

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/204292>

Please be advised that this information was generated on 2020-12-04 and may be subject to change.

Validation of Whole-slide Digitally Imaged Melanocytic Lesions: Does Z-Stack Scanning Improve Diagnostic Accuracy?

Bart Sturm¹, David Creytens², Martin G. Cook³, Jan Smits⁴, Marcory C. R. F. van Dijk⁵, Erik Eijken⁶, Eline Kurpershoek⁴, Heidi V. N. Küsters-Vandeveld⁷, Ariadne H. A. G. Ooms⁴, Carla Wauters⁷, Willeke A. M. Blokx⁸, Jeroen A. W. M. van der Laak¹

¹Department of Pathology, Radboud University Medical Center, Nijmegen, The Netherlands, ²Department of Pathology, Ghent University Hospital, Ghent, Belgium,

³Department of Histopathology, Royal Surrey County Hospital, Guildford, United Kingdom, ⁴Pathan B.V., Rotterdam, The Netherlands, ⁵Pathology-DNA, Rijnstate Hospital, Arnhem, ⁶Laboratory for Pathology East Netherlands (LabPON), Hengelo, ⁷Department of Pathology, Canisius Wilhelmina Hospital, Nijmegen, The Netherlands, ⁸Department of Pathology, University Medical Center Utrecht, Utrecht, The Netherlands

Received: 11 July 2018

Accepted: 17 December 2018

Published: 21 February 2019

Abstract

Background: Accurate diagnosis of melanocytic lesions is challenging, even for expert pathologists. Nowadays, whole-slide imaging (WSI) is used for routine clinical pathology diagnosis in several laboratories. One of the limitations of WSI, as it is most often used, is the lack of a multiplanar focusing option. In this study, we aim to establish the diagnostic accuracy of WSI for melanocytic lesions and investigate the potential accuracy increase of z-stack scanning. Z-stack enables pathologists to use a software focus adjustment, comparable to the fine-focus knob of a conventional light microscope. **Materials and Methods:** Melanocytic lesions ($n = 102$) were selected from our pathology archives: 35 nevi, 5 spitzoid tumors of unknown malignant potential, and 62 malignant melanomas, including 10 nevoid melanomas. All slides were scanned at a magnification comparable to use of a $\times 40$ objective, in z-stack mode. A ground truth diagnosis was established on the glass slides by four academic dermatopathologists with a special interest in the diagnosis of melanoma. Six nonacademic surgical pathologists subspecialized in dermatopathology examined the cases by WSI. **Results:** An expert consensus diagnosis was achieved in 99 (97%) of cases. Concordance rates between surgical pathologists and the ground truth varied between 75% and 90%, excluding nevoid melanoma cases. Concordance rates of nevoid melanoma varied between 10% and 80%. Pathologists used the software focusing option in 7%–28% of cases, which in 1 case of nevoid melanoma resulted in correcting a misdiagnosis after finding a dermal mitosis. **Conclusion:** Diagnostic accuracy of melanocytic lesions based on glass slides and WSI is comparable with previous publications. A large variability in diagnostic accuracy of nevoid melanoma does exist. Our results show that z-stack scanning, in general, does not increase the diagnostic accuracy of melanocytic.

Keywords: Melanoma, validation, whole-slide image, whole-slide imaging, z-stack

INTRODUCTION

Classification of specific types of melanocytic lesions may be challenging for pathologists. Most importantly, differentiation between a benign and malignant lesion requires detection of subtle details such as cytomorphologic atypia and mitotic figures. A constant high quality of hematoxylin and eosin (H and E)-stained slides, enabling assessment of all the required information at high magnification, is of utmost importance. The depth of field is lower for high magnification levels. Therefore, adjustment with the fine focusing knob of the microscope brings minute details in focus.

Today, the technique of whole slide imaging (WSI) has reached the point that glass slides can be scanned within minutes and

the image quality qualifies for routine pathology diagnostics. In Europe, some WSI instruments have regulatory approval for diagnostic use marked by CE IVD (European Conformity for *in vitro* Diagnostic Medical Devices).^[1] The technology has recently been approved by the United States Food and Drug administration (FDA) as an alternative to microscopic inspection of glass slides for primary diagnostics.^[2] Several

Address for correspondence: Dr. Jeroen A.W.M. van der Laak, Department of Pathology, Radboud University Medical Center, 824, P.O. Box 9101, 6500HB, Nijmegen, The Netherlands. E-mail: jeroen.vanderlaak@radboudumc.nl

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: reprints@medknow.com

How to cite this article: Sturm B, Creytens D, Cook MG, Smits J, van Dijk MC, Eijken E, *et al.* Validation of whole-slide digitally imaged melanocytic lesions: Does Z-stack scanning improve diagnostic accuracy?. *J Pathol Inform* 2019;10:6.

Available FREE in open access from: <http://www.jpathinformatics.org/text.asp?2019/10/1/6/252683>

Access this article online

Quick Response Code:



Website:
www.jpathinformatics.org

DOI:
10.4103/jpi.jpi_46_18

pathology labs already render clinical diagnosis for routine patient care using WSI.^[3] According to laboratory quality guidelines a validation process should be carried out before implementing WSI for clinical purposes.^[4] For this purpose, the college of American pathologists (CAP) provided a guideline for validation of WSI.^[5]

A number of dermatopathology WSI validation studies have been published.^[6-10] These studies show considerable variation in sample sizes, case selection, and types of lesions studied. Only two studies focused on the diagnosis of melanocytic lesions,^[7,9] where image quality is of paramount importance. In these studies, concordance varied significantly (75.6% vs. 93.2%), which may be due to differences in the study design. One of the limitations of WSI, as it is most often used, is the lack of a multiplanar focusing option. Although many WSI devices are capable of providing software focusing (using so-called z-stacks: a series of images captured at the same spatial location with varying focus plane), this is hardly ever used in practice. The main reasons for this are the strongly increased scanning time and required storage space. To the best of our knowledge, no previous studies have investigated the potential increase in diagnostic accuracy of using software focusing for WSI-based diagnostics in dermatopathology.

In the present study, we addressed this issue in the diagnosis of challenging melanocytic lesions. The first objective of this study is to establish the diagnostic accuracy of WSI for melanocytic lesions including difficult spitzoid lesions and nevoid melanoma. Our second objective is to study the potential accuracy increase of using z-stack scanning including recognition of dermal mitoses.

MATERIALS AND METHODS

Case selection and slide scanning

Benign nevi and malignant melanoma cases from 2011 to 2012 were randomly selected from the archives of the Pathology Department of the Radboud UMC. The set was expanded with additional challenging cases comprising halo nevi, dysplastic nevi, and spitzoid nevi, on the basis of the original pathology reports. In addition, a set of ten nevoid melanomas was included, which were initially offered for consultation to the Pathology Department of the Radboud UMC. These cases were diagnosed as nevoid melanoma by a dermatopathologist with a special interest in the diagnosis of melanoma (WB). Diagnoses for all ten cases were subsequently confirmed by an independent consultant dermatopathologist (MC). In addition, five of these cases were also confirmed by the EORTC melanoma group (fall meeting, Barcelona, 2011).

From every case ($n = 102$ in total), one representative H and E slide was selected. All cases were scanned using a Panoramic 250 Flash II scanner (3DHitech, Hungary; using a $\times 20$ objective lens, specimen-level pixel size, $0.243 \mu\text{m} \times 0.243 \mu\text{m}$, and applying 80% of JPEG compression rate). This optical setup results in an on-screen magnification comparable to a $\times 400$ magnification when using a conventional

microscope. The slides were scanned in z-stack mode, using seven levels with an interplane distance of $0.6 \mu\text{m}$. Resulting images were examined on a computer screen at a resolution that is comparable to the use of a $\times 400$ magnification at a traditional light microscope.

Reference standard

Three academic pathologists (referred to as “reference pathologists”) specialized in dermatopathology (WB, WM, and DC) revised the original glass slides of 92 cases, except for the nevoid melanomas ($n = 10$). Cases were offered with clinical information (age, gender, and location on the skin) and were classified as either benign, malignant, or melanocytic tumor of unknown malignant potential (MELTUMP) with subsequent specific diagnosis in concordance with the WHO classification of melanocytic lesions.^[11] For lesions classified as malignant or MELTUMP, the presence of dermal mitotic activity had to be assessed. In addition, lesions classified as MELTUMP were stratified into low risk and high risk.^[12]

The reference standard was achieved by applying a four-tier system: benign ($n = 1$), MELTUMP low risk ($n = 2$), MELTUMP high risk ($n = 3$), and malignant ($n = 4$). If at least two reference pathologists agreed on a diagnosis and the third pathologist disagreed with a difference of at maximum one category, the majority diagnosis was taken as the reference standard. Cases, where one reference pathologist classified a lesion two categories or more differently, were admitted blindly to all three reference pathologists again, in an attempt to achieve consensus in a second round. If still no consensus was reached, the case was labeled as “discordant” and excluded from the study.

Dermal mitotic activity was defined as “present” if at least two reference pathologists had registered this. Cases, where only one of the three pathologists registered a dermal mitosis, were re-assessed specifically for dermal mitotic activity by one of the three reference pathologists (WB).

Whole-slide imaging assessment

The set of whole-slide images ($n = 102$) was assessed by six nonacademic surgical pathologists (referred to as “study pathologists”), with a subspecialty in dermatopathology (AO, CW, EE, EK, HK, and MvD). Cases were diagnosed using Panoramic Viewer (3DHitech, version 1.15.4) on a 24-inch ultra HD color calibrated monitor (Dell UltraSharp; resolution: 3840×2160 pixels). Calibration of the monitor was performed using a Datacolor Spyder3 Elite colorimeter. The participating pathologists followed a short course using the slide viewer, particularly emphasizing the software focusing option. They were instructed to assess a WSI first without using the software focusing option. Subsequently, the study pathologists were allowed to immediately re-assess the WSI using the software focusing option, if deemed necessary. All cases, including nevoid melanoma, were offered to the study pathologists with clinical information (age, gender, and location on the skin) and were classified identical to the classification performed by the reference pathologists (i.e., using a four-tier system).

Statistical analysis

For statistical analysis, the four-tier scheme defined above was downsized to a three-tier system by combining high-risk and low-risk MELTUMP, as clinical management of these lesions is largely identical. Concordance of pathologists with the consensus diagnosis was expressed as the number and percentage of cases with identical diagnosis (in the three-tier system) for every subclass as well as overall. As an overall measure of concordance of the study pathologists with the consensus diagnosis, Kappa statistics with 95% confidence intervals (CIs) were calculated. The effect of using software focusing was analyzed using Chi-Square tests with 95% CIs.

RESULTS

A consensus diagnosis by the reference pathologists could initially be achieved in 81 (88%) of 92 cases. After blindly admitting discordant cases ($n = 11$) a second time to the individual reference pathologists, consensus was reached in 89 (97%) of the cases. Therefore, the set of WSI assessed by the study pathologists contained 99 cases (35 benign, 4 MELTUMP, and 60 malignant, including 10 nevoid melanoma). The concordance of the individual reference pathologists with the consensus diagnosis is shown in Table 1. Except for MELTUMP cases, concordance exceeded 91% for every pathologist for both benign and malignant cases. MELTUMP concordances were lower (average 66.7%), but the set of cases was very small ($n = 4$). It should be noted that

in Table 1, reference pathologists' diagnoses are compared with the reference diagnosis, which itself was partly based on the studied diagnoses. The purpose of this analysis is to better understand the quality of the reference standard used here and should not be considered an independent measure for individual pathologists' performance.

WSI assessment by the six study pathologists was primarily studied without the use of the software focusing option, as this is the generally accepted method in which WSI is used. WSI assessment by the six study pathologists against the consensus diagnosis is shown in Table 2. Results of the ten nevoid melanoma cases are displayed separately. Concordance generally exceeds 80% for individual pathologists. Interestingly, the majority of study pathologists classified most MELTUMP cases as either benign or malignant. Furthermore, two pathologists (Path 3 and Path 4) misdiagnosed the large majority of nevoid melanoma, which are notoriously difficult, as being benign in almost all cases. In one case, a nevoid melanoma was classified as a MELTUMP high risk (Path 3). Kappa statistics for these assessments are shown in Table 3, for the set of cases not including the nevoid melanoma ($n = 89$). As nevoid melanomas occur rarely in pathology diagnostic practice, we consider this the most representative way of analysis. The facultative use of software focusing did not show a significant change (Chi-square test $P = 0.85$) in concordance with the consensus diagnosis as shown in Table 4. Although not significant, study pathologist 5 changed the diagnosis once from benign (incorrect diagnosis) to a melanoma (correct diagnosis) after use of the software focusing.

In the malignant and MELTUMP cases ($n = 54$), dermal mitosis was initially detected in 35 cases by at least one of the three reference pathologists. In ten of these 35 cases, only one of the three reference pathologists mentioned dermal mitosis. After reviewing these ten cases (WB), only three cases with dermal mitotic activity remained, establishing 28 out of 54 melanomas and MELTUMP cases with dermal mitotic activity.

Recognition of dermal mitotic activity did not improve significantly using the software focusing option concerning the

Table 1: Concordance of individual reference pathologists with consensus diagnosis based on glass slides

	Benign	MELTUMP	Malignant	Overall
Number of cases	35	4	50	89
Reference pathologists				
1	33 (94)	3 (75)	50 (100)	86 (97)
2	32 (91)	3 (75)	48 (96)	83 (93)
3	33 (94)	2 (50)	48 (96)	83 (93)
Average (%)	93	67	97	94

Number of cases concordant with consensus diagnosis. Percentage within brackets. "Average" shows the average concordance (%) over the three pathologists. MELTUMP: melanocytic tumor of unknown malignant potential

Table 2: Concordance of individual study pathologists with consensus diagnosis without use of z-stack

	Benign	MELTUMP	Malignant	Nevoid melanomas	Overall ^a	Overall ^b
Number of cases	35	4	50	10	89	99
Path 1	29 (83)	2 (50)	45 (90)	8 (80)	76 (85)	84 (85)
Path 2	34 (97)	0 (0)	45 (90)	7 (70)	79 (89)	86 (87)
Path 3	33 (94)	1 (25)	38 (76)	2 (20)	72 (81)	74 (75)
Path 4	33 (94)	0 (0)	34 (68)	1 (10)	67 (75)	68 (69)
Path 5	32 (91)	0 (0)	43 (86)	8 (80)	75 (84)	83 (84)
Path 6	34 (97)	0 (0)	46 (92)	8 (80)	80 (90)	88 (89)
Average (%)	93	13	84	57	84	81

Number of cases concordant with consensus diagnosis. Percentage within brackets. "Average" shows the average concordance (%) over the six pathologists. ^aOverall number of cases excluding nevoid melanoma cases; ^bOverall number of cases including nevoid melanoma cases. MELTUMP: melanocytic tumor of unknown malignant potential

nonnevoid melanoma cases. Pathologist 1 changed three times the status of dermal mitotic activity: twice into true positive and once into false negative [Figure 1]. Pathologist 5 changed three times, all into false positive. The remaining pathologists did not change the status of dermal mitotic activity. Noteworthy, concerning the nevoid melanoma cases is the fact that study pathologist 5 changed the status of dermal mitotic activity in five of the ten cases from negative into positive. As mentioned

above, this subsequently changed the diagnosis once from benign (incorrect) into (nevoid) melanoma. In the remaining four cases, the diagnosis remained melanoma.

DISCUSSION

Multiple studies have been carried out to assess the interobserver variability concerning diagnosis of melanocytic lesions on glass slides. These studies vary in design, for example, number of cases, selection of cases, and experience of pathologists. In this study, we did not include the use of ancillary techniques like immunohistochemistry, because we believe it will not influence the diagnosis on digital slides compared to glass slides. Nevertheless, various studies that (partly) approximate the study we performed are available. For example, Elmore *et al.* studied concordance of 240 melanocytic lesions among 187 pathologists with 48 cases per pathologist.^[13] In that particular study, the cases were classified according to the Melanocytic Pathology Assessment Tool and Hierarchy for Diagnosis (MPATH-Dx) into five classes, i.e., I (e.g., nevus or mild atypia); II (e.g., moderate atypia); III (e.g., severe atypia or

Table 3: Kappa values (95% confidence interval) of individual study pathologists for all cases excluding nevoid melanoma

	κ (n=89)
Path 1	0.74 (0.60-0.85)
Path 2	0.79 (0.67-0.91)
Path 3	0.66 (0.51-0.79)
Path 4	0.55 (0.39-0.70)
Path 5	0.72 (0.58-0.83)
Path 6	0.81 (0.69-0.91)

Table 4: Concordance of individual study pathologists with consensus diagnosis (n=99) with and without use of software focusing, i.e., z-stack

	z-stack use (n)	Concordance without z-stack (%)	Change of diagnosis (n)	Concordance with z-stack (%)
Path 1	28	82	3	82
Path 2	7	100	0	100
Path 3	26	73	0	73
Path 4	3	67	0	67
Path 5	22	82	1	86
Path 6	12	92	0	92

Number of cases z-stack is used, number of change of diagnosis and concordance (%) with consensus before and after z-stack is used

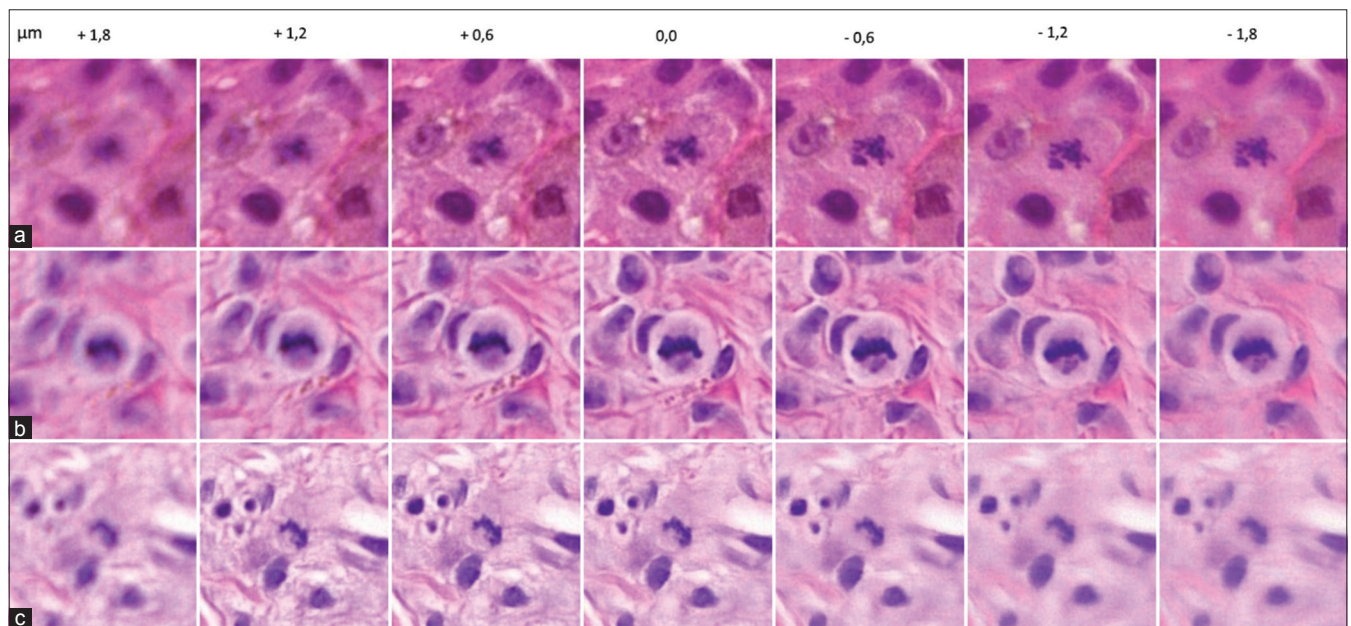


Figure 1: Examples of changes in the status of dermal mitotic activity by pathologist 1 after the use of software focusing. Case 7 (a) changed into true positive. Case 59 (b) changed into false negative. Case 97 (c) changed into true positive. Note that the best field of view of the mitosis is situated mainly below and above the zero z-plane in cases 7 and 97, respectively

melanoma *in situ*); IV (e.g., pathologic stage T1a [pT1a] early invasive melanoma); and V (e.g., \geq pT1b invasive melanoma), and 53 and 134 participants were academic and nonacademic pathologists, respectively. Overall interobserver concordance with the diagnosis of benign nevi was 92% (90–94), early invasive melanoma 43% (39–46), and (clearly) invasive melanoma 72% (69–75). Lott *et al.* carried out a comparable study of 48 cases classified according to the MPATH-Dx among 16 pathologists considered to be experts in the typing of melanocytic lesions, including 14 academic and 2 nonacademic.^[14] This study showed that concordance of benign lesions was 87.8%, (...), early invasive and (clearly) invasive melanoma grouped together 87.0% presuming the most severe diagnosis. Interobserver agreement with weighted kappa coefficients was 0.72 (95% CI: 0.71–0.73) assuming the most severe diagnosis. A final example concerning glass slides is the study carried out by Corona *et al.* This interobserver study concerned 140 melanocytic lesions, of which 120 were melanoma among four dermatopathologists with at least 10 years of experience.^[15] Overall concordance concerning melanoma was 81%, and kappa interobserver agreement was 0.61 (95% CI: 0.54–0.68). Obviously, the figures obtained in these studies are heavily dependent on the selection of cases studied.

In the past two decades, a number of WSI validation studies have been performed.^[3,16] WSI studies specifically for melanocytic lesions are small in number, especially those applying the CAP guidelines.^[5] Shah *et al.* studied 181 cases for concordance of skin lesions, of which 60 concerned melanocytic lesions.^[9] It was found that pathologists achieved an overall interobserver concordance of 82.8% concerning light microscopy versus light microscopy and 80.0% concerning WSI versus WSI in these 60 cases. It was found that pathologists achieved an overall intraobserver concordance of 75.6% (95% CI: 68.5–81.5) concerning WSI versus light microscopy.

In a larger study, Leinweber *et al.* reported on 560 melanocytic lesions, of which 280 were clearly malignant and 280 clearly benign, which were examined by four academic pathologists.^[7] In this study, interobserver agreement was not evaluated. Overall intraobserver concordance was 93.2% (90.4–96.4).

The results of our study, if we exclude all nevoid melanoma cases, are in line with the studies mentioned above. Agreement between the reference pathologists with the consensus diagnosis appears to be higher in our study. Reasons for this can be multiple, for example, methodology, quality of the H and E slides, a second read of discordant cases, methodology of reaching consensus, and less difficult cases. Agreement between the study pathologists is in accordance with other WSI validation studies, although the number of studies available is relatively small. Overall concordance is 84% (75%–90%; $n = 89$) compared to 80.0% ($n = 60$) in the study of Shah *et al.*^[9]

Recognition of nevoid melanoma is a challenge and is a known pitfall among pathologists.^[17] The results in this study show a bimodal distribution. Four pathologists recognized seven to

eight of the ten cases on WSI; two pathologists recognized, respectively, two and one of the cases only. Although the literature yields no exact data, in our opinion, these results reflect the existing daily practice. That stated, it may be possible the bimodal distribution can be attributed to the fact some study pathologists have no prior experience examining WSI. Pathologists 1 and 2 have significant WSI experience; the remaining pathologists have not. Therefore, it appears that WSI experience is not of significant importance to recognize nevoid melanoma as such.

So far, only two studies focused on the possible advantage of software focusing, i.e., z-stack scanning, in histology.^[18,19] Specifically, in dermatopathology, concerns have emerged that WSI may not sufficiently allow identification of mitosis, which is an important feature for the diagnostic evaluation of many melanocytic lesions.^[20] In the field of surgical neuropathology, Pekmezci *et al.* stated that the major reason for undergrading in brain tumor cases was the inability to recognize mitosis in WSI review.^[21]

Studies concerning software focusing are mainly done in the field of cytology.^[22,23] We note that in the field of histology, only one study has been reported.^[18] Kalinski *et al.* showed that software focusing could be of value to improve the detection of the microorganism *Helicobacter pylori* in gastric biopsies. The assessment of inflammation and intestinal metaplasia did not improve. Snead *et al.* stated that the omission of multiplanar scanning in histology samples does not seem to be a common problem, where sections in recent days are cut at 4 μ m by most pathology departments.^[19] To our knowledge, this statement is not substantially scientifically proven. Due to this lacuna in the literature, the study we performed may be the first to prove to underline this statement concerning melanocytic lesions. We believe that our study shows that, in general, the use of software focusing appears not to be of great significance to render a more accurate diagnosis in melanocytic lesions, including nevoid melanoma. Furthermore, in general, it does not provide a better recognition of mitosis. False negative cases were mainly missed dermal mitoses that were unmistakably present. The cases in which pathologist 1 and 5 changed the status of dermal mitotic activity after the use of software focusing were reviewed ($n = 6$, all non-nevoid melanomas) (BS). In the cases of pathologist 5, i.e. all false positive mitoses, it was not possible to find the dermal mitoses on WSI and the glass slides. Although in one case an epidermal mitosis could have been mistaken for a superficial dermal mitosis. Interestingly, in the ten nevoid melanoma cases, one of the six pathologists changed the status of dermal mitotic activity in five cases from negative into positive and subsequently changed the diagnosis once from benign into nevoid melanoma. Although the dermal mitotic activity was not confirmed by the study pathologists due to the study design, this fact may be relevant resulting in the following hypothesis: “concerning melanocytic lesions, in very challenging cases, a software focusing option may be of added value to recognize (dermal) mitosis with more confidence, resulting in a more accurate diagnosis.” Therefore,

in a melanoma consulting pathology practice, z-stack scanning may be of utmost importance to achieve and maintain the highest degree of diagnostic accuracy. Although this statement has to be proven or falsified by future research.

A downside of WSI z-stack scanning is that the scanning time and storage needed for each slide are linear with the number of z-planes scanned. In our study, the file size of each WSI was seven times the size of a normal WSI.

CONCLUSION

Diagnostic accuracy of melanocytic lesions based on glass slides and WSI is comparable with previous publications, however, small in number they are. A large variability in diagnostic accuracy of nevoid melanoma does exist, which reflects daily practice, but may partly be caused by the use of WSI. Importantly, our results show that z-stack scanning, in general, does not increase diagnostic accuracy of melanocytic lesions concerning classification and detection of dermal mitosis. Establishing this knowledge is important because it may prevent unnecessary investments in, for example, data storage.

Acknowledgments

We like to thank Prof. Dr. W. J. Mooi, a pathologist at VU University Medical Center, Amsterdam, The Netherlands, for his expertise, efforts, and valuable comments. We like to thank Dr. M. Nap, a former pathologist at Zuyderland MC, Heerlen, The Netherlands, and others for their help and technical support.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

REFERENCES

1. The Royal College of Pathologists. Best practice recommendations for implementing digital pathology. 2018. p. 7.
2. FDA allows Marketing of First whole Slide Imaging System for Digital Pathology [press release]. Available from: <http://www.FDA.gov>. [Last accessed on 2017 Apr 12].
3. Goacher E, Randell R, Williams B, Treanor D. The diagnostic concordance of whole slide imaging and light microscopy: A systematic review. *Arch Pathol Lab Med* 2017;141:151-61.
4. International Organization of Standardization. ISO 15189:2012 Medical Laboratories – Requirements for Quality and Competence. International Organization of Standardization; 2012.
5. Pantanowitz L, Sinard JH, Henricks WH, Fatheree LA, Carter AB, Contis L, *et al*. Validating whole slide imaging for diagnostic purposes in pathology: Guideline from the College of American Pathologists Pathology and Laboratory Quality Center. *Arch Pathol Lab Med* 2013;137:1710-22.
6. Velez N, Jukic D, Ho J. Evaluation of 2 whole-slide imaging applications

- in dermatopathology. *Hum Pathol* 2008;39:1341-9.
7. Leinweber B, Massone C, Kodama K, Kaddu S, Cerroni L, Haas J, *et al*. Teledermatopathology: A controlled study about diagnostic validity and technical requirements for digital transmission. *Am J Dermatopathol* 2006;28:413-6.
8. Al-Janabi S, Huisman A, Vink A, Leguit RJ, Offerhaus GJ, Ten Kate FJ, *et al*. Whole slide images for primary diagnostics in dermatopathology: A feasibility study. *J Clin Pathol* 2012;65:152-8.
9. Shah KK, Lehman JS, Gibson LE, Lohse CM, Comfere NI, Wieland CN, *et al*. Validation of diagnostic accuracy with whole-slide imaging compared with glass slide review in dermatopathology. *J Am Acad Dermatol* 2016;75:1229-37.
10. Nielsen PS, Lindebjerg J, Rasmussen J, Starklint H, Waldstrøm M, Nielsen B, *et al*. Virtual microscopy: An evaluation of its validity and diagnostic performance in routine histologic diagnosis of skin tumors. *Hum Pathol* 2010;41:1770-6.
11. Melanocytic tumours. In: LeBoit PE, Burg G, Weedon D, Sarasain A, editors. *World Health Organisation Classification of Tumours Pathology and Genetics of Skin Tumours*. Lyon: IARC Press; 2006. p. 49-118.
12. Cerroni L, Barnhill R, Elder D, Gottlieb G, Heenan P, Kutzner H, *et al*. Melanocytic tumors of uncertain malignant potential: Results of a tutorial held at the XXIX symposium of the International Society of Dermatopathology in Graz, October 2008. *Am J Surg Pathol* 2010;34:314-26.
13. Elmore JG, Barnhill RL, Elder DE, Longton GM, Pepe MS, Reisch LM, *et al*. Pathologists' diagnosis of invasive melanoma and melanocytic proliferations: Observer accuracy and reproducibility study. *BMJ* 2017;357:j2813.
14. Lott JP, Elmore JG, Zhao GA, Knezevich SR, Frederick PD, Reisch LM, *et al*. Evaluation of the Melanocytic Pathology Assessment Tool and Hierarchy for Diagnosis (MPATH-dx) classification scheme for diagnosis of cutaneous melanocytic neoplasms: Results from the International Melanoma Pathology Study Group. *J Am Acad Dermatol* 2016;75:356-63.
15. Corona R, Mele A, Amini M, De Rosa G, Coppola G, Piccardi P, *et al*. Interobserver variability on the histopathologic diagnosis of cutaneous melanoma and other pigmented skin lesions. *J Clin Oncol* 1996;14:1218-23.
16. Sturm B, Fleskens SJ, Bot FJ, van Velthuysen ML, Speel EJ, Slootweg PJ, *et al*. Virtual microscopy is a valid alternative for the diagnostic assessment of laryngeal premalignancies. *Histopathology* 2014;64:602-4.
17. McKee PH. Clues to the diagnosis of atypical melanocytic lesions. *Histopathology* 2010;56:100-11.
18. Kalinski T, Zwönitzer R, Sel S, Evert M, Guenther T, Hofmann H, *et al*. Virtual 3D microscopy using multiplane whole slide images in diagnostic pathology. *Am J Clin Pathol* 2008;130:259-64.
19. Snead DR, Tsang YW, Meskiri A, Kimani PK, Crossman R, Rajpoot NM, *et al*. Validation of digital pathology imaging for primary histopathological diagnosis. *Histopathology* 2016;68:1063-72.
20. Onega T, Reisch LM, Frederick PD, Geller BM, Nelson HD, Lott JP, *et al*. Use of digital whole slide imaging in dermatopathology. *J Digit Imaging* 2016;29:243-53.
21. Pekmezci M, Uysal SP, Orhan Y, Tihan T, Lee HS. Pitfalls in the use of whole slide imaging for the diagnosis of central nervous system tumors: A pilot study in surgical neuropathology. *J Pathol Inform* 2016;7:25.
22. Bongaerts O, van Diest PJ, Pieters M, Nap M. Working toward consensus among professionals in the identification of classical cervical cytomorphological characteristics in whole slide images. *J Pathol Inform* 2015;6:52.
23. Hanna MG, Monaco SE, Cuda J, Xing J, Ahmed I, Pantanowitz L. Comparison of glass slides and various digital-slide modalities for cytopathology screening and interpretation. *Cancer Cytopathol* 2017;125:701-9.