



COLLEGE of AMERICAN  
PATHOLOGISTS

# ARCHIVES

of Pathology & Laboratory Medicine

## **EARLY ONLINE RELEASE**

Note: This article was posted on the *Archives* Web site as an Early Online Release. Early Online Release articles have been peer reviewed, copyedited, and reviewed by the authors. Additional changes or corrections may appear in these articles when they appear in a future print issue of the *Archives*. Early Online Release articles are citable by using the Digital Object Identifier (DOI), a unique number given to every article. The DOI will typically appear at the end of the abstract.

The DOI for this manuscript is doi: 10.5858/arpa.2020-0723-CP

The final published version of this manuscript will replace the Early Online Release version at the above DOI once it is available.

# Validating Whole Slide Imaging Systems for Diagnostic Purposes in Pathology

## Guideline Update From the College of American Pathologists in Collaboration With the American Society for Clinical Pathology and the Association for Pathology Informatics

Andrew J. Evans, MD, PhD; Richard W. Brown, MD; Marilyn M. Bui, MD, PhD; Elizabeth A. Chlipala, BS, HTL(ASCP)QIHC; Christina Lacchetti, MHSc; Danny A. Milner Jr, MD, MSc(Epi), MBA; Liron Pantanowitz, MD; Anil V. Parwani, MD, PhD; Kearin Reid, MLIS, MT(ASCP); Michael W. Riben, MD; Victor E. Reuter, MD; Lisa Stephens, MBA, HTLA(ASCP)<sup>CM</sup>; Rachel L. Stewart, DO, PhD; Nicole E. Thomas, MPH, CT(ASCP)<sup>CM</sup>

• **Context.**—The original guideline, “Validating Whole Slide Imaging for Diagnostic Purposes in Pathology,” was published in 2013 and included 12 guideline statements. The College of American Pathologists convened an expert panel to update the guideline following standards established by the National Academies of Medicine for developing trustworthy clinical practice guidelines.

**Objective.**—To assess evidence published since the release of the original guideline and provide updated recommendations for validating whole slide imaging (WSI) systems used for diagnostic purposes.

**Design.**—An expert panel performed a systematic review of the literature. Frozen sections, anatomic pathology specimens (biopsies, curettings, and resections), and hematopathology cases were included. Cytology cases

were excluded. Using the Grading of Recommendations Assessment, Development, and Evaluation approach, the panel reassessed and updated the original guideline recommendations.

**Results.**—Three strong recommendations and 9 good practice statements are offered to assist laboratories with validating WSI digital pathology systems.

**Conclusions.**—Systematic review of literature following release of the 2013 guideline reaffirms the use of a validation set of at least 60 cases, establishing intra-observer diagnostic concordance between WSI and glass slides and the use of a 2-week washout period between modalities. Although all discordances between WSI and glass slide diagnoses discovered during validation need to be reconciled, laboratories should be particularly concerned if their overall WSI–glass slide concordance is less than 95%.

(*Arch Pathol Lab Med.* doi: 10.5858/arpa.2020-0723-CP)

Accepted for publication March 26, 2021.

Supplemental digital content is available for this article. See text for hyperlink.

From the Department of Pathology, Mackenzie Health, Richmond Hill, Ontario, Canada (Evans); the Department of Pathology, Memorial Hermann Southwest Hospital, Houston, Texas (Brown); the Department of Pathology, Moffitt Cancer Center, Tampa, Florida (Bui); Premier Laboratory, LLC, Longmont, Colorado (Chlipala); Policy and Advocacy, American Society of Clinical Oncology, Alexandria, Virginia (Lacchetti); American Society for Clinical Pathology, Chicago, Illinois (Milner); the Department of Pathology, University of Michigan, Ann Arbor (Pantanowitz); the Department of Pathology, Ohio State University Medical Center, Columbus (Parwani); Governance Services (Reid) and Surveys (Thomas), College of American Pathologists, Northfield, Illinois; the Department of Pathology, University of Texas MD Anderson Cancer Center, Houston (Riben); the Department of Pathology, Memorial Sloan-Kettering Cancer Center, New York, New York (Reuter); the Department of Anatomic Pathology, Cleveland Clinic, Cleveland, Ohio (Stephens); and Janssen Research & Development, Spring House, Pennsylvania (Stewart).

Authors' disclosures of potential conflicts of interest and author contributions are found in the Appendix at the end of this article.

Corresponding author: Andrew J. Evans, MD, PhD, Mackenzie Health, Room 1311, 10 Trench St, Richmond Hill, ON, Canada L4C 4Z3 (email: Andrew.evans@mackenziehealth.ca).

In 2013, the Pathology and Laboratory Quality Center for Evidence-based Guidelines (the Center) of the College of American Pathologists (CAP) released a guideline on the validation of whole slide imaging (WSI) for diagnostic purposes.<sup>1</sup> For the purposes of this guideline, validation is defined as a process that demonstrates WSI will perform as expected for its intended use and environment prior to using it for patient care. The specific aims are to ensure pathologists make accurate diagnoses to at least the same level as light microscopy, and to identify and control interfering artifacts or technological risks that WSI could introduce to patient safety. The Center chose to develop this guideline in 2010 because of growing interest in using WSI for diagnostic purposes coupled with an expected wider adoption and a lack of evidence-based guidelines for clinical laboratories to validate the technology. The fundamental question addressed by the original guideline panel was, “What needs to be done to validate a WSI system for diagnostic purposes before it is placed into clinical service?” Following a comprehensive literature search, the Australian

National Health and Medical Research Council methodology was used to grade recommendations based on strength of evidence, consistency, clinical impact, generalizability, and applicability to WSI with oversight by a methodologist consultant. The Australian National Health and Medical Research Council methodology generated 4 categories of statements based on decreasing strength of evidence: recommendations, suggestions, expert opinions, or no recommendation offered.<sup>2</sup> The 2013 guideline had 12 recommendations. In terms of impact since its online release on May 1, 2013, the guideline has received 253 separate citations in publications originating in 39 countries. Most of the citations have occurred in comparative studies, narrative reviews, validation studies, and conference papers in journals originating in the United States (52%; 131 of 253), United Kingdom (13%; 33 of 253), Italy (9%; 24 of 253), and Canada (9%; 22 of 253).<sup>3</sup> The guideline has also had 11 871 downloads from *Archives of Pathology & Laboratory Medicine* (Allen Press Technical Support, email communication, September 13, 2020).

According to Center policy, guidelines are subject to review and revision every 4 years, or at 2 years should significant new evidence appear in the literature. By the end of 2017, the amount of WSI validation literature had substantially increased, demonstrating excellent diagnostic concordance between WSI and glass slide diagnoses. The US Food and Drug Administration granted 2 WSI systems approval for primary diagnosis, and several laboratories around the world had already been using WSI for this purpose.<sup>4-9</sup> Although other digital pathology modalities exist (eg, email of static JPEG images, robotic static/dynamic telemicroscopy, real-time video telemicroscopy), the Center expert panel chose to focus on WSI because other WSI vendors have or will have received Food and Drug Administration clearance by the 510(k) route and WSI will form the basis of artificial intelligence/machine learning technology. In addition, the integration of laboratory information systems with digital pathology solutions is essential for patient safety, especially for high-volume digital reporting. Laboratory information system integration has been available for several years for WSI, but not for other digital pathology modalities.

This guideline update differs from the 2013 guideline in 2 respects. First, the CAP collaborated with the Association for Pathology Informatics and the American Society for Clinical Pathology on the development of the guideline. In addition to having representative members on the expert panel, both organizations assisted in dissemination of the open comment period. Second, the guideline revision process used the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) approach as opposed to the Australian National Health and Medical Research Council system previously used. The GRADE approach is used by many organizations worldwide and is considered a standard in guideline development.<sup>10</sup>

As a result of using GRADE, recommendations are now classified as either strong or conditional. This differs from the 4-tiered strength of recommendations used in the original guideline. Additionally, GRADE introduces the concept of good practice statements (GPSs) for issues that are important and intuitive but lack evidence on which to base a recommendation. Although these changes impact both the strength and number of recommendations, the guideline update affirms almost all of the principles established in the original guideline.

## DESIGN

This evidence-based guideline was developed and revised following the standards established by the National Academy of Medicine.<sup>11</sup> A detailed description of the methods and the systematic review (including the quality assessment and complete analysis of the evidence) used to create this guideline can be found in the [supplemental digital content](#) (SDC), also containing 12 tables and 2 figures.

## EXPERT PANEL COMPOSITION

The CAP in collaboration with the Association for Pathology Informatics and the American Society for Clinical Pathology convened multidisciplinary expert and advisory panels to revise the 2013 guideline. The expert panel included pathologists with considerable experience in WSI, 2 histotechnologists, and a research methodologist. The CAP approved the appointment of the members. Detailed information about the panel composition and the role of the panels can be found in the SDC.

## CONFLICT OF INTEREST POLICY

The collaborators agreed upon a conflict of interest (COI) policy (effective November 2017) and members of the expert panel disclosed all financial interests from 12 months prior to appointment through the development of the guideline. Individuals were instructed to disclose any relationship that could be interpreted as constituting an actual, potential, or apparent conflict. Complete disclosures of the expert panel members are listed in the Appendix. Disclosures of interest judged by the oversight group to be manageable conflicts are as follows: M.B.: research grants, National Institutes of Health/National Cancer Institute (Bethesda, Maryland), and consultancies, Hologic, Inc (Bedford, Massachusetts) and ContextVision (Stockholm, Sweden); E.C.: ownership/partnership, Premier Laboratory, LLC (Boulder, Colorado); A.P.: boards/advisory boards, ContextVision (Stockholm, Sweden), PathPresenter (New York, New York), and Digital Pathology Association; LP.: consultancies, Hamamatsu Photonics KK (Hamamatsu City, Japan), Leica Biosystems (Wetzlar, Germany), and Ibex Medical Analytics (Tel Aviv-Yafo, Israel), research grants, Ibex Medical Analytics (Tel Aviv-Yafo, Israel), Huron Digital Pathology (St Jacobs, Ontario, Canada), and Lunit (Gangnam-gu, Seoul, Korea); V.R.: research grants, Leica Biosystems (Wetzlar, Germany), Philips (Amsterdam, Netherlands), 3DHISTECH Ltd (Budapest, Hungary), and Paige (New York, New York). It would be challenging, if not impossible, to assemble a panel with the level of expertise required to write this guideline where no member has COIs related to WSI. The chair was required to have no COIs related to the subject matter of this guideline. The majority of the expert panel members (7 of 12) were assessed as having no relevant COIs. The CAP provided funding for the administration of the project; no industry funds were used in the development of the guideline. All panel members volunteered their time and were not compensated for their involvement, except for the contracted methodologist. See the SDC for complete information about the COI policy.

## GUIDELINE OBJECTIVE

As per the 2013 guideline, the expert panel addressed the overarching question, “What needs to be done to validate (as defined above) a WSI digital pathology system for

diagnostic purposes before it is placed in clinical service?" The panel considered intended uses of WSI, preparation types, numbers of cases, systems/components, personnel, and processes.

## LITERATURE SEARCH AND COLLECTION

A comprehensive literature search for relevant evidence was completed by the CAP's medical librarian using Ovid MEDLINE and Elsevier Embase on June 26, 2018, encompassing the publication dates of January 1, 2012, through June 26, 2018, and supplemental searches were completed using the Cochrane Library. The search strategy used controlled vocabulary (ie, MeSH, Emtree) and keywords derived from the key questions. Database searches were supplemented with a search for unindexed literature, including a review of clinical trials and pertinent organizations' Web sites. Expert panel members were also polled for relevant unpublished data at the onset of the project. The literature searches were rerun on June 14, 2019, and July 15, 2020, to identify articles published from June 26, 2018, through July 15, 2020. Detailed information regarding the literature search is available in the SDC, including the search terms used.

## INCLUSION CRITERIA

Studies were selected for inclusion in the systematic review of evidence if they met the following criteria: (1) the study referred to WSI; (2) the study pertained to clinical use or investigative research; and (3) the study was a peer-reviewed, full-text article. Detailed information about the inclusion criteria is available in the SDC.

## EXCLUSION CRITERIA

Articles were excluded from the systematic review if they were conference abstracts that were not published in peer-reviewed journals, qualitative studies, follow-up studies, mixed-methods studies, editorials, commentaries, narrative reviews, case reports, or letters; studies including fewer than 30 cases per study arm; studies in animal models or cell lines; full-text articles that were not available in English; studies that discussed only cytology cases; studies involving static and robotic digital imaging, purely technical components, only educational applications, or image analysis; or studies that did not address WSI validation. Despite literature that is now beginning to accrue on cytopathology as a use case for WSI, the evidence was considered to be immature relative to that for surgical pathology. As such, validation of WSI for cytopathology was considered out of scope for the guideline update pending additional research. Similar issues applied for peripheral blood smears and bone marrow aspirates, which were likewise out of scope. The guideline focuses on the use of WSI platforms by pathologists to make visual interpretations from images. It does not include recommendations on the use of automated image analysis systems. Detailed information about the exclusion criteria is available in the SDC.

## QUALITY ASSESSMENT

The research methodologist performed a risk of bias assessment for all fully published studies meeting inclusion criteria. The methodologist assessed key indicators based on study design and methodologic rigor; a rating for quality of evidence (Supplemental Table 1) was designated. An overall

GRADE rating was given for each recommendation by outcome. Refer to the SDC for further details.

## ASSESSING THE STRENGTH OF RECOMMENDATIONS

Following the quality of evidence assessment, completion of the GRADE Evidence to Decision Framework,<sup>12</sup> and discussion of the definitions and implications of strength of recommendations (Table 1), the expert panel designated the recommendations as either strong or conditional.

## GUIDELINE REVISION

As per the initial guideline, this revision will be reviewed in 4 years, or earlier in the event of publication of substantive and high-quality evidence that could potentially alter the original guideline recommendations. If necessary, the entire expert panel will reconvene to discuss potential changes. When appropriate, the panel will recommend revision of the guideline to the CAP and the Association for Pathology Informatics and American Society for Clinical Pathology collaborators for review and approval.

## DISCLAIMER

The Center was developed by the CAP as a forum to create and maintain laboratory practice guidelines. Guidelines are intended to assist physicians and patients in clinical decision-making and to identify questions and settings for further research. With the rapid flow of scientific information, new evidence may emerge between the time a laboratory practice guideline is developed and when it is published or read. Laboratory practice guidelines are not continually updated and may not reflect the most recent evidence. Laboratory practice guidelines address only the topics specifically identified therein and are not applicable to other interventions, diseases, or stages of diseases. Furthermore, guidelines cannot account for individual variation among patients and cannot be considered inclusive of all proper methods of care or exclusive of other treatments. It is the responsibility of the treating physician or other health care provider, relying on independent experience and knowledge, to determine the best course of treatment for the patient. Accordingly, adherence to any laboratory practice guideline is voluntary, with the ultimate determination regarding its application to be made by the physician in light of each patient's individual circumstances and preferences. The CAP and its collaborators make no warranty, express or implied, regarding laboratory practice guidelines and specifically exclude any warranties of merchantability and fitness for a particular use or purpose. The CAP and its collaborators assume no responsibility for any injury or damage to persons or property arising out of or related to any use of this statement or for any errors or omissions.

## RESULTS

A total of 1827 studies met the search term requirements and were carried forward for title and abstract review. Based on review of these abstracts, 173 articles met the inclusion criteria and were selected for full-text review. A total of 62 articles were included for data extraction. Each study was reviewed by 2 expert panel members at each phase. Studies with discordant reviews were referred to the chair for a final decision on inclusion or exclusion. Excluded articles were available as discussion or background references. Additional information about the systematic review is available in the



**Table 1. Definitions for Strength of Recommendation**

Category	Definition	Implication
Strong recommendation	One for which a guideline panel is confident that the desirable effects of an intervention outweigh its undesirable effects (strong recommendation for an intervention) or that the undesirable effects of an intervention outweigh its desirable effects (strong recommendation against an intervention)	Most individuals in this situation would want the recommended course of action and only a small proportion would not
Conditional (weak) recommendation	One for which the desirable effects probably outweigh the undesirable effects (weak recommendation for an intervention) or undesirable effects probably outweigh the desirable effects (weak recommendation against an intervention) but appreciable uncertainty exists	The majority of individuals in this situation would want the suggested course of action, but many would not

SDC, including a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) table outlining details of the systematic review. Refer to the write-up for each recommendation for specific details about supporting evidence.

The panel convened 12 times (10 teleconferences and 2 face-to-face meetings) to develop the scope, draft recommendations, review and respond to solicited feedback, and assess the quality of evidence that supports the final recommendations. A nominal group technique was used for consensus decision-making to encourage unique input with balanced participation among group members. An open comment period was posted on the CAP Web site from June 24, 2019, to July 15, 2019, during which the 3 draft recommendations and 9 GPSs were posted for public feedback. A total of 146 comments were submitted from 154 participants, with all draft recommendation and GPSs receiving at least 90% outright agreement or agreement with some modification. Refer to the SDC for more details. The expert panel approved the final recommendations with a supermajority vote.

An independent review panel, masked to the expert panel and vetted through the COI process, recommended approval by the CAP Council on Scientific Affairs. The final recommendations are summarized in Table 2.

## RECOMMENDATIONS

**1. Strong Recommendation.**—The validation process should include a sample set of at least 60 cases for one application, or use case (eg, hematoxylin-eosin–stained sections of fixed tissue, frozen sections, hematology), that reflect the spectrum and complexity of specimen types and diagnoses likely to be encountered during routine practice. Note: the validation process should include another 20 cases to cover additional applications such as immunohistochemistry or other special stains if these applications are relevant to an intended use and were not included in the 60 cases mentioned above.

This recommendation was reaffirmed from the 2013 guideline.

The quality of evidence is moderate to support this recommendation.

The evidence base supporting this recommendation comprised 32 diagnostic studies<sup>13–44</sup> and 1 systematic review.<sup>45</sup> Data collection occurred prospectively in 6 studies<sup>15–17,20,26,41</sup> and retrospectively in 18.\* One study<sup>33</sup> had both prospective data collection at 1 site and

retrospective data collection using archived specimens at 2 other sites. Seven noninferiority studies<sup>18,22–24,29,32,34</sup> were also included. The GRADE rating for concordance was strong recommendation, moderate quality evidence. Refer to Supplemental Tables 2 through 5 in the SDC for the individual study quality assessment, GRADE certainty rating, and evidence to decision summary for recommendation 1.

If the primary aim of validation is to demonstrate that a WSI system will perform as expected for a specific application such as primary diagnosis, it follows that the process should evaluate cases that reasonably represent the spectrum and proportion of diagnoses to be encountered in that application. The number of cases being evaluated should be sufficient to allow pathologists to establish trust in diagnoses made using WSI as well as to identify and mitigate risks associated with the technology. At the same time, the number of cases must strike a balance in terms of the amounts of time and resources required to complete the validation process. It is neither practical nor possible to evaluate a predetermined number of cases for every conceivable diagnosis during a validation study. To this end, the 2018 Royal College of Pathologists best practice recommendations for implementing digital pathology state that the sample size and duration of the validation process can vary according to specific circumstances and do not mention a particular number of cases to be evaluated.<sup>46</sup>

The number of cases recommended in the 2013 guideline and reaffirmed in this update is not intended to be a rigid number that laboratories must follow when validating WSI systems. It is logical that laboratories embarking on WSI validation will want some evidence-based guidance on a minimum number of cases they need to evaluate. The recommendation of at least 60 cases was determined from a systematic review of published validation studies showing that concordance between WSI and glass slide diagnoses is not improved or worsened when sets of more than 60 cases are used (Table 3; Figure). Systematic review of literature published between 2012 and 2018 (eg, after the 2013 guideline had been written) revealed 33 studies with case numbers ranging from 40 to 8069 where the reported endpoint was either intraobserver concordance<sup>†</sup> or discordance (from which concordance could be derived) between WSI and glass slides.<sup>‡</sup> The weighted mean concordance, directly determined or derived, from these 33 studies was 95.2%, with a median of 95.0% and an interquartile range of

\* References 13, 14, 19, 21, 25, 27, 28, 30, 31, 35–40, 42–44.

<sup>†</sup> References 15–17, 19, 21, 25, 26, 28, 29, 31, 33–37, 40–44.

<sup>‡</sup> References 13, 14, 18, 20, 22–24, 27, 30, 32, 38, 39, 45.

**Table 2. Summary of the Recommendations**

Recommendation	Strength of Recommendation
1. The validation process should include a sample set of at least 60 cases for one application, or use case (eg, hematoxylin-eosin-stained sections of fixed tissue, frozen sections, hematology), that reflect the spectrum and complexity of specimen types and diagnoses likely to be encountered during routine practice. The validation should include another 20 cases to cover additional applications such as immunohistochemistry or other special stains if these applications are relevant to an intended use and were not included in the 60 cases mentioned above	Strong
2. The validation study should establish diagnostic concordance between digital and glass slides for the same observer (ie, intraobserver variability). If concordance is less than 95%, laboratories should investigate and attempt to remedy the cause	Strong
3. A washout period of at least 2 weeks should occur between viewing digital and glass slides	Strong

91.0% to 97.1%. On breaking these studies down into subgroups by number of cases, there were only 2 studies where the mean number of cases was less than 60.<sup>20,21</sup> The mean number of cases for these studies was 41 and the mean WSI to glass slide intraobserver concordance was 85% (82.8% and 86.7%, respectively). Studies with means of 64 and 231 cases each yielded an identical mean concordance of 93%. Likewise, those with a mean of 750 cases also yielded a mean concordance of 93%. As such, no new evidence emerged to support changing the 2013 recommendation of at least 60 cases. More studies using fewer than 60 cases would be required to refine the above recommendation concerning a minimum number of cases.

This recommendation received 111 responses during the open comment period, of which 94.6% (n = 105) either agreed with the recommendation as written or agreed with suggested modifications. There were 22 comments. Some comments appeared to suggest that the number of cases was determined in an arbitrary manner, which, as described above, is not the case. Some comments suggested that 60 cases represented too few cases to adequately cover the complexity of cases encountered in clinical practice, whereas others suggested that the number was too large and impractical. The question of how one defines a “case” was also raised. Does a case include all parts and slides associated with a given accession number in a laboratory information system or can representative parts and/or slides be used? The PIVOTAL trial conducted by Philips to obtain Food and Drug Administration approval did not use all parts and slides for each of the 1992 cases that were evaluated.<sup>22</sup> As mentioned above, the issue becomes one of balance between completeness and practicality with respect to the time and resources. It is the opinion of the expert panel that laboratories should be free to decide on how to define a case, with the proviso being that the selected material allows pathologists to conduct a reasonable and thorough assessment of the WSI system prior to introducing it into patient care. Finally, the issue of validating multiple WSI scanners distributed over a multisite network was raised and whether

each scanner requires its own validation set of 60 cases. Laboratories may use their own judgment to determine whether the applications or use cases (eg, frozen sections, consultations, quality assurance or primary diagnosis) and expected case mix between sites are sufficiently different to warrant separate validations. Because the validation exercise is also intended to identify issues related to histology and their impact on image quality, assessing differences in the quality of histology between sites represents a critical consideration when deciding whether separate validations are required at each site. If WSI scanners from different vendors with different proprietary viewing software are being used across the network, it is reasonable to recommend that separate validation studies be performed for each system.

**2. Strong Recommendation.**—The validation study should establish diagnostic concordance between digital and glass slides for the same observer (ie, intraobserver variability). If concordance is less than 95%, laboratories should investigate and attempt to remedy the cause.

This recommendation was updated from the 2013 guideline.

The quality of evidence is moderate to support this recommendation.

The evidence base supporting this recommendation comprised 32 diagnostic accuracy studies<sup>13–44</sup> and 1 systematic review.<sup>45</sup> Data collection occurred prospectively in 6 studies<sup>15–17,20,26,41</sup> and retrospectively in 18.<sup>5</sup> One study had both prospective data collection at 1 site and retrospective data collection using archived specimens at 2 other sites.<sup>33</sup> Seven noninferiority studies<sup>18,22–24,29,32,34</sup> were also included. Noninferiority designs compare diagnoses made by glass slides and WSI during the validation process to ground truth diagnoses made with glass slides as part of patient care. This approach provides a measure of diagnostic correctness, assuming the ground truth is correct. The original diagnoses may have been made by pathologists other than those completing the validation, providing information on inter-observer variability. Comparing original and validation diagnoses made by glass slides provides baseline diagnostic concordance when cases are reviewed by the same modality. Validation diagnoses made by WSI are compared with original and validation glass slide diagnoses, the aim being to determine whether WSI creates additional discordance on top of that seen with glass slides alone. Based on a statistically determined margin, WSI is assessed as being inferior or noninferior to glass slides.

**Table 3. Concordance of Whole Slide Imaging and Glass Slides by Number of Cases**

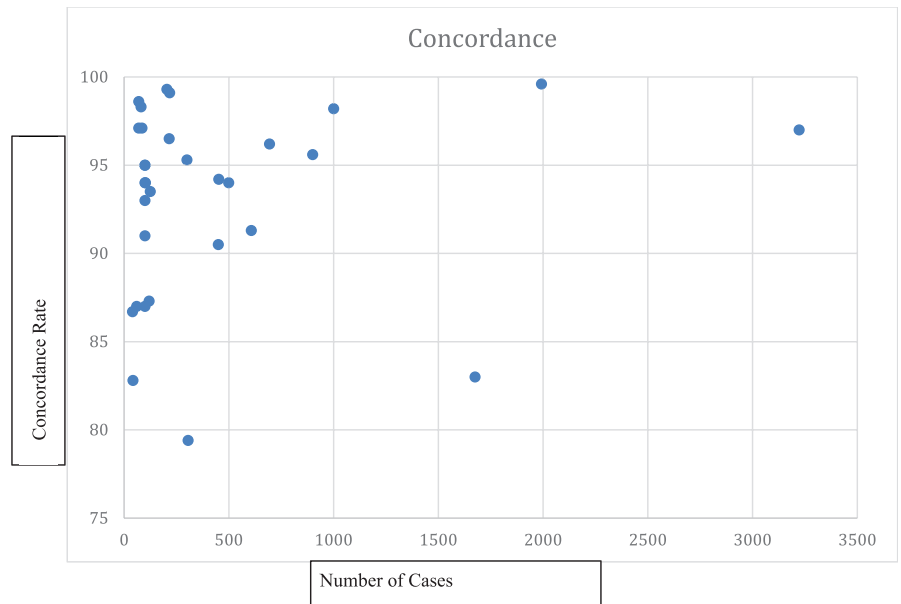
	Average No. of Cases <sup>a</sup>			
	<60	60	200	750
Average no. cases within group	41	64	231	767
Mean concordance, % <sup>b</sup>	85	93	93	93

<sup>a</sup> Approximate average number of cases.

<sup>b</sup> References 13–45.

<sup>5</sup> References 13, 14, 19, 21, 25, 27, 28, 30, 31, 35–40, 42–44.

A scatter plot of the total number of cases per study in comparison with the rate of diagnostic concordance. Concordance rates, as reported in each study or as calculated from reported discordance rates when concordance rates were not available, are plotted.



The GRADE rating for diagnostic concordance was strong recommendation, moderate quality evidence. Refer to Supplemental Tables 6 through 9 in the SDC for the individual study quality assessment, GRADE certainty rating, and evidence to decision summary for recommendation 2.

As outlined in the 2013 guideline, validation is a process designed to demonstrate that new technology or instrumentation performs as expected for its intended use prior to its application for patient care. This recommendation refers to the initial validation of the hardware and software components of a WSI system that must function correctly for pathologists to render diagnostic interpretations. The central question to be addressed is whether the same pathologist makes the same interpretation of a given case regardless of whether it is reviewed by WSI or as glass slides. Concordance, defined in this context as intraobserver agreement between diagnoses made by glass slides and WSI, is an ideal end point. Discordance, defined as disagreement between diagnoses, may also be used. Discrepant diagnoses between modalities can be subclassified as major or minor, where major discrepancies are defined as those that would impact patient management. Arbitration of discrepant diagnoses can be done manually by individual pathologists or panels of pathologists,<sup>22</sup> or more objectively via standardized discordance tables.<sup>47</sup> The intent of this recommendation is to evaluate the performance characteristics of a new technology being introduced into a diagnostic environment and its influence on the ability of pathologists to make diagnoses relative to glass slides. The process is not intended to assess diagnostic correctness or to validate an individual pathologist's diagnostic competency. An interobserver study design compares the interpretations of 2 pathologists each reviewing the same case/slide by different modalities. This approach does not optimally examine the influence of WSI on discordant diagnoses that could represent differences of opinion independent of the modality by which the slides were reviewed. Further, validation of a WSI system should be carried out by pathologists who have already been trained and are technically competent in the use of the

specific WSI system being implemented (see below for discussion on GPS 6).

Laboratories engaged in validating WSI for patient care will logically seek advice concerning an acceptable pass/fail point for concordance/discordance with diagnoses made by traditional glass slide review. The ideal pass/fail number would be 100% concordance (or 0% discordance); however, this does not reflect the subjective nature of pathology as practiced with glass slides where interobserver and intraobserver variability is an established reality. The weighted mean percentage concordance across the 33 studies in our systematic review was 95.2%. This analysis formed the basis for the 95% concordance mark in this recommendation. Discordance between WSI and glass slides was reported in 24 studies, of which 5 studies<sup>20,23,39,40,48</sup> classified the discordance as minor (average rate of 6.7%) and 7 studies<sup>18,20,23,25,29,32,36</sup> classified the discordance as major (average rate of 4.2%). Eleven studies reported specifically on the concordance and minor discordance rate between WSI and glass slides, with minor discordance ranging from 1.4% to 10.1% in 7 studies,<sup>15,18,25,29,30,32,36</sup> with an average minor discordance rate of 5.1%. It should be noted that validation studies using a noninferiority design have shown no significant difference in major discordance rates (ie, discrepant diagnoses that would affect patient management) between WSI and glass slides.<sup>18,22,24,29,32,34</sup>

This recommendation received 111 responses during the open comment period, with 94% (n = 104) agreeing with the recommendation either as written or with suggested modifications. Comments on the minimum concordance value suggested that 95% was too low, or conversely a discordance rate of 5% was unacceptably high, especially if the discordances are major and affect patient management. One comment suggested the need to capture rates of deferral of WSI cases to glass slide review during validation studies. Additional information suggested included the types of cases involved, the reasons for deferral, and whether deferral behavior is shown by all pathologists completing the validation or only a few pathologists who might be inherently reluctant to use WSI for diagnostic purposes. Capturing rates of deferral to glass slide review

**Table 4. Concordance of Whole Slide Imaging (WSI) and Glass Slides by Washout Period**

	Washout Period for WSI and Glass Slides, wk		
	<2	2-8	>8
Mean concordance between WSI and glass slides, % <sup>a</sup>	Not reported	93.2	93.3

<sup>a</sup> References 15–17, 19, 26, 28, 30, 31, 33, 34, 36, 40, 48, 50.

was a recommended quality management metric in the 2014 American Telemedicine Association clinical guidelines for telepathology.<sup>49</sup>

The 95% figure is not intended to be a pass/fail mark, where less than 95% means validation has failed and WSI should not be used for patient care. Achieving less than 95% concordance merely suggests the results are below average based on the peer-reviewed literature used to develop this guideline. As such, the panel's recommendation is that laboratories investigate and attempt to resolve systematic issues that may have contributed to a concordance rate of less than 95%. Possible areas to examine include, but are not limited to, the types of cases that were found to be problematic, whether the discordances are attributable to only 1 or 2 pathologists or a larger group of reviewers, and whether the discordant cases are related to correctable histology and/or scanner issues. One approach to exploring discordances for specific diagnoses that arose in an initial validation set would be to review additional cases of that type to better understand the issues involved. If factors contributing to discordant WSI diagnoses cannot be rectified to the satisfaction of the pathologists using the WSI system, those cases could routinely be deferred to glass slide review. Problematic specimen types could then become the focus of additional research as WSI technology evolves.

**3. Strong Recommendation.**—A washout period of at least 2 weeks should occur between viewing digital and glass slides.

This recommendation was reaffirmed from the 2013 guideline.

The quality of evidence is moderate to support this recommendation.

The evidence base supporting this recommendation comprised 14 studies.<sup>||</sup> Data collection occurred prospectively in 4 studies<sup>15–17,26</sup> and retrospectively in 8.<sup>¶</sup> One study had both prospective data collection at 1 site and retrospective data collection using archived specimens at 2 other sites.<sup>33</sup> A noninferiority study<sup>34</sup> was also included. Refer to Supplemental Tables 10 through 12 in the SDC for the individual study quality assessment, GRADE certainty rating, and evidence to decision summary for recommendation 3.

This recommendation is intended to address the issue of recall bias when cases are reviewed by 2 different modalities by the same observer. Most, if not all, pathologists can identify cases that they describe as “once seen, never forgotten.” Nonetheless, there is merit in attempting to control for bias that may lead to less than thorough review and/or rubber-stamping of diagnoses previously made by one modality when cases are reviewed by the second modality. Literature specifically designed to identify an optimal washout period for WSI validation studies is nonexistent, although studies based on comparing glass

slide diagnoses made on the same cases at different time points have been performed. Campbell et al<sup>51</sup> reported the influence of a 2- versus 4-week washout period on diagnoses made by 3 pathologists. The pathologists each reviewed groups of 60 glass slides from a set of 120. Following intervals of 2 and 4 weeks, the pathologists reviewed a second set of 60 slides that included a mix of cases seen or not seen in the initial review. The participants were asked to identify cases they had seen during the first review and were asked to rate their level of confidence with respect to recalling those cases. The pathologists correctly recalled 40% of the cases after 2 weeks and 31% of the cases after 4 weeks, indicating an appreciable degree of recall that could influence data collected during intraobserver studies.

A total of 14 studies (see Table 4) with respect to the influence of the length of a washout period on intraobserver concordance between glass slide and WSI diagnoses were included. The washout period ranged from less than 4 weeks to greater than 8 weeks, with 5 (35.7%) reporting a washout of 1 to 4 weeks and 9 (64.3%) using a washout of more than 8 weeks. No influence was found when concordance data from these studies were stratified according to washout duration. As such, no new evidence was identified on systematic review to support changing the washout period of at least 2 weeks recommended in the 2013 guideline.

This recommendation received 111 responses during the open comment period, 90% (n = 100) of which agreed with the recommendation as written or agreed with comments concerning modifications. Suggestions for washout periods other than the recommended 2-week minimum ranged from not using a washout period to increasing it to a minimum of 4 weeks. Some contributors believed that 2 weeks was too long and would create a barrier to using the guideline. Others believed 2 weeks was too short and would increase the likelihood of recall bias, although this concern is not supported by our systematic review. Those believing that a washout period was unnecessary also indicated that washout periods could remove the opportunity to identify cases where all tissues on a glass slide may not have been captured during the scanning process. Although failure of WSI scanners to capture all tissue is a rare occurrence,<sup>52</sup> it has been documented and could have potentially devastating consequences on patient care with medicolegal repercussions.

As per the other recommendations in this guideline, the 2-week washout period serves as guidance for laboratories seeking evidence-based advice on this specific aspect of WSI validation. Laboratories are free to use washout periods of any duration they might deem more practical or better able to minimize the negative impact of recall bias on data being collected to inform the use of their WSI system for diagnostic work. One solution to the possibility of not detecting incomplete tissue capture by a WSI scanner would be to build in a technical quality control step whereby all scanned slides used in the validation study are reviewed by a

<sup>||</sup> References 15–17, 19, 26, 28, 30, 31, 33, 34, 36, 40, 48, 50.

<sup>¶</sup> References 19, 28, 30, 31, 36, 40, 48, 50.



scanning technologist in conjunction with the corresponding glass slides to ensure all tissue was captured. Detection of any such events should trigger communication with the WSI vendor to begin work on correcting the problem. Image analysis algorithms based on artificial intelligence may emerge as a preanalytical solution for detecting missing tissue on scanned slides before cases are assigned to pathologists for review.<sup>53</sup>

### GOOD PRACTICE STATEMENTS

As mentioned in the introduction, GPSs are defined as statements having a “high level of certainty that the recommendation will do more good than harm (or the reverse), but where there is little direct evidence.”<sup>54,55</sup> Unlike recommendations, GPSs are not evidence based. Nine of the recommendation statements from the 2013 guideline were reaffirmed and/or revised as GPSs.

**GPS 1.** All pathology laboratories implementing WSI technology for clinical diagnostic purposes should carry out their own validation studies.

**GPS 2.** Validation should be appropriate for and applicable to the intended clinical use and clinical setting of the application in which WSI will be used. Validation of WSI systems should involve specimen preparation types relevant to intended use (eg, formalin-fixed, paraffin-embedded tissue; frozen tissue; immunohistochemical stains). If a new application for WSI is contemplated, and it differs materially from the previously validated use, a separate validation for the new application should be performed.

**GPS 3.** The validation study should closely emulate the real-world clinical environment in which the technology will be used.

**GPS 4.** The validation study should encompass the entire WSI system. It is not necessary to separately validate each individual component (eg, computer hardware, monitor, network, scanner) of the system or the individual steps of the digital imaging process.

**GPS 5.** Laboratories should have procedures in place to address changes to the WSI system that could impact clinical results. This statement was revised from the 2013 guideline.

**GPS 6.** Pathologists adequately trained to use the WSI system must be involved in the validation process.

**GPS 7.** The validation process should confirm all of the material present on a glass slide to be scanned is included in the digital image.

**GPS 8.** Documentation should be maintained recording the method, measurements, and final approval of validation for the WSI system to be used in the anatomic pathology laboratory.

**GPS 9.** Pathologists should review cases/slides in a validation set in random order. This applies to both the review modality (ie, glass slides or digital) and the order in which slides/cases are reviewed within each modality. This statement was revised from the 2013 guideline.

Responses during the open comment period indicated an average of 98% (range, 94%–100%) agreement with the GPSs as written or agreement with some modification. Several of the GPSs were revised from the 2013 guideline by the expert panel or were the subject of specific questions during the open comment period. Good practice statement 2, concerning validation of WSI for specific intended uses, states that separate validation is required for new applica-

tions that differ materially from the application for which the initial validation was performed. Examples include the following: if the initial validation was performed for primary diagnosis involving hematoxylin-eosin, histochemically and immunohistochemically stained paraffin sections, a separate validation is required if frozen section interpretation is the new application. Similarly, if the initial validation did not include biopsies requiring review at high resolution (eg, gastric biopsies with *Helicobacter pylori*), it would be prudent to conduct a separate validation with such cases, as they may represent limitations on the use of WSI.<sup>24,27</sup> Good practice statement 3 covers the concept of validation studies emulating the real-world environment in which the technology is to be used. It is recognized that validation studies cannot perfectly replicate real-life diagnostic activities. Laboratories are free to incorporate whatever they feel would be appropriate to achieve this goal, including clinical information, serial/deeper hematoxylin-eosin levels, immunohistochemical stains, or ancillary test data. Good practice statement 4, concerning the validation of the entire WSI system as opposed to individual components, received comments on how to deal with situations where new monitors/displays were introduced after validation had been completed. These points also relate to GPS 5 and the need for procedures to address changes to the WSI system that could impact clinical results. (The panel developed a resource to help laboratories determine the actions needed for various changes to the WSI system. It can be accessed on the CAP's WSI guideline Web page: <https://www.cap.org/protocols-and-guidelines/cap-guidelines/current-cap-guidelines/validating-whole-slide-imaging-for-diagnostic-purposes-in-pathology>.) Good practice statement 6, concerning the need for pathologists to be adequately trained on the use of the WSI system prior to embarking on a validation study, received requests for a definition of *adequately trained*. Following systematic review, no evidence-based recommendations could be made about the type of training or the metrics used to determine technical competency of pathologists using WSI systems for diagnostic purposes. Further, expert panel members noted that some studies did not report whether user training had been provided. As such, adequate training is defined at the discretion of the laboratory medical director. The same applies for the number of pathologists participating in the validation process. Having all pathologists in a given institution complete a validation study certainly has its benefits; however, this may not be practical, and there is no evidence indicating that it is necessary. The expert panel will provide suggestions for training in a document made available on the CAP Web site following release of the guideline. Good practice statement 9, concerning order of review of cases within and between the glass slide and WSI arms of a validation study, was discussed in detail by the expert panel. Although some studies in the systematic review of WSI validation literature indicate that such randomization of review order was performed,<sup>22</sup> there is no specific evidence that changing the order actually influences the data collected. Even further, although it may be intuitive that randomizing the order of cases between modalities will minimize recall bias, no studies were found in the pathology literature where random versus nonrandom review order was specifically compared in terms of intraobserver variability.

See Table 5 for a summary of the GPSs.

**Table 5. Summary of Good Practice Statements (GPSs)**

<p><b>GPS 1.</b> All pathology laboratories implementing WSI technology for clinical diagnostic purposes should carry out their own validation studies</p> <p><b>GPS 2.</b> Validation should be appropriate for and applicable to the intended clinical use and clinical setting of the application in which WSI will be used. Validation of WSI systems should involve specimen preparation types relevant to intended use (eg, formalin-fixed, paraffin-embedded tissue; frozen tissue; immunohistochemical stains). If a new application for WSI is contemplated, and it differs materially from the previously validated use, a separate validation for the new application should be performed</p> <p><b>GPS 3.</b> The validation study should closely emulate the real-world clinical environment in which the technology will be used</p> <p><b>GPS 4.</b> The validation study should encompass the entire WSI system. It is not necessary to separately validate each individual component (eg, computer hardware, monitor, network, scanner) of the system or the individual steps of the digital imaging process</p> <p><b>GPS 5.</b> Laboratories should have procedures in place to address changes to the WSI system that could impact clinical results. This statement was revised from the 2013 guideline</p> <p><b>GPS 6.</b> Pathologists adequately trained to use the WSI system must be involved in the validation process</p> <p><b>GPS 7.</b> The validation process should confirm all of the material present on a glass slide to be scanned is included in the digital image</p> <p><b>GPS 8.</b> Documentation should be maintained recording the method, measurements, and final approval of validation for the WSI system to be used in the anatomic pathology laboratory</p> <p><b>GPS 9.</b> Pathologists should review cases/slides in a validation set in random order. This applies to both the review modality (ie, glass slides or digital) and the order in which slides/cases are reviewed within each modality. This statement was revised from the 2013 guideline</p>
---

Abbreviation: WSI, whole slide imaging.

## DISCUSSION

This guideline provides evidence-based recommendations for pathologists and laboratories implementing WSI for diagnostic purposes including frozen sections, consultation, quality assurance reviews, and primary diagnosis. The term *diagnostic purposes* in the context of WSI refers to an activity where a pathologist interprets a scanned slide and that interpretation either contributes to or is the sole basis for information in a pathology report that is used for patient management. Although WSI is frequently used to review cases at tumor boards/multidisciplinary case conferences, tumor boards do not constitute diagnostic activities in and of themselves, even though they may lead to slide reviews that result in amendments to original pathology reports. As such, WSI systems do not require validation before they are used for tumor boards.

As with the initial version released in 2013, this update was generated by an expert panel that considered the necessary steps involved in validating these systems for diagnostic use. A systematic review of literature that appeared after the release of the 2013 guideline was conducted to identify and grade evidence that might inform each step in the validation process. In developing the final recommendations and GPSs in this guideline, professional bodies such as the CAP need to anticipate questions that will arise as laboratories embark on WSI validation. To place this guideline in proper context, several points must be acknowledged. Guidelines are not intended to be standards of care. Several other guidelines concerning the implementation of WSI for clinical use exist.<sup>46</sup> Laboratories are free to deviate from published guidelines where the advice is either impractical or not applicable to their situation. Finally, the panel recognizes that the recommendations offered in this manuscript may be more stringent than what is required by the CAP Laboratory Accreditation Program. For accreditation purposes, laboratories should follow the requirements of their accreditation agencies.

## CONCLUSIONS

The revised guideline contains 3 strong recommendations and 9 GPSs, all of which were included in the original 2013 guideline. Importantly, systematic review of literature following release of the original guideline reaffirms strong recommendations concerning the use of a validation set

comprising at least 60 cases, establishing diagnostic concordance between WSI and glass slides for the same observer, and the use of a 2-week washout period. The mean diagnostic concordance between WSI and glass slides reported by studies assessed in our systematic review was 95.2%. Although all discordances between WSI and glass slide diagnoses discovered during validation studies need to be reconciled, laboratories should undertake a systematic review of their data if the overall WSI–glass slide concordance is less than 95%.

We thank the American Society for Clinical Pathology, the Association for Pathology Informatics, and their staff for their involvement and support. We also thank the advisory panel members for their guidance throughout the development of the guideline and for their thoughtful review of this work: Walter H. Henricks, MD; Jason Hipp, MD, PhD; Dennis O'Neill, MD; David McClintock, MD; Paul J. van Diest, MD, PhD; Chee Leong Cheng, MBBS; and Veronica Klepeis, MD, PhD.

## References

- 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(12):1710–1722.
- Hillier S, Grimmer-Somers K, Merlin T, et al. FORM: an Australian method for formulating and grading recommendations in evidence-based clinical guidelines. *BMC Med Res Methodol.* 2011;11:23.
- RELEX Group plc. Scopus database. Elsevier BV; 2021. <https://www.scopus.com/>. Accessed January 12, 2021.
- US Food and Drug Administration. FDA allows marketing of first whole slide imaging system for digital pathology [news release]. <https://www.fda.gov/news-events/press-announcements/fda-allows-marketing-first-whole-slide-imaging-system-digital-pathology>. Published April 12, 2017. Accessed June 18, 2018.
- Retamero JA, Aneiros-Fernandez J, Del Moral RG. Complete digital pathology for routine histopathology diagnosis in a multicenter hospital network. *Arch Pathol Lab Med.* 2020;144(2):221–228.
- Volynskaya Z, Chow H, Evans A, Wolff A, Lagmay-Traya C, Asa SL. Integrated pathology informatics enables high-quality personalized and precision medicine: digital pathology and beyond. *Arch Pathol Lab Med.* 2018;142(3):369–382.
- Fraggetta F, Garozzo S, Zannoni GF, Pantanowitz L, Rossi ED. Routine digital pathology workflow: the Catania experience. *J Pathol Inform.* 2017;8(1): 51.
- Stathonikos N, Nguyen TQ, Spoto CP, Verdaasdonk MAM, van Diest PJ. Being fully digital: perspective of a Dutch academic pathology laboratory. *Histopathology.* 2019;75(5):621–635.
- Thorstenson S, Molin J, Lundström C. Implementation of large-scale routine diagnostics using whole slide imaging in Sweden: digital pathology experiences 2006–2013. *J Pathol Inform.* 2014;5(1):14.
- Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1, introduction—GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol.* 2011; 64(4):383–394.

11. Graham R, Mancher M, Miller Wolman D, Greenfield S, Steinberg E, eds; Institute of Medicine Committee on Standards for Developing Trustworthy Clinical Practice Guidelines. *Clinical Practice Guidelines We Can Trust*. Washington, DC: National Academies Press; 2011.
12. Alonso-Coello P, Schunemann HJ, Moberg J, et al. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices, 1: introduction. *BMJ*. 2016;353:i2016.
13. Araujo ALD, Amaral-Silva GK, Fonseca FP, et al. Validation of digital microscopy in the histopathological diagnoses of oral diseases. *Virchows Arch*. 2018;473(3):321–327.
14. Rakha EA, Aleskandarani M, Toss MS, et al. Breast cancer histologic grading using digital microscopy: concordance and outcome association. *J Clin Pathol*. 2018;71(8):680–686.
15. Villa I, Mathieu MC, Bosq J, et al. Daily biopsy diagnosis in surgical pathology: concordance between light microscopy and whole-slide imaging in real-life conditions. *Am J Clin Pathol*. 2018;149(4):344–351.
16. Williams BJ, Hanby A, Millican-Slater R, Nijhawana A, Verghese E, Treanor D. Digital pathology for the primary diagnosis of breast histopathological specimens: an innovative validation and concordance study on digital pathology validation and training. *Histopathology*. 2018;72(4):662–671.
17. Tabata K, Mori I, Sasaki T, et al. Whole-slide imaging at primary pathological diagnosis: validation of whole-slide imaging-based primary pathological diagnosis at twelve Japanese academic institutes. *Pathol Int*. 2017;67(11):547–554.
18. Bauer TW, Slaw RJ, McKenney JK, Patil DT. Validation of whole slide imaging for frozen section diagnosis in surgical pathology. *J Pathol Inform*. 2015;6(1):49.
19. Al-Janabi S, Huisman A, Jonges GN, Ten Kate FJ, Goldschmeding R, van Diest PJ. Whole slide images for primary diagnostics of urinary system pathology: a feasibility study. *J Renal Inj Prev*. 2014;3(4):91–96.
20. Gui D, Cortina G, Naini B, et al. Diagnosis of dysplasia in upper gastrointestinal tract biopsies through digital microscopy. *J Pathol Inform*. 2012;3(1):27.
21. Fertig RM, Gaudi S, Cervantes J, et al. Feasibility study in teledermatopathology: an examination of the histopathologic features of mycosis fungoides and spongiotic dermatitis. *J Cutan Pathol*. 2017;44(11):919–924.
22. Mukhopadhyay S, Feldman MD, Abels E, et al. Whole slide imaging versus microscopy for primary diagnosis in surgical pathology: a multicenter blinded randomized noninferiority study of 1992 cases (pivotal study). *Am J Surg Pathol*. 2018;42(1):39–52.
23. Kent MN, Olsen TG, Feeser TA, et al. Diagnostic accuracy of virtual pathology vs traditional microscopy in a large dermatopathology study. *JAMA Dermatol*. 2017;153(12):1285–1291.
24. Snead DR, Tsang YW, Meskiri A, et al. Validation of digital pathology imaging for primary histopathological diagnosis. *Histopathology*. 2016;68(7):1063–1072.
25. Gomez-Gelvez JC, Kryvenko ON, Chabot-Richards DS, Foucar K, Inamdar KV, Karner KH. Comparative analysis reveals potential utility of digital microscopy in the evaluation of peripheral blood smears with some barriers to implementation. *Am J Clin Pathol*. 2015;144(1):68–77.
26. Loughrey MB, Kelly PJ, Houghton OP, et al. Digital slide viewing for primary reporting in gastrointestinal pathology: a validation study. *Virchows Arch*. 2015;467(2):137–144.
27. Thrall MJ, Wimmer JL, Schwartz MR. Validation of multiple whole slide imaging scanners based on the guideline from the College of American Pathologists Pathology and Laboratory Quality Center. *Arch Pathol Lab Med*. 2015;139(5):656–664.
28. Arnold MA, Chenever E, Baker PB, et al. The College of American Pathologists guidelines for whole slide imaging validation are feasible for pediatric pathology: a pediatric pathology practice experience. *Pediatr Dev Pathol*. 2015;18(2):109–116.
29. Ordi J, Castillo P, Saco A, et al. Validation of whole slide imaging in the primary diagnosis of gynaecological pathology in a university hospital. *J Clin Pathol*. 2015;68(1):33–39.
30. Houghton JP, Ervine AJ, Kenny SL, et al. Concordance between digital pathology and light microscopy in general surgical pathology: a pilot study of 100 cases. *J Clin Pathol*. 2014;67(12):1052–1055.
31. Campbell WS, Hinrichs SH, Lele SM, et al. Whole slide imaging diagnostic concordance with light microscopy for breast needle biopsies. *Hum Pathol*. 2014;45(8):1713–1721.
32. Bauer TW, Slaw RJ. Validating whole-slide imaging for consultation diagnoses in surgical pathology. *Arch Pathol Lab Med*. 2014;138(11):1459–1465.
33. 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(12):1733–1739.
34. Bauer TW, Schoenfeld L, Slaw RJ, Yeran L, Sun Z, Henricks WH. Validation of whole slide imaging for primary diagnosis in surgical pathology. *Arch Pathol Lab Med*. 2013;137(4):518–524.
35. Al-Janabi S, Huisman A, Nap M, Clarijs R, van Diest PJ. Whole slide images as a platform for initial diagnostics in histopathology in a medium-sized routine laboratory. *J Clin Pathol*. 2012;65(12):1107–1111.
36. Al-Janabi S, Huisman A, Willems SM, Van Diest PJ. Digital slide images for primary diagnostics in breast pathology: a feasibility study. *Hum Pathol*. 2012;43(12):2318–2325.
37. Campbell WS, Lele SM, West WW, Lazenby AJ, Smith LM, Hinrichs SH. Concordance between whole-slide imaging and light microscopy for routine surgical pathology. *Hum Pathol*. 2012;43(10):1739–1744.
38. Fonyad L, Krenacs T, Nagy P, et al. Validation of diagnostic accuracy using digital slides in routine histopathology. *Diagn Pathol*. 2012;7(1):35.
39. Al-Janabi S, Huisman A, Vink A, et al. Whole slide images for primary diagnostics in dermatopathology: a feasibility study. *J Clin Pathol*. 2012;65(2):152–158.
40. Al-Janabi S, Huisman A, Vink A, et al. Whole slide images for primary diagnostics of gastrointestinal tract pathology: a feasibility study. *Hum Pathol*. 2012;43(5):702–707.
41. Hanna MG, Reuter VE, Hameed MR, et al. Whole slide imaging equivalency and efficiency study: experience at a large academic center. *Mod Pathol*. 2019;32(7):916–928.
42. Larghi A, Fornelli A, Lega S, et al. Concordance, intra- and inter-observer agreements between light microscopy and whole slide imaging for samples acquired by EUS in pancreatic solid lesions. *Digest Liver Dis*. 2019;51(11):1574–1579.
43. Cima L, Brunelli M, Parwani A, et al. Validation of remote digital frozen sections for cancer and transplant intraoperative services. *J Pathol Inform*. 2018;9(1):34.
44. Sturm B, Creyten D, Cook MG, et al. Validation of whole-slide digitally imaged melanocytic lesions: does z-stack scanning improve diagnostic accuracy? *J Pathol Inform*. 2019;10(1):6.
45. Williams BJ, DaCosta P, Goacher E, Treanor D. A systematic analysis of discordant diagnoses in digital pathology compared with light microscopy. *Arch Pathol Lab Med*. 2017;141(12):1712–1718.
46. Cross S, Furness P, Igal L, Snead D, Treanor D. Best practice recommendations for implementing digital pathology. Royal College of Pathologists. <https://www.rcpath.org/uploads/assets/f465d1b3-797b-4297-b7fedc00b4d77e51/Best-practice-recommendations-for-implementing-digital-pathology.pdf>. Published January 2018. Accessed June 21, 2018.
47. Wack K, Drogowski L, Treloar M, et al. A multisite validation of whole slide imaging for primary diagnosis using standardized data collection and analysis. *J Pathol Inform*. 2016;7(1):49.
48. Malarkey DE, Willson GA, Willson CJ, et al. Utilizing whole slide images for pathology peer review and working groups. *Toxicol Pathol*. 2015;43(8):1149–1157.
49. Pantanowitz L, Krupinski EA. American Telemedicine Association 2014 meeting: what did you miss? *J Pathol Inform*. 2014;5(1):30.
50. Wilbur DC, Brachtel EF, Gilbertson JR, Jones NC, Vallone JG, Krishnamurthy S. Whole slide imaging for human epidermal growth factor receptor 2 immunohistochemistry interpretation: accuracy, precision, and reproducibility studies for digital manual and paired glass slide manual interpretation. *J Pathol Inform*. 2015;6(1):22.
51. Campbell WS, Talmon GA, Foster KW, Baker JJ, Smith LM, Hinrichs SH. Visual memory effects on intraoperator study design: determining a minimum time gap between case reviews to reduce recall bias. *Am J Clin Pathol*. 2015;143(3):412–418.
52. Fraggetta F, Yagi Y, Garcia-Rojo M, et al. The importance of eSlide macro images for primary diagnosis with whole slide imaging. *J Pathol Inform*. 2018;9(1):46.
53. Janowczyk A, Zuo R, Gilmore H, Feldman M, Madabhushi A. HistoQC: an open-source quality control tool for digital pathology slides. *JCO Clin Cancer Inform*. 2019;3(1):1–7.
54. Guyatt GH, Alonso-Coello P, Schunemann HJ, et al. Guideline panels should seldom make good practice statements: guidance from the GRADE Working Group. *J Clin Epidemiol*. 2016;80(12):3–7.
55. Tugwell P, Knottnerus JA. When does a good practice statement not justify an evidence based guideline? *J Clin Epidemiol*. 2015;68(5):477–479.



**APPENDIX. Disclosed Interests and Activities From May 2017 to September 2020<sup>a</sup>**

Name	Interest/Activity Type	Entity	
Marilyn M. Bui, MD, PhD	Board, advisory boards	DPA <i>Pathology and Laboratory Medicine International</i> <i>International Journal of Clinical and Experimental Pathology</i> <i>Journal of Pathology Informatics</i> <i>Journal of the American Society of Cytopathology</i> <i>Open Journal of Pediatrics</i> <i>OA Case Reports</i>	
	Consultancies	Aiforia AstraZeneca Bristol-Myers Squibb ContextVision Epizyme Hologic Visiopharm Rosen Louik & Perry, PC de la Parte & Gilbert, PA Josepher & Patteese, PA	
	Grants/collaborative agreements	NIH/NCI Bankhead Coley Florida Biomedical Research Program The GeM Consortium funded by Boston's Childrens' Hospital	
	Leadership in other associations	DPA American Society of Cytopathology Florida Society of Pathologists Chinese American Pathologists Association	
	Elected or appointed positions in other national/international medical organizations		
	Lecture fees (honoraria)	The Osler Institute	
	Royalties	Pathology Review, Osler Institute	
		CAP Press	
		Demos Medical	
		ASCP	
		DPA	
		National Society for Histotechnology	
		Colorado BioScience Association	
	Elizabeth A. Chlipala, BS, HTL(ASCP)QIHC	Lecture fees (honoraria)	ASCP
		Elected or appointed positions in other national/international medical organizations	DPA National Society for Histotechnology Colorado BioScience Association
Liron Pantanowitz, MD		Grants/collaborative agreements	Ibex Medical Analytics
			Huron
			Lunit
		Hamamatsu Photonics KK	
Anil V. Parwani, MD, PhD	Consultancies	Ibex Medical Analytics Leica Biosystems	
	Leadership in other associations	ASC	
Victor E. Reuter, MD	Board, advisory boards	ContextVision	
	Consultancies	Paige Cepheid	
Victor E. Reuter, MD	Grants/collaborative agreements	Leica Biosystems	
		Phillips	
		3DHistech Ventana	

Abbreviations: ASC, American Society of Cytopathology; ASCP, American Society for Clinical Pathology; CAP, College of American Pathologists; DPA, Digital Pathology Association; GeM, Genomics of MPNSTs; NCI, National Cancer Institute; NIH, National Institutes of Health.

<sup>a</sup> The following individuals have reported no relevant disclosures: Richard Brown, MD; Andrew J. Evans, MD, PhD; Christina Lacchetti, MSc; Danny A. Milner Jr, MD, MSc(Epi), MBA; Kearin Reid, MLIS, MT(ASCP); Michael Riben, MD; Lisa Stephens, MBA, HTLA(ASCP)<sup>CM</sup>; Rachel Stewart, DO, PhD; Nicole E. Thomas, MPH, CT(ASCP)<sup>CM</sup>.