent with breast primary
0052	A	High-grade serous papillary carcinoma with undifferentiated features	Acute inflammation and necrotic debris; no evidence of malignancy	Malignant mixed Müllerian tumor (sarcomatoid carcinoma)
0062	A	Endometrioid adenocarcinoma, grade 1	Multilocular serous cystadenomas (bilateral) and borderline endometrioid tumor (unilateral)	Mixed clear-cell adenocarcinoma/ endometrioid adenocarcinoma arising from multicystic serous cystadenofibroma
0062	C	Endometrioid adenocarcinoma, grade 1	Serous cystadenoma and borderline serous (vs. endometrioid) tumor	Right—serous cystadenofibroma; left— serous carcinoma arising in cystadenofibroma (minor discordance)
0062	D	Endometrioid adenocarcinoma, grade 1	Serous cystadenoma	Right—serous cystadenofibroma; left— villoglandular carcinoma, favor endometrioid (minor discordance)
0342	A	Sparse superficial strips and fragments of endometrial tissue	CIN3 with benign endometrial mucosa	Benign lower uterine segment mucosa
0346	A	Endometrial polyp. Admixture of late secretory and menstrual pattern endometrium	Complex atypical endometrial hyperplasia	Endometrial polyp and benign interval-phase endometrial mucosa with breakdown
0349	A	Sparse superficial strips of atrophic endometrial epithelium	CIN3 and changes suggestive of endocervical adenocarcinoma in situ	Benign endocervical mucosa
0361	A	Endometrial polyp with florid secretory glands in a background of weakly proliferative stromal breakdown	Complex atypical endometrial hyperplasia in endometrial polyp	Benign secretory endometrial mucosa
0361	B	Endometrial polyp with florid secretory glands in a background of weakly proliferative stromal breakdown	Endometrioid adenocarcinoma, FIGO grade 1	Benign endometrium
0361	D	Endometrial polyp with florid secretory glands in a background of weakly proliferative stromal breakdown	Atypical endometrial hyperplasia fragments of endometrial polyp	Benign endometrial polyp
0370	A	Secretory pattern endometrium	Complex atypical endometrial hyperplasia	Simple endometrial hyperplasia without atypia
0384	A	Proliferative pattern endometrium with stromal breakdown	Endometrial adenocarcinoma, FIGO grade 1	Focal simple endometrial hyperplasia without atypia
0411	B	CIN1	CIN2	CIN1
0418	A	CIN2	Endocervical glandular hyperplasia	Atypical endocervical glandular proliferation (minor discordance)
0418	B	CIN2	No pathologic abnormality	CIN3
0418	C	CIN2	Free-floating fragments of unoriented dysplastic squamous epithelium, favor LGSIL	CIN2
0424	D	Focal acute cervicitis	Focal high-grade squamous intraepithelial lesion (moderate dysplasia)	Squamous metaplasia

*The 4 reading pathologists at this site were designated A, B, C, and D.
†Original sign-out diagnosis.
‡Concordance or minor discordance compared with reference standard.

greatly increasing the educational value of the images provided.[52,53] The use of WSI also eliminates the need for providing glass slides and recuts to students for educational purposes and ensures that every student views the same image. The reader is referred to reviews that address these issues in greater detail.[50,54] Digital pathology also has the potential to underpin more advanced approaches to image analysis of tissues to provide quantitative data at the point of scanning that can support case selection, prioritization and diagnostic evaluation of tissues to support tumor grading, biomarker measurement, patient stratification, immuno-oncology and precision medicine.

The wide variety of cases included in this series allowed us to perform a detailed analysis of major discordance rates by organ system in order to determine if there were specific organ systems, specimen types, or diagnostic categories where WSI was consistently inferior to microscopy. Although we did identify a few organ systems where the major discordance rate for WSI (vs. reference standard) was slightly higher than the major discordance rate for microscopy (vs. reference standard),

TABLE 10. Comparison of Current Study to Selected Prior Studies Comparing WSI and Microscopy for Primary Diagnosis in Surgical Pathology

Study	Campbell	Bauer	Snead	Current Study
Year of publication	2012	2013	2016	2017
No. cases	212 (312 parts)	607 (1025 parts)	3017	1992
No. pathologists interpreting slides in the study	2	2	17	16
No. reads*	424	1025	3017	15925
Training period on WSI was part of study	Not specified	Yes	Yes	Yes
Device used for WSI	iScan	Aperio ScanScopeXT	Omnyx VL4	Philips IntelliSite Pathology Solution
Magnification at which scanning was performed	20×	20×	40× (60× for renal biopsies and some others)	40×
Wash-out period	Not specified	1 y	3 wk	4 wk
Consecutive cases selected	Yes	Yes	Yes	No
Separate pathologist validated enrolled cases	Not specified	No	No	Yes
Cases randomized for reading order	Not specified	Not specified	Not specified	Yes
Batches alternated for reading modality	Not specified	Not specified	Not specified	Yes
Reading pathologists blinded to original sign-out diagnosis	Yes	Yes	Yes	Yes
Reading pathologists evaluated each case by both WSI and microscopy during study	No	No	No	Yes
Adjudication only by pathologists different from reading pathologists	Yes	Yes	No	Yes
Adjudication only by pathologists from different institution than reading pathologists	No	No	No	Yes
Adjudication pathologists blinded†	No	Not specified	Not Specified	Yes
Statistician blinded to modality and randomization scheme	Not specified	Not specified	Not specified	Yes
Areas of concern for WSI	None	Over-calling of eosinophilic esophagitis, missed intramucosal adenocarcinoma, one case with higher Gleason score, one active colitis called normal	Over-calling of dysplasia, under-detection of organisms	Endocrine pathology, neoplastic kidney pathology

*Number of diagnoses rendered specifically for the study (excluding original sign-out microscopic diagnosis performed for patient care).
†Adjudication pathologists were blinded to reading modality, clinical diagnosis, reading pathologist, and "correct diagnosis."

a case-by-case analysis revealed no consistent vulnerabilities for WSI when compared with microscopy. As 4 pathologists interpreted each case by both modalities after the original sign-out diagnosis, one would expect that if there were a consistent technical problem that precluded an accurate diagnosis with WSI, it would manifest as major discordances between WSI and the reference standard for a given case but concordant diagnoses between microscopy and the reference standard on the same case. Further, one would expect this to occur with all 4 pathologists who viewed the case. For example, if identification of nuclear features of papillary thyroid carcinoma was a consistent Achilles' heel of WSI, one would expect that all 4 pathologists would misinterpret cases of papillary thyroid carcinoma by WSI while making the correct diagnosis by microscopy. Instead, our analysis showed that even in the most problematic areas (eg, thyroid pathology), there was not even a single case where all 4 pathologists consistently erred when using WSI while making the correct diagnosis by microscopy. These findings lend additional support to the contention that cases where pathologists make an incorrect diagnosis by WSI and a correct diagnosis by microscopy represent random error by individual pathologists rather than a systematic or technical problem attributable to the use of WSI. It is important to note that although this manuscript focuses heavily on potential vulnerabilities of WSI, there were also areas where microscopy

performed worse than WSI (Table 4). The choice to not subject these areas to the same degree of scrutiny as WSI was made since potential areas of vulnerability of WSI are of greater concern to pathologists.

The strengths of this study include the multicenter, blinded, randomized design, the inclusion of a wash-out period, representation of both academic pathologists as well as pathologists based in commercial laboratories, reading of cases by pathologists who were not experts in the organ systems they were assigned to interpret, and the inclusion of margins and lymph nodes in many cases with resected tumors, closely simulating real-life settings. The ability to recall cases (memory bias)—a major concern in any intraobserver variability study—was minimized by using a large number of cases, consecutive cases with many "routine" diagnoses, a wash-out period, and randomizing the reading order.

In summary, this study demonstrates that WSI is noninferior to microscopy for the purpose of making a primary diagnosis in surgical pathology. This conclusion applies across a wide range of organ systems, sampling methods, specimen types, stains, and practice settings. Our findings have the potential to significantly alter the workflow of surgical pathologists in coming years and pave the way for a purely digital workflow analogous to the process currently used by radiologists.

REFERENCES

1. 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;137:1710–1722.
2. 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;68:499–505.
3. Bauer TW, Slaw RJ. Validating whole-slide imaging for consultation diagnoses in surgical pathology. *Arch Pathol Lab Med.* 2014;138: 1459–1465.
4. Jones NC, Nazarian RM, Duncan LM, et al. Interinstitutional whole slide imaging teleconsultation service development: assessment using internal training and clinical consultation cases. *Arch Pathol Lab Med.* 2015;139:627–635.
5. 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: 1949–1953.
6. 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;18:109–116.
7. 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;43:702–707.
8. Al-Janabi S, Huisman A, Vink A, et al. Whole slide images for primary diagnostics in dermatopathology: a feasibility study. *J Clin Pathol.* 2012;65:152–158.
9. Al-Janabi S, Huisman A, Nikkels PG, et al. Whole slide images for primary diagnostics of paediatric pathology specimens: a feasibility study. *J Clin Pathol.* 2013;66:218–223.
10. Al-Janabi S, Huisman A, Jonges GN, et al. Whole slide images for primary diagnostics of urinary system pathology: a feasibility study. *J Renal Inj Prev.* 2014;3:91–96.
11. Campbell WS, Hinrichs SH, Lele SM, et al. Whole slide imaging diagnostic concordance with light microscopy for breast needle biopsies. *Hum Pathol.* 2014;45:1713–1721.
12. Eccher A, Neil D, Ciangherotti A, et al. Digital reporting of whole-slide images is safe and suitable for assessing organ quality in preimplantation renal biopsies. *Hum Pathol.* 2016;47:115–120.
13. Fine JL, Grzybicki DM, Silowash R, et al. Evaluation of whole slide image immunohistochemistry interpretation in challenging prostate needle biopsies. *Hum Pathol.* 2008;39:564–572.
14. Gage JC, Joste N, Ronnett BM, et al. A comparison of cervical histopathology variability using whole slide digitized images versus glass slides: experience with a statewide registry. *Hum Pathol.* 2013; 44:2542–2548.
15. Ho J, Parwani AV, Jukic DM, et al. Use of whole slide imaging in surgical pathology quality assurance: design and pilot validation studies. *Hum Pathol.* 2006;37:322–331.
16. Jen KY, Olson JL, Brodsky S, et al. Reliability of whole slide images as a diagnostic modality for renal allograft biopsies. *Hum Pathol.* 2013;44:888–894.
17. Kalinski T, Zwönitzer R, Sel S, et al. Virtual 3D microscopy using multiplane whole slide images in diagnostic pathology. *Am J Clin Pathol.* 2008;130:259–264.
18. Kondo Y, Ijima T, Noguchi M. Evaluation of immunohistochemical staining using whole-slide imaging for HER2 scoring of breast cancer in comparison with real glass slides. *Pathol Int.* 2012;62:592–599.
19. Krishnamurthy S, Mathews K, McClure S, et al. Multi-institutional comparison of whole slide digital imaging and optical microscopy for interpretation of hematoxylin-eosin-stained breast tissue sections. *Arch Pathol Lab Med.* 2013;137:1733–1739.
20. Loughrey MB, Kelly PJ, Houghton OP, et al. Digital slide viewing for primary reporting in gastrointestinal pathology: a validation study. *Virchows Arch.* 2015;467:137–144.
21. Nassar A, Cohen C, Agersborg SS, et al. A multisite performance study comparing the reading of immunohistochemical slides on a computer monitor with conventional manual microscopy for estrogen and progesterone receptor analysis. *Am J Clin Pathol.* 2011;135:461–467.

22. Nassar A, Cohen C, Albitar M, et al. Reading immunohistochemical slides on a computer monitor—a multisite performance study using 180 HER2-stained breast carcinomas. *Appl Immunohistochem Mol Morphol.* 2011;135:19–212-217.
23. Nunes C, Rocha R, Buzelin M, et al. High agreement between whole slide imaging and optical microscopy for assessment of HER2 expression in breast cancer: whole slide imaging for the assessment of HER2 expression. *Pathol Res Pract.* 2014;210:713–718.
24. 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;68:33–39.
25. Reyes C, Ikpatt OF, Nadji M, et al. Intra-observer reproducibility of whole slide imaging for the primary diagnosis of breast needle biopsies. *J Pathol Inform.* 2014;5:5.
26. Rodriguez-Urrego PA, Cronin AM, Al-Ahmadie HA, et al. Inter-observer and intraobserver reproducibility in digital and routine microscopic assessment of prostate needle biopsies. *Hum Pathol.* 2011;42:68–74.
27. Shah KK, Lehman JS, Gibson LE, et al. Validation of diagnostic accuracy with whole-slide imaging compared with glass slide review in dermatopathology. *J Am Acad Dermatol.* 2016;75:1229–1237.
28. van der Post RS, van der Laak JA, Sturm B, et al. The evaluation of colon biopsies using virtual microscopy is feasible. *Histopathology.* 2013;63:114–121.
29. Velez N, Jukic D, Ho J. Evaluation of 2 whole-slide imaging applications in dermatopathology. *Hum Pathol.* 2008;39:1341–1349.
30. Weinstein RS, Descour MR, Liang C, et al. An array microscope for ultrarapid virtual slide processing and telepathology. Design, fabrication, and validation study. *Hum Pathol.* 2004;35:1303–1314.
31. Snead DR, Tsang YW, Meskiri A, et al. Validation of digital pathology imaging for primary histopathological diagnosis. *Histopathology.* 2016;68:1063–1072.
32. Bauer TW, Schoenfield L, Slaw RJ, et al. Validation of whole slide imaging for primary diagnosis in surgical pathology. *Arch Pathol Lab Med.* 2013;137:518–524.
33. Fonyad L, Krenac T, Nagy P, et al. Validation of diagnostic accuracy using digital slides in routine histopathology. *Diagn Pathol.* 2012;7:35.
34. Gilbertson JR, Ho J, Anthony L, et al. Primary histologic diagnosis using automated whole slide imaging: a validation study. *BMC Clin Pathol.* 2006;6:4–19.
35. Jukic DM, Drogowski LM, Martina J, et al. Clinical examination and validation of primary diagnosis in anatomic pathology using whole slide digital images. *Arch Pathol Lab Med.* 2011;135: 372–378.
36. Al-Janabi S, Huisman A, Nap M, et al. Whole slide images as a platform for initial diagnostics in histopathology in a medium-sized routine laboratory. *J Clin Pathol.* 2012;65:1107–1111.
37. Brunelli M, Beccari S, Colombari R, et al. iPathology cockpit diagnostic station: validation according to College of American Pathologists Pathology and Laboratory Quality Center recommendation at the Hospital Trust and University of Verona. *Diagn Pathol.* 2014;9 (suppl 1):S12.
38. Buck TP, Dilorio R, Havrilla L, et al. Validation of a whole slide imaging system for primary diagnosis in surgical pathology: a community hospital experience. *J Pathol Inform.* 2014;1:43.
39. Campbell WS, Lele SM, West WW, et al. Concordance between whole-slide imaging and light microscopy for routine surgical pathology. *Hum Pathol.* 2012;43:1739–1744.
40. Cheng CL, Azhar R, Sng SH, et al. Enabling digital pathology in the diagnostic setting: navigating through the implementation journey in an academic medical centre. *J Clin Pathol.* 2016;69:784–792.
41. 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;67:1052–1055.
42. Pagni F, Bono F, Di Bella C, et al. Virtual surgical pathology in underdeveloped countries: the Zambia project. *Arch Pathol Lab Med.* 2011;135:215–219.
43. Goacher E, Randell R, Williams B, et al. The diagnostic concordance of whole slide imaging and light microscopy. *Arch Pathol Lab Med.* 2017;141:151–161.

44. Raab SS, Nakhleh RE, Ruby SG. Patient safety in anatomic pathology. Measuring discrepancy frequencies and causes. *Arch Pathol Lab Med.* 2005;129:459–466.

45. Mills AM, Gradecki SE, Horton BJ, et al. Diagnostic efficiency in digital pathology. A comparison of optical vs. digital assessment in 510 surgical pathology cases. *Am J Surg Pathol.* 2017. In Press.

46. Elsheikh TM, Asa SL, Chan JK, et al. Interobserver and intra-observer variation among experts in the diagnosis of thyroid follicular lesions with borderline nuclear features of papillary carcinoma. *Am J Clin Pathol.* 2008;130:736–744.

47. McKenney JK, Simko J, Bonham M, et al. The potential impact of reproducibility of Gleason grading in men with early stage prostate cancer managed by active surveillance: a multi-institutional study. *J Urol.* 2011;186:465–469.

48. van der Kwast TH, Evans A, Lockwood G, et al. Variability in diagnostic opinion among pathologists for single small atypical foci in prostate biopsies. *Am J Surg Pathol.* 2010;34:169–177.

49. Huisman A, Looijen A, van den Brink SM, et al. Creation of a fully digital pathology slide archive by high-volume tissue slide scanning. *Hum Pathol.* 2010;41:751–757.

50. Pantanowitz L, Szymas J, Yagi Y, et al. Whole slide imaging for educational purposes. *J Pathol Inform.* 2012;3:46.

51. Ayad E. Virtual telepathology in Egypt, applications of WSI in Cairo University. *Diagn Pathol.* 2011;6 (suppl 1):S1.

52. Fuller MY, Mukhopadhyay S, Gardner JM. Using the Periscope live video-streaming application for global pathology education: a brief introduction. *Arch Pathol Lab Med.* 2016;140:1273–1280.

53. Hwang DH, Szeto DP, Perry AS, et al. Pulmonary large cell carcinoma lacking squamous differentiation is clinicopathologically indistinguishable from solid-subtype adenocarcinoma. *Arch Pathol Lab Med.* 2014;138:626–635.

54. Taylor CR. From microscopy to whole slide digital images: a century and a half of image analysis. *Appl Immunohistochem Mol Morphol.* 2011;19:491–493.