Research Article

Pitfalls in the use of whole slide imaging for the diagnosis of central nervous system tumors: A pilot study in surgical neuropathology

Melike Pekmezci[1], Sanem Pinar Uysal[2], Yelda Orhan[2], Tarik Tihan[1], Han Sung Lee[1]

[1]Department of Pathology, Division of Neuropathology, University of California, San Francisco, CA, USA, [2]School of Medicine, Koc University, Istanbul, Turkey

E-mail: *Dr. Melike Pekmezci - melike.pekmezci@ucsf.edu
*Corresponding author

Abstract

Background: Whole slide imaging (WSI) finds increasingly higher value in everyday surgical pathology in addition to its well-established use for educational and research purposes. However, its diagnostic utility, especially in subspecialty settings such as neuropathology, is not fully validated. Neuropathology practice is unique with smaller overall tissue size and frequent need for high-power evaluation. In addition, tumor grade is an integral part of the initial diagnosis. The purpose of this study is to assess the feasibility of primary pathology diagnosis of surgical neuropathology specimens using WSI. **Materials and Methods:** We reviewed consecutive surgical neuropathology cases diagnosed in our institution during a 2-month period and identified a single diagnostic slide, which was scanned at 40× magnification. Two neuropathologists who were blinded to the original diagnoses reviewed the whole slide image and rendered a diagnosis including tumor grade when applicable. They reviewed the single diagnostic slide after a wash-out period. Intra- and inter-observer discrepancies, as well as reasons for discrepancies, were evaluated. **Results:** The concordance rates were 94.9% and 88% for two neuropathologists. Two critical issues leading to discrepancies were identified: (1) identification of mitoses and (2) recognition of nuclear details. **Conclusions:** Given the current study is exclusively for surgical neuropathology cases, an all-encompassing conclusion about the utility of WSI for diagnostic purposes may not be available. Nevertheless, pathologists should be aware of the potential pitfalls due to identification of mitotic figures and nuclear details. We recommend independent validation for each subspecialty of pathology to identify subspecialty-specific concerns, so they can be properly addressed.

Key words: Central nervous system tumors, digital pathology, neuropathology, whole slide imaging

INTRODUCTION

Whole slide imaging (WSI) is increasingly used in everyday surgical pathology practice as its technical and logistic challenges are being addressed with every new generation of scanners and software.[1,2] WSI is one of the imaging modalities, in which the entire tissue samples on glass slides are converted into digital images using one

of the available platforms and can be viewed at any time as if one is using a light microscope. The assessment of pathology specimens using WSI is appealing for remote consultations, quality assurance, and quantitative analyses. Furthermore, ease of data transfer, archiving, and use as an educational tool provide additional benefits over other virtual microscopy modalities such as robotic telepathology.[3] However, diagnostic accuracy of WSI will need to be validated in comparison to the current gold standard, i.e., traditional light microscopy.

There are numerous validation studies in the literature, investigating the use of WSI as a substitute for direct microscopic evaluation of glass slides for pathology diagnoses.[1-14] The designs of these studies show significant variations including sample size, case selection, diagnostic setting, type of imaging hardware and software, image resolution, and study outcomes. The significance of morphologic features such as growth pattern, nuclear pleomorphism, or mitotic activity could differ considerably among various tissues analyzed and this could result in different levels of concordance among WSI studies. The above mentioned studies on specific category of specimens or tumors have yielded variable results. The results and methods used for various lesions and pathology subspecialties also vary in the literature, further complicating the development of general guidelines applicable to every specimen.[4-6] There are limited number of studies focusing on the application of WSI in some subspecialties and even fewer evaluating the role of WSI as primary diagnosis in surgical neuropathology. Many earlier studies either did not include neuropathology specimens[7,8] or did not specifically addressed issues related to surgical neuropathology.[9,10] In one study, the only major discrepancy was recorded for neuropathology specimens, representing a 4.9-fold increase in major discrepancy rate for these cases.[11]

The studies using telepathology focus on intraoperative evaluation of neuropathology cases, mostly via robotic systems and simultaneous visualization.[12,13] These studies may not be directly comparable to studies with WSI and do not specifically address the issue of providing final diagnoses with the use of this technology.

Neuropathology is a subspecialty in which the overall tissue size is considerably smaller, cytologic smears are integral to intraoperative and final diagnosis, and evaluation at high power magnification is frequently required. For the majority of neoplastic cases, histological typing and grading are critical components of the diagnosis and a discrepancy in grade (i.e. WHO Grade II vs. IV) may have equally significant consequences in clinical management when compared to discrepancy in tumor types (i.e., pilocytic astrocytoma versus glioblastoma).

The purpose of this study is to assess the feasibility of primary pathology diagnosis of surgical neuropathology specimens using WSI. We analyzed discordant results to identify specific pitfalls and limitations of WSI as well as potential areas of improvement.

MATERIALS AND METHODS

We have identified all consecutive surgical neuropathology cases diagnosed in our institution during a 2-month period. We have included common primary central nervous system neoplasms that are often encountered in everyday clinical practice and may have diagnostic challenges. We have excluded pituitary adenomas, diagnoses related to degenerated disc diseases or other reactive lesions, nonprimary lesions such as metastatic carcinomas and melanomas, vascular malformations, and other benign or descriptive diagnoses such as meningocele, dermoid cyst, or focal cortical dysplasia. In addition, we have excluded the cases for which the slides were not available for WSI scanning.

The lead author (MP) served as the study coordinator and reviewed all available slides to select a single representative diagnostic slide for each case. All cases, except for one consultation case, were hematoxylin and eosin (H&E)-stained slides prepared in our laboratory with the same method. The remaining one case was evaluated for adequate staining quality by the lead author before inclusion in the study. The selected slides were scanned at a 40× magnification with ScanScope® XT (Aperio Technologies Inc., Vista, CA, USA). The files were uploaded to a cloud-based server (PathForce[dx] LLC, Seattle, WA, USA) and were reviewed by two neuropathologists (TT, HSL), who had not previously seen any of these cases as part of their work flow, using a login restricted web-based image viewer software (SimagisLive, Smart Imaging Technologies Co., Houston, TX, USA) on high-resolution 24-inch monitors. The reviewers were blinded to the original diagnoses and were provided with the same amount of information available during original sign out session. The reviewers were expected to independently assess the virtual slides, render a diagnosis, and provide the WHO grade when applicable. Following initial review of H&E images, WSI for immunohistochemical (IHC) stains was provided only upon the study neuropathologists' request and only if they were performed during the original work-up. WSI-based diagnoses, WHO grades, comments, and IHC and special stain requests were recorded. Following a washout period of 2–6 months, both neuropathologists were provided with the original microscopic glass slides and the same clinical information used for WSI and the same parameters were recorded.

After completion of the review by both neuropathologists, the study coordinator collated data, compared reviewers'

slide- and WSI-based diagnoses as well as the original report diagnoses to determine the concordance. The results were then reassessed by the coordinator and reviewers for the discordant cases to achieve a consensus diagnosis and to evaluate the significance and the reasons for discordances. Data evaluation and generation of crosstabs were performed with SPSS Advanced Statistical Package, version 16.0.1. (IBM Corporation, Armonk, NY, USA).

RESULTS

Among a total of 319 consecutive surgical neuropathology cases reviewed during the study period, 97 cases were included in the study after applying the inclusion/exclusion criteria. Sixteen of these cases were sent from other institutions for consultation. All cases were reviewed on WSI by both study neuropathologists. The original slides from 16 patients were not available to review for neuropathologist #2 (consultation cases returned to primary institution or original slides send out for clinical trials).

The concordance rates between the WSI- and slide-based diagnoses were 94.9% for neuropathologist #1 and 88% for neuropathologist #2. The list of discordant cases is provided in Table 1. Neuropathologist #1 had 5 cases and neuropathologist #2 had 10 cases for which the diagnoses of WSI and glass slide were discordant with each other. The concordance rates between WSI-based diagnosis and original diagnosis were 92.8% for neuropathologist #1 and 85.6% for neuropathologist #2. Slide-based diagnoses and original diagnoses were concordant in 97.9% of

cases for neuropathologist #1 and 91.6% of cases for neuropathologist #2.

Representative images captured from the WSI files of discordant cases (case #11 and #52) are presented in Figure 1. These cases were diagnosed as WHO Grade II for astrocytoma and WHO Grade I for meningioma on digital images. The light microscopic diagnoses were anaplastic astrocytoma, WHO Grade III for case #11 and atypical meningioma, WHO Grade II for case #52.

IHC stains were performed in 22 (22.7%) cases to establish a definitive diagnosis at the time of the original pathology evaluation. Reviewer #1 requested IHC on
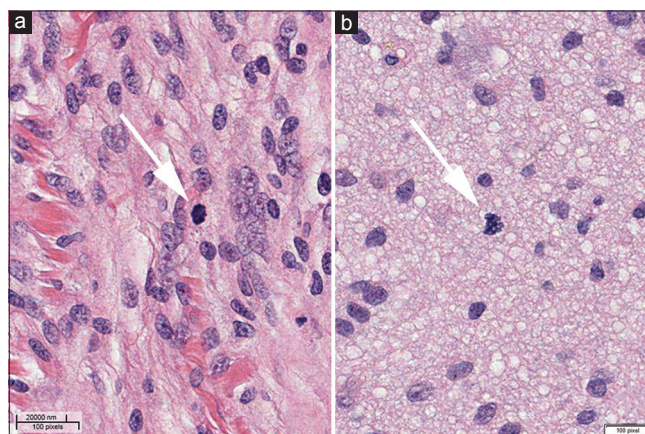
Figure 1: Screen images captured from whole slide image files at high magnification highlighting the challenges in identification of mitoses. The arrows identify mitotic figures that are unequivocally clear on light microscopy but are less distinctive in the digital image. (a) Atypical meningioma, case #52; (b) anaplastic astrocytoma, case #11

Table 1: Discordance between whole slide imaging-based and slide-based diagnoses

Case	Reviewer	WSI-based diagnosis	Slide-based diagnosis	Other reviewer's diagnosis
		Discordances due to identification of mitoses		
53	1	Astrocytoma	Anaplastic Astrocytoma*	Astrocytoma
11	2	Astrocytoma	Anaplastic astrocytoma*	Anaplastic astrocytoma
45	1	Meningioma	Atypical meningioma*	Atypical meningioma
92	1	Meningioma	Atypical meningioma*	Meningioma
52	2	Meningioma	Atypical meningioma*	Atypical meningioma
97	2	Meningioma	Atypical meningioma*	Atypical meningioma
		Discordances due to interpretation of nuclear detail		
65	1	GBM with oligodendroglioma component	Anaplastic oligodendroglioma*	Anaplastic oligodendroglioma
85	2	Atypical neurofibroma	Neurofibroma*	Neurofibroma
22	1	Anaplastic astrocytoma	Oligodendroglioma*	Anaplastic astrocytoma
		Discordances due to diagnostic interpretation		
62	2	Atypical meningioma	Rhabdoid meningioma*	Rhabdoid meningioma
71	2	Astroblastoma	Anaplastic ependymoma*	GBM with focal ependymal features
23	2	Atypical meningioma	Meningioma*	Meningioma
28	2	Anaplastic oligodendroglioma*	Metastatic carcinoma	Anaplastic oligodendroglioma
19	2	DNET*	Oligodendroglioma	DNET
47	2	DSRCT*	Rhabdomyosarcoma	DSRCT

*Diagnoses are concordant with the original pathology report. WSI: Whole slide imaging, GBM: Glioblastoma, DNET: Dysembryoblastic neuroepithelial tumor, DSRCT: Desmoplastic small round cell tumor

14 (14.4%) cases at the time of WSI review and on 17 (17.5%) cases at the time of slide review. Reviewer #2 requested IHC on 16 of 97 cases (16.5%) at the time of WSI review and on 15 of 83 cases (18.1%) at the time of slide review. These numbers did not yield significantly different results.

DISCUSSION

Our results indicate 94.9% and 88% intraobserver agreement for the use of WSI in the primary diagnosis of surgical neuropathology cases. The focus of this study is to evaluate the concordance of WSI-based and glass slide diagnoses rendered by the same pathologist rather than possible diagnostic discrepancies between experts due to interpretation that can be encountered in surgical neuropathology. This is in parallel with the more recent validation studies[8,11] and CAP recommendations[14] as one of their 12 guideline statements specifically says, "The validation study should establish diagnostic concordance between digital and glass slides for the same observer, i.e., intraobserver variability." They suggested that this approach would help differentiate the potential weaknesses and technical properties of virtual microscopy from individualistic interpretive tendencies as the reason for any discrepancy.

Further review of the discordant cases showed no overlap between the two reviewers even though the challenges that led to these discordances were similar. Upon discussion of the discrepancies with the reviewers to achieve a consensus, the most likely underlying reasons were difficulty in identification of mitoses and nuclear details. It is also important to note that we have encountered these problems despite the fact that we have used a high (×40) scanning resolution. We did not compare our high scanning resolution to ×20 scanning results; however, a previous study comparing ×20 and ×40 resolution scans did not find any tangible differences in diagnosis.[4]

While it is difficult to identify the exact reasons for some of the discordant diagnoses, a few problems emerge as potential explanations:
*   The majority of discordant cases shows a trend toward under-grading by WSI, most often due to difficulties in the identification of mitotic figures in the WSI. As shown in Figure 1, many mitotic figures readily identified by the light microscope on glass slides are at best difficult to recognize as mitoses on WSI. This potential pitfall resulted in under-grading meningiomas and infiltrating gliomas since in both group of tumors, the frequency of mitoses are critical in determining the WHO grade. To the best of our knowledge, no study has previously identified problems in diagnosis and grading based on the inability to identify mitoses on a WSI platform.

Therefore, it is entirely reasonable to suggest that the discordance in diagnosis can be attributed to this challenge, particularly for surgical neuropathology. In one meningioma case, the diagnosis in WSI was higher grade (atypical) than the slide-based diagnosis but the reason for considering a higher grade in this case was not the observation of increased mitoses. Thus, in virtually all cases of under-grading, one can say with relative certainty that the major reason was the inability of the observer to recognize mitoses in WSI review.
*   The loss of nuclear details and distortion of the chromatin pattern may at least partially explain some of the discordances. This suggests that at least a component of diagnostic variation by the neuropathologist would be the difficulty in reliably identifying the nuclear characteristics of the tumors. This is particularly relevant in the differentiation of the oligodendroglial versus astrocytic tumors. Most experts agree that variable hyperchromasia versus homogeneously fine chromatin patterns are the most critical morphological difference between these tumors and before molecular characterization constituted the main criterion to separate one from the other. We have identified three cases in which the difficulty in recognizing nuclear detail may have resulted in diagnostic discordance [Table 1].
*   In five cases, we were not able to identify any issue that may be associated with the nature of generating or interpreting WSI and, in these cases, the diagnostic discrepancy was more of an interpretive nature than technical. While it may be possible that the pathologist may feel uneasy about the use of the WSI for diagnosis in some circumstances, the nature of the discrepancy cannot be solely explained by this conjecture and is most likely subjective. This is suggested by the fact that in three of these cases, WSI diagnosis was concordant with the other neuropathologist's diagnoses while in two others, the slide-based diagnosis showed concordance [Table 1]. Suffice it to say that these discordant cases do not seem to be associated with the application or interpretation of WSI technology.

In this study, we have not considered an "expert review" of cases since expert neuropathologists with more than 15 years of experience signed out all cases and both reviewers are Board Certified neuropathologists. Typically, intraobserver variability measured among such experts is expected to be extremely low while the interpretive discordances identified as interobserver variability are often challenging issues with little or no clear set of criteria. We have not performed an intraobserver variation study on glass slides only, which was not the purpose of this study.

We also recognize that the design of this study omits some of the critical components of everyday practice. For

example, additional data from medical records, review of the imaging studies, or discussions with the neurosurgeon or clinical team as well as special stains to further specify the diagnosis would affect the final report and none were available to the reviewers in our study. The need to narrow the diagnosis to one specific entity without being able to perform the above may be considered as a source for discordance in this study. Nevertheless, careful consideration of these variables is necessary in the design of future validation studies.

The current study is exclusively for surgical neuropathology and it includes a select group of complex cases that could be recognized by review of a single representative slide. We have also excluded the cases with ambiguous diagnoses, including infectious and inflammatory cases. In such cases, it is not easy to provide a specific diagnosis or it is often possible to select a single representative slide. These excluded cases may represent a different study population with different technical problems, and our conclusions may not be applicable to these entities.[4,11] While mitotic figures and chromatin pattern may be important in other subspecialties, our results should be extrapolated with caution.

There are several weaknesses in our study, including relatively small number of cases precluding further statistical analysis or pattern identification. In addition, use of a single representative slide from each case may have introduced a bias to our results, which we attempted to prevent the use of the same slide for both WSI- and slide-based diagnoses. Nevertheless, with all the caveats, we believe that reporting of our results is critical since they identify two potential pitfalls that could limit the utility of WSI for accurate diagnosis and grading of tumors in surgical neuropathology.

In summary, an all-encompassing conclusion about the validity of WSI may not be readily available. Instead, we suggest that the technical properties have not been entirely optimized for WSI to fully replace the traditional microscopy. Awareness of the shortcomings will lead to better and faster optimization of WSI and replacement of traditional microscopy in time, which is not a matter of if but when.

Conflicts of interest
There are no conflicts of interest.

REFERENCES

1. Rojo MG, García GB, Mateos CP, García JG, Vicente MC. Critical comparison of 31 commercially available digital slide systems in pathology. Int J Surg Pathol 2006;14:285-305.
2. Hanna MG, Pantanowitz L, Evans AJ. Overview of contemporary guidelines in digital pathology: What is available in 2015 and what still needs to be addressed? J Clin Pathol 2015;68:499-505.
3. Szymas J, Wolf G, Papierz W, Jarosz B, Weinstein RS. Online Internet-based robotic telepathology in the diagnosis of neuro-oncology cases: A teleneuropathology feasibility study. Hum Pathol 2001;32:1304-8.
4. Al Habeeb A, Evans A, Ghazarian D. Virtual microscopy using whole-slide imaging as an enabler for teledermatopathology: A paired consultant validation study. J Pathol Inform 2012;3:2.
5. 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;3:91-6.
6. Ordi J, Castillo P, Saco A, Del Pino M, Ordi O, Rodríguez-Carunchio L, *et al.* Validation of whole slide imaging in the primary diagnosis of gynaecological pathology in a University Hospital. J Clin Pathol 2015;68:33-9.
7. 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;43:1739-44.
8. Houghton JP, Ervine AJ, Kenny SL, Kelly PJ, Napier SS, McCluggage WG, *et al.* Concordance between digital pathology and light microscopy in general surgical pathology: A pilot study of 100 cases. J Clin Pathol 2014;67:1052-5.
9. Bauer TW, Schoenfield L, Slaw RJ, Yerian L, Sun Z, Henricks WH. Validation of whole slide imaging for primary diagnosis in surgical pathology. Arch Pathol Lab Med 2013;137:518-24.
10. Buck TP, Dilorio R, Havrilla L, O'Neill DG. Validation of a whole slide imaging system for primary diagnosis in surgical pathology: A community hospital experience. J Pathol Inform 2014;5:43.
11. Arnold MA, Chenever E, Baker PB, Boué DR, Fung B, Hammond S, *et al.* The College of American Pathologists guidelines for whole slide imaging validation are feasible for pediatric pathology: A pediatric pathology practice experience. Pediatr Dev Pathol 2015;18:109-16.
12. Horbinski C, Wiley CA. Comparison of telepathology systems in neuropathological intraoperative consultations. Neuropathology 2009;29:655-63.
13. Thrall MJ, Rivera AL, Takei H, Powell SZ. Validation of a novel robotic telepathology platform for neuropathology intraoperative touch preparations. J Pathol Inform 2014;5:21.
14. Pantanowitz L, Valenstein PN, Evans AJ, Kaplan KJ, Pfeifer JD, Wilbur DC, *et al.* Review of the current state of whole slide imaging in pathology. J Pathol Inform 2011;2:36.