

Supplementary Materials

S1. Considerations and recommended functional requirements for a whole slide scanner (WSS)

The process of scanning to acquire images might be the most significant aspect of the DP system. When implementing a digital pathology (DP) system using a WSS, it is important to understand that a whole slide imaging (WSI) is a high-resolution copy of a glass slide image. In other words, the actual glass slide image might not be 100% accurately replicated into a digital image due to various factors involved in image acquisition using a scanner. During the scanning process, some image data can be missed or inadvertently omitted because of inappropriately set scan parameters or tissue samples that are too thin or too small (e.g., fine needle aspirate of breast fat tissue or highly necrotic tissue) or an automatic tissue detection system error. Therefore, the person in charge must verify that all important areas of interest are included in the scan range and prepare a plan to prevent errors that can occur for those reasons. The scanner must provide a function that allows users to check whether the digital image was scanned satisfactorily. The manager who evaluates the quality of scanned images must carefully examine whether faint stains, pen marks, foreign objects, air bubbles during sealing, or damage to the cover slide affected the quality of scanned digital images and whether errors such as misalignment of strips or tiles when combining have occurred. It is recommended to prepare a workflow that can selectively figure out it.

S2. Recommended functional requirements for image database systems

The image database system must be able to guarantee that the identification information of the glass slide matches that of the digital image.

S3. Recommended functional requirements for image database systems

Even if the version of the image-archiving software changes, the use of a preserved image should not be problematic.

S4. Recommended functional requirements for image database systems

The type of storage method must include the concept of backup or mirroring to ensure that data are safely preserved (e.g., using Network attached storage, NAS or a redundant array of inexpensive disks, RAID, and the others).

S5. Considerations and recommended functional requirements for image display devices and image viewing software

Image display devices, including monitors, are part of an imaging chain (also called a visualization pipeline), as are optical components such as scanner lenses and image acquisition components such as a charge-coupled device, an electronic component for data processing.

Consideration should be given to the following factors that determine the quality of image display devices: the type of display (such as the size of the device), the structure of the light source (light-emitting or light-receiving), the liquid crystal alignment mode (in-plane switching or vertical alignment), and the structure of the liquid-crystal display and flat panel; the mechanical characteristics of the device, including the resolution (dots per inch), luminance, contrast ratio and contrast, color temperature, color profile of the monitor, viewing angle, response rate, image retention, and burn-in; the mechanical characteristics of the image display system associated with the speed and capacity of the graphic memory in the computer system; and environmental factors such as room lighting, window placement, distance from the observer eye level, and differences in user heights.

When making diagnoses using a DP system, it is often necessary to check clinical information from the patient electronic medical records (EMR) or radiologic data from the PACS, for which multiple monitors can be used. When comparing or observing two or more DP images using multiple monitors, the monitors should have been manufactured in the same year and should be the same model to minimize differences in the images caused by the different monitors.

Because technologies for image display devices such as monitors are advancing rapidly, it is difficult to define the minimum requirements for a DP system based simply on numeric values. Moreover, it is difficult to define the absolute requirements because the luminance, luminance ratio, and contrast can change depending on the office environment. Therefore, optimal functional conditions should be defined according to the situation and needs of each institution.

S6. Considerations and recommended functional requirements for image display devices and image viewing software

Image viewing software can include the following functions: an observation field display that shows overview images (also referred to as preview or macro-images), with the part of the total overview image being

observed indicated within a square; an annotation function that displays the object magnification and length scale (accumulation) on the images and allows users to insert figures or words to mark areas of interest; a function to screen-capture partial or entire images of interest displayed on the monitor; a function that allows side-by-side comparison of DP images from different tests performed on the same patient, such as immunohistochemical or special stains, or DP images from similar cases for reference; and basic morphometric functions, such as measuring the length and area of certain microstructures. Whether these functions can be adequately performed in the workflow of real practice should be determined in advance.

S7. Issues related to integration/links with laboratory information systems (LIS) and EMR systems

The DP system must be linked appropriately to the LIS that stores and manages test records from the pathology laboratory and the HIS or EMR system that manages clinical patient records inside the hospital. The American Telemedicine Association (US), Canadian, and European guidelines recommend linkage and management using standard methods such as HL7.

The DP system must include metadata associated with the digital images (i.e., overview images [preview, macro-images], scan parameters, and data on the scanned area). When the system is linked to the LIS, data such as the test number, tissue information, block number, and staining information must be appropriately linked. Moreover, the linked systems must be checked to confirm the smooth operation of both systems and the link between them under actual workflow conditions.

S8. Guidelines and considerations for validation needed for the implementation of DP systems and internal QC needed during operation

- 1) All pathology laboratories operating WSI-based DP systems for clinical diagnosis must conduct in-house validation studies (Expert consensus)

Variable factors in the testing process could influence DP system performance and validity; thus, a validation study before system implementation is essential. Simply because the DP system has already been approved by relevant authorities through a verification process and is being operated according to the manufacturer's recommended operating protocol does not guarantee the validity of the system for the samples and environment at each institution. The validation results must be appropriately documented and maintained accordingly.

12) Pathology laboratories must maintain documentation regarding the validation of the DP system, including the methods, results, and final approval (Expert consensus)

Validation is implemented by providing documented evidence that the requirements for DP systems have been well met in proper system operation. Therefore, pathology laboratories must keep and manage the documents demonstrating their successful validation of their DP systems, including the methods used, results, and final approval. During the validation, system users should be educated and trained to operate the system and supporting documents showing that this education has been conducted must be prepared and maintained. The final document must contain the signature of the DP system manager or designated representative. In addition, the inclusion of a statement in the pathologic report that a DP system was used for diagnosis is recommended.

S9. Guidelines and considerations for validation needed for the implementation of DP systems and internal QC needed during operation

2) The validation study should be conducted under conditions that are consistent with the clinical use intended by the DP system manufacturer (Recommendation)

Validation is intended to prove that the WSI system is operating as expected according to its intended purpose. Therefore, the specific methods and design of the validation must be consistent with the purpose at the time that the WSI system was manufactured. For example, even if a DP system that was manufactured to run gynecological liquid-based cytology slides has been successfully validated using gynecological liquid-based cytology samples before implementation, it would not be safe to assume that this system would demonstrate the same quality for centrifuged urine cytology samples. Therefore, separate testing must be conducted when the DP system is to be used for purposes other than originally planned. However, if the overall process of sample preparation and interpretation is the same, then a single validation study could be sufficient. For example, when testing immunohistochemical stain slides, having the sample preparation process would obviate the need to individually test all antibodies.

S10. Guidelines and considerations for validation needed for the implementation of DP systems and internal QC needed during operation

3) The validation study should be designed to be as similar as possible to the actual clinical settings in which the technology will be used (Recommendation)

It is not advisable to conduct the validation study by selecting samples that show “typical” pathologic findings for each diagnosis favorable for testing. The validation must represent common cases and should include a sufficient number of borderline cases that could be difficult to diagnose using the DP system, such that the spectrum of diagnostic complexity and difficulty found in actual workflow is adequately represented. In addition to a comparison of diagnostic accuracy, the validation must also include an assessment of its performance with respect to cases that are expected to be more difficult than by microscopy, such as dysplasia grading, calcium oxalate crystal detection, mitosis counting, eosinophil counting, microorganism detection, and viral inclusion detection. This process can be used to facilitate user training and learning, as well as proper validation. For frozen section cases, whether the turn-around time from scanning to diagnosis is similar to that of microscopic diagnosis must also be assessed. If the system is used in a single institution, comparative assessment with other laboratories is not necessary. However, if samples prepared in other institutions are used, then advanced testing of the method is necessary to simulate the same workflow.

S11. Guidelines and considerations for validation needed for the implementation of DP systems and internal QC needed during operation

4) The validation study should cover the entire DP system (Recommendation)

The validation study is a quality assurance (QA) process intended to test the entire process; thus, separate testing of individual system components (e.g., the computer system, monitor, and scanner) or processes is unnecessary.

S12. Guidelines and considerations for validation needed for the implementation of DP systems and internal QC needed during operation

5) Significant changes in the composition of the DP system necessitate re-validation (Expert consensus)

Validation must be repeated whenever significant changes are made to the composition of the DP system, such as the use of a new type of scanner or hardware or software upgrades. The validation could be performed with a smaller number of samples (i.e., 20 samples) if the new scanner was manufactured by the same manufacturer; is the same model as the previously validated scanner; and is used with the same network, image database system, image viewing software, and image display device. Minor changes can be managed according to internal guidelines.

S13. Guidelines and considerations for validation needed for the implementation of DP systems and internal QC needed during operation (Supplementary Material S13) [1].

6) Validation is intended to be conducted by at least one pathologist who has been acclimated to the DP system (Recommendation)

The validation process assumes that a pathologist who has been acclimated to the DP system will make a diagnosis. Therefore, validation should be performed by someone familiar with using the DP system, rather than inexperienced individuals, to eliminate results biased by the tester's level of education and training. Moreover, although the system does not need to be validated by every pathologist who uses it, the validation could include other laboratory personnel (e.g., laboratory managers, histo- or cytotechnicians, and residents), IT managers, or technical advisors. The validation should also include personnel who perform slide scanning.

7) The validation must be performed on at least 60 samples for a single applicable field (e.g., histopathologic H&E-stained slides, frozen sections, cytology slides, blood smear slides) according to the type of sample or test. For additional applicable fields (e.g., immunohistochemical staining, special staining), validation could be performed by adding 20 or more samples (Recommendation).

The number of people involved in the validation and the scale of the validation could vary significantly between institutions. Moreover, it is difficult to accurately calculate the minimum number of samples needed to guarantee 100% validity. The manager of the DP system at each institution must fully consider the scale and characteristics handled by the institution as well as the relevant personnel and include samples with varying degrees of diagnostic difficulty when selecting the appropriate number of samples needed to ensure reliable operation of the DP system. A prospective validation process during 1–3 months of actual operation, as well as a retrospective validation study using prior tests, could be also considered.

S14. Guidelines and considerations for validation needed for the implementation of DP systems and internal QC needed during operation

8) Validation must be carried out using a comparative analysis of concordance between microscopic and WSI-based diagnoses made by a single observer (intra-observer variability assessment) (Suggestion)

Validation is intended to assess the diagnostic concordance between microscopic and WSI-based diagnoses; thus, it must be conducted as an intra-observer variability assessment with repeated assessments by the same observer. The degree of diagnostic concordance can be assessed using a 3-tier system according

to the clinical implications (i.e., major discordance that could drastically affect patient prognosis and treatment; minor discordance that could affect the diagnosis severity without causing changes in patient prognosis and treatment; and minimal discordance with little or no difference in the diagnosis severity and patient prognosis or treatment). The goal of comparative analysis should be to identify the cause of problems related to image quality, such as artifacts during digital image scanning, rather than diagnostic variability resulting from interpretational changes by the observer.

S15. Guidelines and considerations for validation needed for the implementation of DP systems and internal QC needed during operation

10) During the validation, a washout period of at least 2 weeks is needed to minimize the influence of recall bias (Recommendation)

An observer remembering tissue slide images previously examined and their corresponding diagnoses can cause recall bias that could affect validation concordance. Therefore, it is important to perform the validation with a sufficient washout period between the observations. Previous studies and major guidelines recommend a washout period of at least 2 weeks; however, a longer washout period might be more favorable as long as it does not burden the operation of the institution.

9) Validation can be performed using either randomly or sequentially arranged samples (Recommendation)

Intuitively, the random arrangement of samples would seem to minimize the influence of recall bias on validation. However, relevant studies have reported no significant difference between random and sequential assessments.

S16. Guidelines and considerations for validation needed for the implementation of DP systems and internal QC needed during operation

11) During validation, data integrity during image acquisition must be assessed by verifying whether all tissues on the glass slide have been properly scanned to form the digital image. (Expert consensus)

In addition to assessing the diagnostic concordance, the assessment of data integrity during image acquisition is also important with respect to the QA of the DP system. Slides with poor staining quality, images of very small tissues that are out-of-focus, and errors or scan failure during image acquisition should be checked and appropriate measures taken during validation. In addition, it is important to check whether the metadata of the digital images and the slide identification numbers (slide labels) match.

S17. Scope of application

In terms of staining methods and sample types applicable to WSI-based diagnosis using DP systems, most basic tissue slides stained with hematoxylin and eosin (H&E), most special stains, immunohistochemical stains, and frozen section slides are expected to be applicable, though they will require appropriate validation studies followed by trial periods until the users reach a stable learning level. However, for the few special cases listed below, a higher level of validation and longer period of trial operation are recommended.

1) Cytology slides: In most cases with cell smears, liquid-based cytology, or blood smears, the results are similar to those from microscopic diagnoses made using glass slides with appropriate focus stacking (Z-stacking) during scanning and sufficient validation. However, the optimal focus stacking method must be carefully selected for certain sample types (e.g., samples with many 3-dimensional structures, such as thyroid gland fine-needle aspirates), staining conditions, and smear conditions. Excessive focus stacking and the acquisition of higher-magnification images ($60 \times 100 \times$ or higher) could lead to over-scaling of the WSI files, which could negatively affect the operation of the entire system. Therefore, the determination of optimal scanning conditions during the implementation of DP for cytology slides is essential and should be based on a balance between desirable scan quality and file size. Partial image acquisition of the slides is not recommended because it can severely impair diagnostic integrity and accuracy. In conclusion, importing cytology slides into a DP system requires a more extensive validation and trial operation period than is needed for other types of slides, including simultaneous comparison periods of diagnostic results using both WSI- and microscope-based diagnosis.

2) Samples clinically or morphologically suspected of lymphoreticular neoplasms: Lymphoreticular neoplasms have similar morphology at low magnification and similar nuclear features, such as chromatin patterns and nucleoli feature that are important for histologic diagnosis. Accordingly, the acquisition of high-magnification images is generally recommended. Although evidence is lacking about diagnostic agreement at different scanning magnifications, results to date have shown less than 2% of major discordance between WSI-based and microscopic diagnoses using a basic $20 \times$ scan. Moreover, the recorded 8% of minor discordance was mostly due to differences in grading follicular lymphoma and was similar to the inter- and intra-observer diagnostic discordance in microscopic diagnosis. Because pathologic diagnosis of lymphoreticular neoplasms is almost always made in combination with the results of additional tests, such as immunohistochemical staining, diagnostic differences in the findings of H&E slide images alone do not seem

to significantly affect diagnostic accuracy. On the other hand, the image comparison function available in the image viewing software of DP systems might offer the benefit of improved accuracy in the interpretation of immunohistochemical staining. Carefully designed validation studies and trial operations based on those considerations are needed.

3) Detection of microorganisms such as *Helicobacter pylori*: The detection of *H. pylori* infections through the microscopic examination of gastric biopsy tissue samples, especially when special staining such as the Giemsa stain is used, showed specificity and sensitivity comparable to that in other *H. pylori* tests. However, concerns remain about whether microorganisms such as *H. pylori* can be detected successfully in the 40 ×-scanned WSI most commonly used today. A recent study demonstrated that increasing the number of focus stacking layers provided results similar to those obtained via microscopic examination. Without focus stacking, WSI-based detection showed an impaired sensitivity of 0.562 and specificity of 0.818. Therefore, when detecting microorganisms, special care is required, such as using appropriate focus stacking for assessment and mentioning the limitations of such examination results in the report.

S18. Considerations and recommended functional requirements for image display devices and image viewing software

As the period of use passes, the monitor should regularly take measures to verify validity or maintain performance by appropriate methods, because of the image display performance such as luminance reduction, contrast ratio reduction, burn-in phenomenon and others.

S19. Considerations and recommended functional requirements for image database systems

Each institution should determine how long digital image data should be preserved. The guidelines by College of American Pathologists (USA), British Royal College of Pathologists (UK), and Federal Association of German Pathologists (Germany) recommend data preservation for at least 10 years, whereas the guidelines from the Japanese Society of Pathology (JSP) (Japan) recommend permanent preservation, with a minimum of 5 years. The JSP guidelines also recommend that data from the past 5 years be preserved in hot storage, meaning that they are available for immediate use. For reference, the preservation period recommended by the Medical Act of South Korea for glass slides containing pathological tissue is 5 years, which is the requirement for test records or findings among general medical records.

S20. Issues related to telepathology, firewalls, protection of personal information, and mobile device use

Strict technical measures must be in place to ensure information security and protect personal information regardless of the type of terminal being used. Therefore, measures are needed to ensure that transmitted data are not easily released outside the network and that transmitted metadata do not contain personal information to minimize the risk to personal data even if a data leak were to occur.

Reference

1. Chong Y, Kim DC, Jung CK, et al. Recommendations for pathologic practice using digital pathology: consensus report of the Korean Society of Pathologists. *J Pathol Transl Med* 2020; 54: 437-52