



Topic: CAP Validating Whole Slide Imaging (WSI) for Diagnostic Purposes in Pathology

Date: September 15, 2016

Why is validation needed? It is a tool to present an image to a pathologist for diagnosis. We don't have to validate our microscopes so why validate WSI systems?

By substituting a whole slide imaging device for the microscope, we have now potentially introduced a new set of technical limitations and artifacts which our traditional training may not have prepared us to handle. For instance, while an out of focus image may be easily recognized using either WSI or a traditional microscope, other situations may not be as readily apparent. What if the WSI device has problems finding small groups of cells or misses cells at the periphery of the slide and therefore doesn't capture any images of them? What if the standard WSI capture settings don't have sufficient resolution to make certain diagnoses that require examination of sub-cellular detail such as neuroendocrine nuclear features? Pathologists are responsible to ensure that the diagnoses they make using WSI are based on best practices. As with other aspects of laboratory medicine, we require validation when a major component of a clinical testing system is changed. Pathologic diagnosis using this new technology should be no different.

WSI systems are not yet cleared for primary diagnostic use by the Food and Drug Administration (FDA). Why are these guidelines coming out now?

The work of the CAP panel began and was well underway before the FDA announced the requirement of premarket approval for use of WSI systems for primary diagnosis. The panel, nonetheless, decided to complete their guideline work for a number of reasons:

1. Since the FDA has not delineated the requirements for a premarket approval study, we hope publishing our guidelines may help facilitate and influence that process
2. The FDA requirement for premarket approval is specifically for primary diagnostic use; the FDA has, to date, remained silent on requirements for the use of WSI for other applications such as interpretation of immunohistochemistry, frozen section, and consultations/second opinions
3. These guidelines can still be applied to use of this technology for purposes outside the scope of FDA oversight, such as clinical trials or veterinary applications
4. These guidelines are currently applicable for international laboratories in concordance with their countries regulatory requirements.

Are these guidelines required by the CAP Laboratory Accreditation Program (LAP)?

There are two requirements in the Laboratory General Checklist that pertain to WSI: GEN.52900 and GEN.52920. Please refer to the checklist to read the requirements.

What is the definition of primary diagnostic use for pathology diagnosis?

"Primary diagnostic use" of any visual image, be it by light microscopy, static images (eg, jpegs), live video images or interactive whole slide digital images, is interpreted as the use of that visual image as the primary basis for establishing a pathologic diagnosis. The diagnosis may be based purely on that image, or may also rely on adjunct testing (eg, immunohistochemistry). If adjunct testing is employed, the image is considered of primary diagnostic use if its interpretation serves as the basis on which adjunct studies are selected and correlated. With regard to digital images being used to render a primary diagnosis, the final diagnosis is made using the digital slide and not the glass slide.

This visual data used for primary diagnosis may represent tissue-based or fluid-based slides and a variety of staining techniques.

Uses of WSI that would not be considered primary diagnostic use include:

- Digital imaging studies for biomarker testing



- Interpreting digital slides with immunohistochemical or in situ hybridization tests that augment or refine the primary diagnosis
- Frozen sections using digital imaging, where a glass slide is subsequently reviewed to provide a final diagnosis.
- Intradepartmental or extradepartmental consultation (ie, second opinion)

How do we categorize diagnostic applications and do all pathologists agree? The ones listed (FS and IHC) seem distinct, but what about IHC vs ISH, smears vs thin-layer, etc?

Categorizing different diagnostic applications for WSI pertains to the recommendation that a validation study be performed “for each intended diagnostic purpose.” Exercise of judgment on the part of the pathologist and laboratory is necessary to determine whether a different validation should be performed for a given diagnostic application or specimen preparation. The medical director must determine whether a different validation should be performed for a given diagnostic application or specimen preparation. Factors important for making this decision might include the types of artifacts to which different preparations are subject, the morphologic features on which pathologists rely when evaluating different preparations/slides/stains, and the capabilities required in the WSI system for a pathologist to interpret the image (eg, magnification level, fluorescence, etc). Evidence from the literature suggested that the specimen preparation type was a much more important performance variable than the source of the tissue or the specific analyte being assessed; therefore, a single validation study may suffice to cover a group of similar intended uses, as long as the overall process of preparation and interpretation is the same. For example, when reading digitized immunohistochemistry slides, the study need only validate that digital slides are able to capture the expected chromogen color(s), intensity, and localization.

Does each special stain require validation?

Special stains belonging to specific categories probably do not need to be validated independently. For example, an immunohistochemical stain using DAB as a chromogen may be effectively validated by 20 cases included in the validation set. Within such a validation set, one might consider including a mix of membranous, cytoplasmic, and nuclear immunoreaction signals. If a different chromogen is employed, another validation set should be considered. The general concept guiding whether a separate validation needs to be performed involves the sameness of stains and procedures, allowing similar stains to be included under the umbrella of a particular validation set.

How can you validate the ability of pathologists to use WSI for frozen section diagnoses “emulating the real-world clinical environment”?

In informatics, an important guiding principle before bringing on any new technology or information system is to make sure that it performs as expected in the test environment before going live. The same is true when introducing a new WSI scanner for clinical use in the laboratory. The WSI system should be validated (tested) before going live for its intended use. That is why the third CAP recommendation suggests that a validation study of a WSI system carried out by a laboratory should closely emulate the real-world clinical environment in which the technology will be used. This is important because approval of the WSI system will be limited to the conditions under which validation occurred. If the WSI system is intended to be used for frozen sections, then each case examined during validation should compare the diagnosis rendered at frozen section using a traditional microscope against the diagnosis rendered when examining the same frozen section slides as whole slide images. Of course, the pathologist rendering the diagnosis on the digital images must have access to the same information when using a traditional microscope (eg, clinical information, specimen location, other pertinent gross examination information as appropriate). If rapid digitization of glass slides is typically required for frozen sections, then the validation process should also include a determination as to whether the WSI system of choice can facilitate accurate diagnosis within the same specified turnaround time parameters. Finally, it is also advisable to use the scanner in the intended location (eg, frozen section room) and with the same workflow (eg, scanning performed by pathology assistants, immediately after a coverslip is added when the slide may still be wet) as it will be utilized clinically for frozen sections.



For many diagnostic applications (frozen sections, primary reads, second opinions, etc), the validation study that needs to be performed should determine whether the WSI images are 'good enough' for making accurate and complete diagnoses. If this validation is done for a challenging application such as frozen sections, why does it need to be repeated for primary reads?

Recommendation #2 states that validation of WSI systems should involve specimen preparation types relevant to the intended use. Specifically, the validation process should include specimens and histologic features one would reasonably expect to encounter in the setting of primary diagnosis. The spectrum of specimens, stains and histologic features encountered with intraoperative consultations in most pathology departments will not be broad enough to allow simple extrapolation of the frozen section validation to primary diagnosis. In this sense, primary diagnosis differs materially from the previously validated frozen section use and a separate validation should be performed. The validation process for primary diagnosis may also identify situations where special scanning protocols are required. An example not likely to be encountered at frozen section would be the need for default scanning of certain slides at 40x (as opposed to the standard 20x) for the identification of *H. pylori* in gastric biopsies.

Do the vendor(s) have to indicate that a new release with hardware/software changes may alter the output of the WSI system? Who defines what a "significant change" is versus a "minor change"?

Getting an entirely new WSI scanner would be considered a significant change to the system, which would require revalidation. When a new release of the hardware or software is being installed, it will be important for the medical director, in consultation with the vendor and other members of the laboratory, to understand what changes have been introduced. As with any other change to equipment (eg, a tissue processor) or software (eg, a laboratory information system upgrade) used in the laboratory, it ultimately will be the decision of the medical director as to whether or not a particular change has the potential to affect the accuracy of the diagnoses being rendered. For example, "minor" changes to the power supply of the scanner, the screen layout of the control software, or naming conventions for the data files are unlikely to be deemed "significant." In contrast, "significant" modifications to the tissue-finding algorithm should certainly trigger a revalidation. The medical director should be responsible for determining what change in the system would be considered significant enough to potentially alter a clinical diagnosis.

Do we need to perform a full revalidation if we change monitors? What about the future and the possibility of signing out cases on home computers and smartphones?

Revalidation is only required whenever there is a significant change made to any component of the WSI system (Guideline Statement #5). A slight change of monitor (eg, the make of monitor, location of the monitor in a pathologist's home office) in our opinion is a minor change that should be handled through a facility's change management process. However, viewing digitized slides on a small display of a tablet computer or smartphone would represent a major change, and therefore require revalidation. There is only limited information on this topic in the medical literature.

What about recommendations for training in use of WSI? Does the complete process require repetition whenever there is a change in staff?

It is important to distinguish validation of the equipment from training of the users. The guidelines pertain to validation of the equipment, and Guideline Statement #6 indicates that at least one "pathologist adequately trained to use the WSI system must be involved in the validation process." This training should include an understanding of the operation of the equipment, options for its configuration, and a familiarity with any limitations of the technology. However, not all potential users need to be involved in the validation process. Once the equipment has been validated for a particular use, the medical director should determine what training is necessary for additional pathologists planning to use the system. While the panel agrees that it is important for staff to be properly trained to use these WSI instruments, the purpose of these recommendations is not to provide specific guidance on such training.



How do you account for the fact that a pathologist's memory of previously viewed cases often extends significantly beyond 2 weeks?

This is an issue where the literature gives no clear guidance since there are limited studies available which address this question. If this is a particular concern when designing a validation procedure, use of a randomization protocol is suggested (eg, randomly assign cases to be viewed first with either WSI or traditional microscopy). With enough cases, any recall bias should be evenly distributed. Due to the need to keep in mind practical limitations, we suggested the 2 week wash out period as being a reasonable compromise. As a point of information, the FDA has recently recommended a 1-month wash-out period.

What level of agreement/kappa would be the threshold for acceptable intraobserver performance? Should there be a benchmark? Can intraobserver performance be assessed alone or in conjunction with interobserver performance?

The panel believes that, as a validation criterion, it is most important that a pathologist is able to repeatedly make the same diagnosis on the same slide, whether viewed with a microscope or by using WSI. As long as the validation study records this intraobserver concordance, there is no reason why interobserver performance can't also be measured. Indeed, the aim of the validation study would be to achieve a high concordance rate between diagnoses made using glass versus digital slides. While we are not aware of any current benchmarks (eg, kappa value), laboratories should note that intraobserver concordance is unlikely to always be perfect because even discrepancy rates for second opinion glass slide review have been reported to range from 1.4% to 30%. Hence, the medical director should be responsible for determining an appropriate diagnostic concordance rate for their laboratory WSI validation study.

Can we implement the guideline now for primary diagnosis?

If you interpret images from outside the United States where FDA and CLIA regulations do not apply, these guidelines can be implemented immediately. Matters are less clear for images scanned and interpreted in the United States. As a Class 3 device under current FDA rules, scanners cannot yet be *marketed* as devices for making primary diagnoses with WSI, and vendors may have constraints selling to you if that is your purpose in purchasing the instrument. It is not clear that use of WSI for primary diagnosis is prohibited as an off-label, laboratory developed test, but no clear ruling exists as yet.

Does this guideline address WSI for image analysis?

No, the guideline does not specifically address issues related to image analysis. It covers only primary diagnosis and visual interpretation of ancillary (eg, immunohistochemical) staining. It does not address issues such as measuring distances or surface areas, annotating images, automated interpretation, quantification of immunohistochemical stains or immunofluorescence or other computer aided diagnostics. The CAP is currently developing a guideline to address quantitative image analysis for HER2 Immunohistochemistry.

Does the guideline address the retention of image files? Would stored digital slides be acceptable to replace glass slides?

The guidelines do not explicitly deal with the question of how long to retain image files. The intended clinical use may provide guidance as to their retention, following existing guidelines for the retention of glass slides, (eg, 10 years for diagnostic cases or abnormal cytology slides). Institutions may choose not to retain the image files for an extended duration if the glass slides are being stored for the required duration, given that an intact glass slide can always be digitized again. While there are archival advantages of WSI in terms of speed of recall, in certain settings storage costs for digital data may be prohibitively high. Stored glass slides that harbor cellular material or tissue are a potential source of DNA or other molecular information that should be taken into consideration. Conversely, retained images can be used to unequivocally demonstrate what the pathologist reviewed at the time of diagnosis, thereby taking into account any image-related artifacts, tissue not visible, resolution, etc.



What are the important patient identification and privacy issues related to WSI? (eg, HIPAA)

If WSI is used for diagnostic or other related clinical purposes, procedures must be in place that ensure that sites using WSI provide reasonable and expected confidentiality and data security, in both data storage and data transmission. The security and privacy requirements of HIPAA apply, as they would for any other potential use of protected health information (PHI). Procedures might include message security, system and user authentication, activity logs, encryption, and access restrictions. With respect to patient identification, as is the case for any laboratory analysis, processes, procedures, and training must be in place to ensure that patient identification linked to glass and digital slides is accurate, maintained, and secure. There are multiple ways to ensure positive patient identification, including use of verbal communications, barcodes or images of slide labels.

Is there currently a billing code for using WSI as outlined in the guideline?

Currently, there is no billing code for primary diagnosis using WSI.

CAP WSI Validation Guideline Expert Panel Authors:

Liron Pantanowitz, MD, Chair
John H. Sinard, MD, PhD
Walter H. Henricks, MD
Lisa A. Fatheree, BS, SCT(ASCP)
Alexis B. Carter, MD
Lydia Contis, MD
Bruce A. Beckwith, MD
Andrew J. Evans, MD, PhD
Christopher N. Otis, MD
Avtar Lal, MD, PhD
Anil V. Parwani, MD, PhD

CAP Digital Pathology Committee

Eric F. Glassy, MD, Chair
Thomas W. Bauer, MD, PhD
Marilyn M. Bui, MD, PhD
Toby C. Cornish, MD, PhD
Andrew J. Evans, MD, PhD
Jason Hipp, MD, PhD
Dennis G. O'Neill, MD
Liron Pantanowitz, MD
Anil V. Parwani, MD, PhD
Bill A. Rampy, DO, PhD
Mohamed El-Sayed Salama, MD
Rebecca E. Waters, MD, DO
Douglas A. Murphy, MT(ASCP), staff

If you have further questions about any of the Center guidelines, please email center@cap.org.

Reference:

Pantanowitz L, Sinard JH, Henricks WH, et al. Validating whole slide imaging for diagnostic purposes in pathology: Guideline from the College of American Pathologists (CAP) Pathology and Laboratory Quality Center. *Arch Pathol Lab Med*. doi: 10.5858/arpa.2013-0093-CP. Effective date: May 1, 2013