Research Article

Conventional Microscopical versus Digital Whole-Slide Imaging-Based Diagnosis of Thin-Layer Cervical Specimens: A Validation Study

Odille Bongaerts[1], Carla Clevers[1], Marij Debets[1], Daniëlle Paffen[1], Lisanne Senden[1], Kim Rijks[1], Linda Ruiten[1], Daisy Sie-Go[2], Paul J Van Diest[1,3], Marius Nap[4]

[1]Department of Pathology, Zuyderland Hospital, Heerlen, [2]Department of Pathology, University Medical Center Utrecht, Utrecht, [4]Nap Pathology Consultance bv, Numansdorp, The Netherlands, [3]Department of Oncology, Johns Hopkins Oncology Center, Baltimore, MD, USA

Received: 29 April 2018 Accepted: 29 June 2018 Published: 27 August 2018

Abstract

Background: Whole-slide imaging (WSI) has been implemented in many areas of pathology, but primary diagnostics of cytological specimens are lagging behind. One of the objectives of viewing scanned whole-slide images from histological or cytological specimens is remote exchange of knowledge and expertise of professionals to increase diagnostic accuracy. We compared the scoring results of our team obtained in double readings of two different data sets: conventional light microscopy (CLM) versus CLM and CLM versus WSI. We hypothesized that WSI is noninferior to CLM for primary diagnostics of thin-layer cervical slides. **Materials and Methods:** First, we determined the concordance rate at different thresholds of the participating cytotechnicians by double reading with CLM of 500 thin-layer cervical slides (Cohort 1). Next, CLM was compared with WSI examination of another 505 thin-layer cervical slides (Cohort 2) scanned at ×20 in single focus plane. Finally, all major discordant cases of Cohort 1 were evaluated by an external expert in the field of gynecological cytology and of Cohort 2 in the weekly case meetings. **Results:** The overall concordance rate of Cohort 1 (CLM vs. CLM) was 97.8% (95% confidence interval [CI]: 96.0%–98.7%) and of Cohort 2 was 95.3% (95% CI: 93.0%–96.9%). **Conclusion:** Concordance rates of WSI versus CLM were comparable with those of CLM versus CLM. We have made a step forward paving the road to implementation of WSI also in routine diagnostic cytology.

Keywords: Cervical cytology, cytology, digital pathology, validation study, whole-slide imaging

Introduction

Digital whole-slide images (WSI) have already been implemented in many areas of pathology: education, teleconsultations, slide conferences and panels, quality assurance, frozen section diagnosis, and image analysis.[1-7] One of the objectives of viewing scanned whole slides from histological or cytological specimen is remote exchange of knowledge and expertise of professionals to increase diagnostic accuracy.[1,2,6,8,9] As a result of the increased implementation of WSI in research applications and clinical practice, guidelines for validation of WSI for use in primary diagnostics of histological specimens were recently proposed by the College of American Pathologists (CAP).[2,10-12] What the CAP guidelines indicate is that building up experience in viewing WSI is necessary before starting a validation process and that the normal working situation should be mimicked as much as possible for validating the use of WSI in primary diagnostics. Before the implementation of WSI as a diagnostic tool, one of the critical concerns is to investigate whether WSI comparable with conventional light microscopy (CLM) can be used in primary diagnostics.

WSI scanners have been classified as a Class II medical device[13] with the recent Food and Drug Administration approval for the

Address for correspondence: Mrs. Odille Bongaerts, Department of Pathology, Zuyderland Hospital, PO Box 6446, 6401 CX Heerlen, The Netherlands.
E-mail: o.vandenbergh@zuyderland.nl

Access this article online

Quick Response Code:

Website: www.jpathinformatics.org

DOI: 10.4103/jpi.jpi_28_18

How to cite this article: Bongaerts O, Clevers C, Debets M, Paffen D, Senden L, Rijks K, *et al*. Conventional microscopical versus digital whole-slide imaging-based diagnosis of thin-layer cervical specimens: A validation study. J Pathol Inform 2018;9:29.

Available FREE in open access from: http://www.jpathinformatics.org/text.asp?2018/9/1/29/239856

Philips setup. Several studies aimed at the validation of the use of WSI in primary diagnostics of histological specimens.[1,3,5,10,14-18] However, only a few studies have investigated the use of WSI in primary diagnostics of cytological specimens[8,19-24] and compared this with the performance of the local diagnostic team on conventional glass slides. One recent study used the CAP guideline for the validation of WSI in primary diagnostics of pediatric cytopathology.[10] Recently, we published a study in which we describe how to demonstrate that all diagnostic criteria that we use in CLM can be recognized in WSI and that by training, confidence can be achieved in our group of cytotechnicians and pathologists.[9]

There are several advantages of WSI for diagnosis of cervical thin-layer cytology preparations: allowing accurate annotation to be added to images by multiple investigators without disturbing the specimen, sharing cases, education and training, proficiency testing, and easy and fast review of previously imaged archived cases. In addition, remote diagnostics to concentrate cytology diagnostics on one location of laboratory conglomerates and to deal with the shortage of skilled cytotechnologists in underserved areas in the world are potentially important applications.

However, before WSI can reliably be used in primary diagnosis of cervical cytology, several issues have to be addressed. The present study addresses the validation of the WSI reading within our group of cytotechnicians with experience in WSI examination of thin-layer cervical slides.[9] To this end, we followed as much as possible the working structure of the routine diagnostic process. This approach is in line with the 2013 CAP guideline for implementing digital diagnostics in pathology. We compared the scoring results of our team obtained in double reading of two different datasets: CLM versus CLM, which will give us information about the diagnostic concordance rate within the team, and CLM versus WSI, which will show us what the effect is on the concordance rate of the same team when the second reading was done on WSI. With this in mind, the primary objective of our study was to demonstrate that WSI is noninferior to CLM for primary diagnostics of thin-layer cervical slides.

Materials and Methods

This study was performed at the Department of Pathology of the Zuyderland Hospital, a regional teaching hospital in the south of The Netherlands. Approval of the Internal Review Board was obtained. All seven cytotechnicians who participated in this study were certified in gynecological cytology with >5 years' experience in CLM. Before the start of this study, all participants were educated reviewing WSI of cervical thin-layer specimens.[9]

Cohorts

The cohorts of this study were composed as follows: in two groups of cervical cytology cases, selected in a different way, the natural distribution of the Bethesda classification[25] obtained by the routine CLM examination was determined. The first group consisted of 5 sets of 1000 cases that were selected at random and the second group consisted of 5 sets of 1000 cases chosen as consecutive submissions. In both groups, the Bethesda classification distribution was very similar and it became evident that all Bethesda classifications were present in each selection of at least 300 cases. However, the number of high-grade lesions was always very low, even though these cases were not screening derived. Therefore, we decided not to use 1 fully random or consecutive cohort, but 2 cohorts that were partly consecutive (300 cases) and partly enriched for high-grade lesions (100 cases) mixed with an additional 100 cases with normal cytology from the same period. The first cohort was for double reading by CLM, and the second cohort, with a similar composition, to avoid a possible bias by triple reading was used for comparison of CLM and WSI reading.

Case selection

The case selection of the two cohorts was done as follows: For Cohort 1, 500 nonpopulation-based screening cervical thin-layer slides were selected consisting of 300 consecutive slides from the normal workflow from the period January–March 2016, enriched with 200 selected aberrant slides from the period 2010–2015 comprising 48 high-grade precursor lesions (high-grade squamous intraepithelial lesion [HSIL], atypical glandular cells [AGC], favor of neoplasia [FN], and adenocarcinoma *in situ* [AIS]) and 49 cancer and 103 normal cases. The normal cases were added to avoid any preoccupation with high-grade lesions once the historical slide registration number would be recognized by the cytotechnicians.

Cohort 2 contained an independent but comparable composition of 505 slides: 308 from the normal workflow from the period May–November 2014 enriched with 197 cases consisting of 51 HSIL/AGC/FN/AIS and 48 cancer and 98 normal cases from the period 2009–2014. None of the cases used for enrichment was present in both cohorts.

Conventional light microscopy

For examination of Cohort 1, all cases had been routinely examined (CLM-1) with CLM by an arbitrary cytotechnician. Then, after the removal of marks, the archived slides of these 500 cases were redistributed among the cytotechnicians in the study and examined in small groups in parallel to the regular daily workload. During re-examination (CLM-2), the cytotechnician had the original clinical information available (age, complaints, menstrual pattern, and aspect of the cervix). We followed the same procedure as in routine diagnostics, which means that the second readings were as much as possible done by a cytotechnician different from the first reader. The following items were scored: the quality of the thin-layer slide and a Bethesda classification.[25]

For examination of Cohort 2, a similar approach was followed. However, in addition, all slides were digitized to allow WSI examination. The 308 cases from the normal workflow were scanned before the CLM examination routinely (CLM-3) by an arbitrary cytotechnician. For those 197 cases used as enrichment, scanning took place after conventional reading and after the removal of marks.

Scanning of thin-layer cervical slides and viewing of whole-slide images

Only the thin-layer slides (SurePath, Becton, Dickinson and Company, Franklin Lakes, New Jersey, USA) of Cohort 2 were scanned in one focal plane using a 3DHistech 250 flash II scanner (3DHistech Kft., Budapest, Hungary) equipped with a CIS VCC-FC60 FR 19 CL camera and a ×20 objective (numerical aperture 0.8), using a camera/adapter magnification of 1.6 and quality factor of 70 (JPEG image compression), resulting in 400–600 Mb storage per WSI. The average scanning time was 1 min and 30 s. The WSI was displayed on a HD LCD color monitor (HP 23-inch 1920-1080) and viewed with the Pannoramic Viewer (3DHistech).

Comparison of conventional light microscopy and WSI-based assessments

In order to facilitate a uniform comparison, the findings of both cohorts were scored according to the Bethesda 2014 classification.[25] A separation was made in "concordant" cases and "discordant" cases. In order to obtain more detailed information, the separation analysis was repeated using different thresholds: NORMAL, atypical squamous cells of undetermined significance/ Atypical glandular cells of undetermined significance (ASC-US/ AG-US), low-grade squamous intraepithelial lesion (LSIL), high grade squamous intraepithelial lesion/carcinoma in situ/atypical glandular cells favour of neoplasia/ adenocarcinoma in situ (HSIL/CIS/AGC-FN/AIS) and CANCER. A final score "concordant" was given to those cases that showed no differences in Bethesda classification between the first and second reading. Furthermore, those cases with minor differences in the Bethesda classification between the first and second reading that did not result in differences in treatment but may lead to a different follow-up schedule were considered as concordant.

A final score "discordant" was given if there was a different score between the first and second reading in the Bethesda classification with impact on patient care by a combination of current treatment and follow-up protocols. Cases with unsatisfactory image quality for a confident WSI-based diagnosis were separately noted but excluded for further analysis.

Usually, major discrepancies in CLM are submitted to an external expert for consultation. To follow that procedure, the major discordant cases of Cohort 1, CLM versus CLM, were re-evaluated by an experienced cytopathologist (Dr. D. M. D. S Sie-Go from the University Medical Centre in Utrecht in The Netherlands). All major discrepant cases and the cases that were considered unsatisfactory for evaluation (UFE) due to image quality of Cohort 2 were discussed in weekly case meetings with the diagnostic team to discuss the possible influence of the digital image on this discrepancy in the same way as in our previous study[9] to reach consensus diagnosis or arriving at a final diagnosis based on histology (if available). Our previously established reference atlas (www.ex-pathcytology.com) appeared useful in these discussions.

Statistical analysis

Concordance rates with 95% confidence interval (CI) were calculated for CLM-1 versus CLM-2, and CLM-3 versus WSI. In these comparisons, cases deemed UFE by either method were excluded. Kappa statistics were performed for Cohort 1 to determine the inter-observer reliability at multiple thresholds with CLM. The kappa for Cohort 2 was performed at the same thresholds to determine the correlation between both techniques (CLM and WSI).

The interpretation of the kappa value is as follows:[26] 0 = agreement equivalent to chance, 0.1–0.20 = slight agreement, 0.21–0.40 = fair agreement, 0.41–0.60 = moderate agreement, 0.61–0.80 = substantial agreement, 0.81–0.99 = near perfect agreement, and 1 = perfect agreement. The positive agreement and the negative agreement were determined. These specific agreements respond to the question whether the group cytotechnicians in the examination of both cohorts arrive at the same specific diagnoses.

Figure 1 highlights the workflow of our study.

Results

For Cohort 1, 498/500 cases could finally be read according to the Bethesda classification and were registered as normal or higher. There was one UFE case in CLM-1, and there was one more case registered as UFE for CLM-2 compared to CLM-1. These two cases were excluded from further analysis.

For Cohort 2, 489/505 cases could finally be read according to the Bethesda classification and were registered as normal or higher. There were in total 16 cases considered as UFE. CLM-3 contained eight cases and an additional eight more cases were registered as UFE in WSI. These 16 cases were excluded from further analyses.

Conventional light microscopy double reading

Table 1 shows the aggregated results of the CLM-1 versus CLM-2 examinations of Cohort 1. There were 11/498 (2.2%) discordant cases of which the outcome of the second reading would result in a different treatment and follow-up (Dutch guidelines).[27] Five CLM-1 normal cases were diagnosed as either LSIL (×1), HSIL/AGC-FN/AIS (×3), or cancer (×1) in CLM-2. Of the CLM-2 normal cases, one was classified as LSIL and two were classified as HSIL/AGC-FN/AIS by CLM-1. One ASC-US/AG-US case was diagnosed as cancer in CLM-1. Finally, 2 CLM-1 ASC-US/AG-US cases were diagnosed as HSIL/CIS/AG-FN-AIS with CLM-2. There were in the second reading (CLM-2) in total 7 overcalls and 4 undercalls compared with CLM-1.

As can be seen from Table 1, 48/498 (9.6%) cases revealed minor discrepancies between CLM-1 and CLM-2. Of these, 19 CLM-1 normal cases were judged ASC-US/AG-US in CLM-2. On the other hand, 7 ASC-US/AG-US CLM-1 cases were assessed as normal in CLM-2. CLM-1 contained 2 cases registered as ASC-US/AG-US which were diagnosed as LSIL in CLM-2 and 3 cases diagnosed in CLM-1 as LSIL was

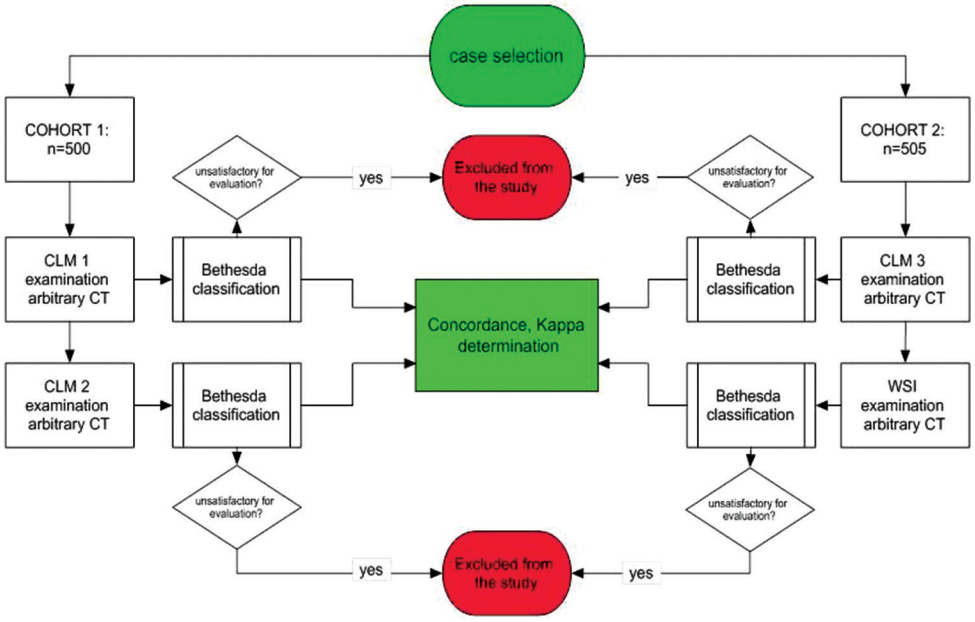

Figure 1: The workflow of the study

Table 1: Cross table showing the results of the two assessments by conventional light microscopy (conventional light microscopy-1 and conventional light microscopy-2)

	CLM-1 (arbitrary cytotechnician)					Total
	Normal	ASC-US/AG-US NOS	LSIL	HSIL/AG-FN-AIS	Cancer	
CLM-2 (arbitrary cytotechnician)						
Normal	338	7	1	2	0	348
ASC-US/AG-US	19	15	3	0	1	38
LSIL	1	2	3	0	0	6
HSIL/CIS/AG-FN-AIS	3	2	0	45	9	59
Cancer	1	0	0	8	38	47
Total	362	26	7	55	48	498

In bold and hatched, discordant cases with consequences in both treatment and follow-up. CLM: Conventional light microscopy, HSIL: High-grade squamous intraepithelial lesion, LSIL: Low-grade squamous intraepithelial lesion, AIS: Adenocarcinoma *in situ*, FN: Favor of neoplasia, ASC-US: Atypical squamous cells of undetermined significance, NOS: Not otherwise specified, AG-US: Atypical glandular cells of undetermined significance, CIS: Carcinoma in situ

classified to be ASC-US/AG-US by CLM-2. Furthermore, the high-grade lesions showed cases with minor discrepancies but with no consequences for treatment and follow-up policy between CLM-1 and CLM-2. In 8 cases from CLM1, classified as HSIL/AGC-FN/AIS, the second CLM reading resulted in cancer, and 9 CLM-1 cancer cases were diagnosed as HSIL by CLM-2. These cases with no differences in treatment or follow-up were regarded as concordant in the comparison.

Table 2 shows an overview of the final classifications of the external expert on the discordant cases of Cohort 1. The expert agreed in 3/11 cases with CLM-1 diagnoses, in 2/11 cases with CLM-2 diagnosis, and choose another Bethesda classification in 6 cases.

The concordance rate between CLM-1 and CLM-2 at the LSIL threshold was 97.8% (95% CI: 0.96–0.99). The general unweighted kappa using the same technic (CLM) was 0.75 (95% CI: 0.68–0.80). The concordance and the specific agreement at different thresholds are shown in Table 3. The concordance rates varied between 91.6% and 99.6%. Table 4 shows the kappa values of Cohort 1 at different thresholds.

Conventional light microscopy versus whole-slide image assessment

Table 5 shows the aggregated results of the CLM-3 versus WSI examinations of Cohort 2. After comparison of the WSI and the CLM-3 assessment, major discrepancies were found in 23/489 (4.7%) cases that should have resulted in different treatment and/or follow-up strategies. Three CLM-3 normal cases were diagnosed as LSIL in WSI. Out of 10 WSI normal cases, 5 were found to be LSIL, 3 HSIL, and 2 cancer in CLM-3. One ASC-US/AG-US case from WSI was diagnosed as HSIL/AG-FN/AIS in CLM-3. Two CLM-3 LSIL cases were diagnosed as HSIL/AG-FN/AIS in WSI. Five CLM-3 cases

Table 2: Expert diagnosis on the major discrepancies of Cohort 1 with double reading by conventional light microscopy (conventional light microscopy-1 and conventional light microscopy-2)

CLM-1	CLM-2	*n*=11	Expert classification
Normal	LSIL	1	Normal
Normal	HSIL/ AGC-FN/AIS	3	2 normal and 1 ASC-US
Normal	Cancer	1	AGC-FN/AIS
ASC-US/AG-US	HSIL/ AGC-FN/AIS	2	1 LSIL and 1 normal
LSIL	Normal	1	Normal
HSIL/AG-FON/AIS	Normal	2	1 normal 1 AGC-em
Cancer	ASC-US/AG-US	1	HSIL

CLM: conventional light microscopy, ASC-US : atypical squamous cells of undetermined significance, AG-US: atypical glandular cells of undetermined significance, LSIL: low grade squamous intraepithelial lesion, HSIL: high grade squamous intraepithelial lesion, AGC-FN: atypical glandular cells favour of neoplasia, CIS: carcinoma *in situ*, AIS: adenocarcinoma *in situ*, AGC-em: atypical glandular cells of the endometrium

Table 3: Concordance and the specific agreement rates at different thresholds between light microscopic (CLM-1) and double reading (CLM-2) of a cohort of 498 cervical cytology cases enriched for high-grade lesions

	CLM-2			
Threshold	Concordance rate	95% CI	PA	NA
CLM 1				
ASC-US/AG-US	91.6% (456/498)	88.8-93.7	0.47	0.95
LSIL	97.8% (487/498)	96.0-98.7	0.46	0.99
HSIL/CIS/AG-FN-AIS	98.2% (489/498)	96.6-99.1	0.79	0.97
Cancer	99.6% (496/498)	98.6-99.9	0.80	0.98

CLM: conventional light microscopy, ASC-US : atypical squamous cells of undetermined significance, AG-US: atypical glandular cells of undetermined significance, LSIL: low grade squamous intraepithelial lesion,HSIL: high grade squamous intraepithelial lesion, AGC-FN: atypical glandular cells favour of neoplasia, CIS: carcinoma *in situ*, AIS: adenocarcinoma *in situ*, CI: confidence interval, PA: agreement, NA: negative agreement

were diagnosed as LSIL in WSI. One CLM-3 cancer case was diagnosed as LSIL in WSI. Finally, one ASC-US/AG-US case from CLM-3 was diagnosed as HSIL/AG-FN/AIS in WSI. In total, there were 11 overcalls and 12 undercalls in WSI compared to CLM-3.

All 23 discordant cases were discussed in the consensus case meetings. Table 6 provides the results of these case meetings. As can be seen from the table in 11/23 of these cases, the CLM-3 diagnosis was preferred, in 7 the WSI diagnosis, and in 5 cases another diagnosis.

Figure 2 shows an example of a major discrepancy case discussed in the consensus case meeting.

The concordance rate between CLM-3 and WSI at the LSIL threshold was 95.3% (95% CI: 93.0%–96.9%). Kappa between the two methods was 0.67 (95% CI: 0.60–0.74). The concordance and the specific agreement rates at different thresholds are shown in Table 7. The concordance rates varied between 89.4% and 99.4%. Table 4 shows the kappa values of Cohort 2 at different thresholds.

Discussion

If we take into account that in noninferiority studies, a discordance of 4% or less is acceptable, the present results are within these limits.[14]

The use of WSI for cytology diagnosis is still subject of debate. Validation of the digital approach for primary diagnostics is an important primary step that needs to be made before this technique can be introduced with confidence in daily routine. Our previous study[9] showed that confidence in WSI-based cytology may be obtained after proper training, and this encouraged us to take the next step and do a comparative study of CLM and WSI. During the study design, we realized ourselves that before we could draw any conclusions from such a comparison, we had to know how well the diagnostic team of cytotechnicians would perform in double reading in a routine situation using CLM. According to the Dutch Guidelines for cervical screening[27] double readings are not meant to be done by the same technician but by a second experienced colleague. In our study design we have tried to avoid that specimens were judged twice by the same cytotechnician (CT) to stay as close as possible to the routine situation. This also explains why in this study we cannot pay attention to any intra-observer variations. In addition, we cannot address the inter-observer variation since not all technicians have scored all samples. We approached the study as much as possible as a routine diagnostic process which means that the study samples were distributed in a similar way as the daily routine cases. A limitation of the study design is that we choose to investigate two cohorts enriched with HSIL, cancer, and normal cases with a historical slide registration number which could be identified by the CT. This limitation results in an increase of vigilance of the CT to identify these high-grade lesions. This increase in vigilance has a restricted effect on the applicability of the use of WSI in a true-screening setting. In other respects, the enrichment gives information about the capability and experience in identifying these lesions by the CT in WSI. The CT becomes experienced in identifying high-grade lesions. This would not happen if we totally simulate a real screen situation. What we did want to know is the similarity in performance of the diagnostic team using CLM or WSI in primary diagnostics of cervical thin-layer slides. This is measured by a double reading CLM and a double reading CLM versus WSI.

The present study showed that on WSI scanned at ×20 on a single focus plane, a concordance rate with CLM of 89.4%–99.4% (depending on the threshold chosen) was obtained for cervical cytology, comparable with the CLM versus CLM concordance rate of 91.6%–99.6%, which is in

Table 4: Kappa values of Cohort 1 and Cohort 2 at different thresholds

Threshold	κ			
	Cohort 1	95% CI	Cohort 2	95% CI
Normal	0.83 (near perfect)	0.78-0.89	0.82 (near perfect)	0.76-0.87
ASC-US/AG-US	0.43 (moderate agreement)	0.25-0.62	0.34 (fair agreement)	0.11-0.57
LSIL	0.45 (moderate agreement)	0.05-0.86	0.12 (slight agreement)	0.0-0.47
HSIL/CIS/AG-FN-AIS	0.76 (substantial agreement)	0.67-0.86	0.66 (substantial agreement)	0.55-0.77
Cancer	0.78 (substantial agreement)	0.68-0.87	0.67 (substantial agreement)	0.55-0.0.79

ASC-US : atypical squamous cells of undetermined significance, AG-US: atypical glandular cells of undetermined significance, LSIL: low grade squamous intraepithelial lesion, HSIL: high grade squamous intraepithelial lesion, AGC-FN: atypical glandular cells favour of neoplasia, CIS: carcinoma *in situ*, AIS: adenocarcinoma *in situ*

Table 5: Cross table showing the results of the assessments by conventional light microscopy and whole slide images

	CLM-3 (arbitrary cytotechnician)					Total
	Normal	AG-US	LSIL	HSIL/AG-FN-AIS	Cancer	
WSI (arbitrary cytotechnician)						
Normal	333	13	5	3	2	356
ASC-US/AG-US	9	9	6	1	0	25
LSIL	3	1	2	5	1	12
HSIL/CIS/AG-FN-AIS	0	1	2	40	17	60
Cancer	0	0	0	6	30	36
Total	346	23	15	55	50	489

In bold and hatched discordant cases with consequences in both treatment and follow-up. CLM: conventional light microscopy, WSI: Whole-slide imaging, ASC-US: atypical squamous cells of undetermined significance, AG-US: atypical glandular cells of undetermined significance, LSIL: low grade squamous intraepithelial lesion, HSIL: high grade squamous intraepithelial lesion, AGC-FN: atypical glandular cells favour of neoplasia, CIS: carcinoma in situ, AIS: adenocarcinoma *in situ*

Table 6: The results of the discordant cases between conventional light microscopy and whole slide image-based reading discussed in the consensus whole slide image-based case meetings

CLM-3	WSI	*n*=23	Consensus classification*
Normal	LSIL	3	1 ASC-US and 2 normal
ASC-USUS-AG-US	HSIL/AGC-FN/AIS	1	HSIL
LSIL	Normal	5	3 normal and 2 LSIL
LSIL	HSIL	2	1 AGC-FN/AIS and 1 HSIL
HSIL/AGC-FN/AIS	Normal	3	1 no consensus** and 2 HSIL
HSIL/AGC-FN/AIS	ASC-US/AG-US	1	1HSIL
HSIL/AGC-FN/AIS	LSIL	5	4 HSIL and 1 LSIL
Cancer	Normal	2	1 no consensus** and 1 HSIL
Cancer	LSIL	1	HSIL

*Consensus classification after discussion in the weekly meeting, **No consensus could be reached on this difficult case. CLM: conventional light microscopy, WSI: whole- slide imaging, ASC-US : atypical squamous cells of undetermined significance, AG-US: atypical glandular cells of undetermined significance, LSIL: low grade squamous intraepithelial lesion,HSIL: high grade squamous intraepithelial lesion, AGC-FN: atypical glandular cells favour of neoplasia, CIS: carcinoma in situ, AIS: adenocarcinoma *in situ*

line with Rowe *et al.*[28] This means that even on WSI at ×20 and a single focal plane, with cellular information partly out of focus, excellent diagnostic results can be obtained with cervical cytology WSI. We expect that after gaining more experience in WSI (the CT now has an experience of several years in CLM and only a few 100 cases for WSI), the concordance in WSI double readings also increases. From the results of the kappa calculations of Cohorts 1 and 2, it appeared that WSI has an effect on the reproducibility of CT compared to CLM at thresholds LSIL and ASC-US, probably due to the small number of ASC-US and LSIL cases in both cohorts[29] making kappa calculations less reliable. At the thresholds normal and HSIL with sufficient numbers of cases, kappa did not differ much in Cohort 2. After analyzing the discrepant cases of both cohorts, it appeared that for Cohort 1, there were more (7) overcalls (false positive) than undercalls (false negative) (4). For Cohort 2, there were one more undercalls (12) than overcalls (11). It seems from the results that WSI is less sensitive compared to CLM. These numbers are too low to draw any significant conclusions. We realize that before the routine implementation of WSI as a new screening tool, steps have to be made forward in the field of quality assurance thinking of more experience and trust in viewing WSI, more standardization of the thin-layer slide preparation technique, optimizing the scan utilities of the scanners. More comparative research has to be done with a larger number of cases to determine the influence of false-positive and false-negative rates. To decrease these rates for CLM and for WSI by CT and pathologists, more practical training and education are needed.[20,30-33] Overall and compared to the published literature on the subject, 2 cohorts of approximately 500 cases each which were enriched for high-grade cases comprise the largest study group used in

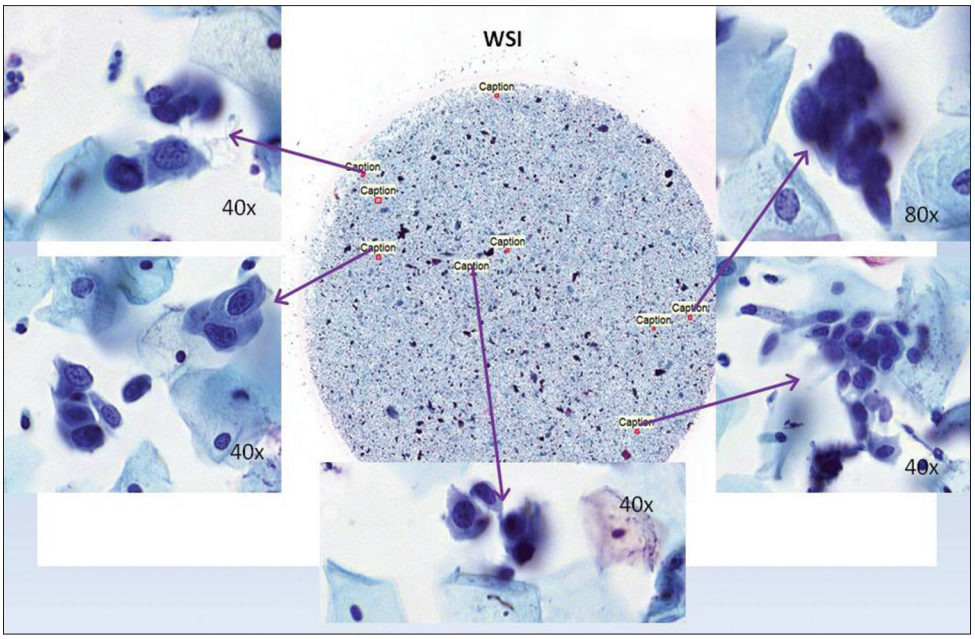

Figure 2: Example of a major discrepant CLM/WSI case discussed in the case meeting. In this case, the differential diagnosis was between, ASC-US (WSI) and moderate dysplasia (CLM 3). Parabasal or squamous metaplastic cells are seen with enlarged nuclei (1½ to 2 times the size of an intermediate squamous cell nucleus). The nuclei show some irregularities and coarse granular chromatin distribution. The N/C ratio is disturbed, grooves are present in some nuclei. Consensus diagnosis was moderate dysplasia on the WSI. CLM: Conventional light microscopy, WSI: Whole-slide imaging, ASC-US: atypical squamous cells of undetermined significance

Table 7: Concordance and specific agreement rates at different thresholds between light microscopic original (conventional light microscopy-3) and whole slide image-based second reading (whole slide images) of a cohort of 489 cervical cytology cases enriched for high-grade lesions

Threshold	WSI			
	Concordance rate	95% CI	PA	NA
CLM 3				
ASCUS/AG-US	89.4% (437/489)	86.3-91.8	0.38	0.98
LSIL	95.3% (466/489)	93.0-96.9	0.15	0.98
HSIL/CIS/AG-FN-AIS	98.4% (481/489)	96.8-99.2	0.69	0.99
Cancer	99.4% (486/489)	98.2-99.8	0.70	0.95

CLM: conventional light microscopy, WSI: Whole slide imaging, ASC-US: atypical squamous cells of undetermined significance, AG-US: atypical glandular cells of undetermined significance, LSIL: low grade squamous intraepithelial lesion,HSIL: high grade squamous intraepithelial lesion, AGC-FN: atypical glandular cells favour of neoplasia, CIS: carcinoma *in situ*, AIS: adenocarcinoma *in situ*, CI: confidence interval, PA: agreement, NA: negative agreement

digital cytology so far. The results of our study allow to give directions for further steps to prepare the introduction of digital cytology as a routine diagnostic technique.

Our study complies with CAP guidelines for validation of WSI in primary diagnostics.[2,11,13] In our normal cervical screening situation[27] cervical slides are evaluated by cytotechnicians with demonstrable experience and certification in cervical cytology. In situations of abnormal blood loss, abnormal portio, abnormal cells and/or diagnostic difficulties the slides were evaluated by a different second cytotechnician.

At the time of this writing, there were few published studies investigating the use of WSI in primary cytology diagnostics.[8,10,18-23,30,31] Steinberg and Ali achieved an accuracy of 98.3% comparing WSI with CLM examination by six participants of ten selected thin-layer cervical slides.[19] A next study investigated the accuracy of 24 experienced cytologists in the examination of WSI from 24 cervical thin-layer slides which appeared to be equal to glass slide microscopy when participants were able to focus through the full thickness of the slide.[8] A third study investigated the accuracy and efficiency of WSI in primary diagnostics of cervical cytology by four cytotechnicians and three pathologists and found that the accuracy of interpretation and time needed per case was superior for glass slides.[24] Except for the latter, these studies and ours show that WSI can well be used in primary diagnosis of cervical cytology. However, some technical issues need to be addressed before WSI can become part of a daily work process of screening cervical cytology: scanning time, number of focus levels, ease of panning and zooming through z-stack, storage capacity, and associated costs. Further, there are logistic issues: The familiarity with and the time needed for evaluating WSI of thin-layer cervical slides. Finally, training and building up experience have definitely contributed to the relative success of our cervical WSI approach. We faced initial reluctance in using WSI in our laboratory like others.[8,10,11,18,22,24] As described in our previous paper, we found a way to cope with human resistance by organizing case meetings to get experience in

viewing WSI and to document cytomorphological features of classical lesions and abnormalities of the cervix as can be identified in WSI. In this way, confidence in viewing WSI was built. As a spinoff of the case meetings in this study, we created a website (www.ex-pathcytology.com) for training purposes or support in diagnostics or further studies.[9] The experience and involvement of the cytotechnicians created in this way was a solid base for the current study. We also expect that continuous educational meetings will improve the experience of our diagnostic team and improve the correct classification of low-grade lesions. Further research should be undertaken investigating inter-observer and intra-observer variation to determine whether WSI is indeed equivalent to CLM in primary diagnostics of thin-layer cervical slides.

After the final judgment of the external pathologist on the discordant cases of Cohort 1, it turned out that in almost half of the discordant cases, a different diagnosis was preferred. This underlines the fact that cytological interpretation is highly dependent on the interpretation of the cells in the slides, the experience in viewing cervical cytology, and training of the cytotechnician or pathologist in viewing cervical cytology.[30] These arguments were also highlighted in a study of Gutierrez where low-to-good concordance rates were obtained between four cytotechnicians and one pathologist on fifty CLM cases.[31] Next to this, there was fading of staining in the conventional slides in the course of time, and there were administrative errors.

In the consensus case meetings for the discordant cases from Cohort 2, it became clear that not every WSI was scanned on the right focus level for the three-dimensional (3D) cell groups, and not every thin-layer slide was really "thin." This former issue could be tackled by using z-stacks that could help in the examination of the 3D cell groups. For example, in one case, the cancer was missed because of the lack of focusing of the 3D cell groups in the WSI. The latter issue can be dealt with by further standardization of the thin-layer SurePath preparation technique and perhaps better diluting samples. We changed the sample volume settings from 200 to 175 µl, but this still did not result in real thin layers in all cases. Both issues will be the subject of further studies.

Another issue is the scanning magnification. In our study, we choose to scan all slides at ×20 instead of ×40. Our arguments in favor of the ×20 scanning magnification were: The focal point method based imaging system of Sure-Path slides (Becton Dickinson, Franklin Lakes NJ USA) works at ×20, the ×20 is the standard magnification of 3DHistech scanner but also has an extra enlargement of 1.6 (Camera adapter magnification) what results in a higher magnification(×32) than the bare ×20 and the extra storage needed scanning at ×40 of 500 thin-layer slides compare to ×20 scans. Arguments against the ×20 were that in the study of Wright *et al*.,[24] it appeared that the diagnostic accuracy increased when evaluating ×40 scans compared to ×20 BioImagene scans. At present, there are, however, no comparative studies between different scanners, so there is yet no evidence on a possible superior image quality of certain scanners above others.

Conclusion

We expect that WSI has the potential to become a diagnostic tool in a routine laboratory setting. However, larger studies more closely mimicking real practice are needed to support these expectations. Proper training and continuous education are important cornerstones. WSI Cytology opens up new possibilities creating virtual work spots for cytologists: to work remotely, to share, to educate, for proficiency testing, easy and fast review of previously imaged archived cases, which are not only limited to cervical cytology. Further, it may help to deal with the shortage of skilled cytotechnologists in underserved areas in the world.

Acknowledgments

We would like to thank Dr. Ruud Clarijs, Dr. Robert Riedl, Dr. Paul Theunissen, and Dr. Bart de Vries, for their professional diagnostic contribution during the WSI consensus-based case meetings.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

References

1. Al Janabi S, Huisman A, van Diest PJ. Digital Pathology: Current status and future perspectives. Histopathology 2012;61:1-9.
2. Pantanowitz L, Sinard JH, Henricks WH, Fatheree LA, Carter AB, Contis L, *et al*. Validating Whole Slide Imaging for diagnostic purposes in Pathology. Arch Pathol Lab Med 2013;137:1710-22.
3. 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 2014;68:33-9.
4. Farahani N, Parwani AV, Pantanowitz L, Whole slide imaging in pathology: advantages, limitations and emerging perspectives. Pathol Lab Med Int 2015;7:23-33.
5. 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.
6. Goacher E, Randell R, Williams B, Treanor D. The diagnostic concordance of Whole Slide Imaging and light microscopy. Arch Pathol Lab Med 2017;141:151-61.
7. Bueno G, Fernández-Carrobles MM, Deniz O, García-Rojo M. New Trends of Emerging Technologies in Digital Pathology. Pathobiology 2016;83:61-9.
8. Evered A, Dudding N. Accuracy and perceptions of virtual microscopy compared with glass slide microscopy in cervical cytology. Cytopathology 2011;22:82-7.
9. Bongaerts O, van Diest PJ, Nap M, Working towards consensus among professionals in the identification of classical cervical cytomorphological characteristics in whole slide imaging. J Pathol Inform 2015;6:52.
10. 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.
11. 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;68:449-505.
12. Association DP. DPA recommends whole slide imaging manufacturers submit the NOVO applications to the FDA for primary diagnosis in the United States.
13. College of American Pathologists. FDA open to whole slide imaging class-II device. 2016.

14. 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.
15. van der Post RS, van der Laak JA, Sturm B, Clarijs R, Schaafsma HE. The evaluation of colonbiopsies using virtual microscopy is reliable. Histopathology 2013;63:114-21.
16. 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.
17. Saco A, Ramírez J, Rakislova N, Mira A, Ordi J. Validation of whole-slide imaging for histolopathological diagnosis: current state. Pathobiology 2016;83:89-98.
18. Mukhopadhyay S, Feldman MD, Abels E, Ashfaq R, Beltaifa S, Cacciabeve NG, *et al*. Whole Slide Imaging Versus microscopy for Primary Diagosis in Surgical Pathology: A Multicenter Blinded Randomized Noninferiority study of 1992 cases (pivotal study). Am J Surg Pathol 2018;42:39-52.
19. Steinberg DM, Ali SZ, Application of virtual microscopy in clinical cytopathology. Diagnostic cytopathology 2001;25:389-96.
20. Khalbuss WE, Pantanowitz L, Parwani AV. Digital imaging in cytopathology. Patholog Res Int 2011;2011:264683. [Epub 2011 Jul 19].
21. Lee RE, McClintock DS, Laver NM, Yagi Y. Evaluation and optimizaion for liquid based preparation cytology in whole slide imaging. J Pathol Inform 2011;2:46.
22. Gerhard R, Teixeira S, Gaspar da Rocha A, Schmitt F. Thyroid fine-needle aspiration cytology: is there a place to virtual cytology. Diagn Cytopathol 2013;41:793-8.
23. Lahrmann B, Valous NA, Eisenmann U, Wentzensen N, Grabe N. Semantic focusing allows fully automated single-layer slide scanning of cervical cytology slides. PLoS One 2013;8:e61441.
24. Wright AM, Smith D, Dhurandhar B, Fairley T, Scheiber-Pacht M, Chakraborty S, *et al*. Digital Slide imaging in cervicovaginal cytology; A pilot study. Arch Pathol Lab Med 2013;137:618-24.
25. Nayar R, Wilbur DC. The PAP Test and Bethesda 2014. Acta Cytologica 2015;59:121-32.
26. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33.159-74.
27. National institute for public health and environment (RIVM). Avaialble from: www.rivm.nl /;2016.
28. Rowe LR, Marshall CJ, Bentz JS. One hundred percent thourough quality control rescreening of liquid based monlayers in cervicovaginal cytopathology. Cancer 2002;96:325-9.
29. Verhagen A. Alessie J. Evidence based dignostiek van het bewegingsapparaat. Bohn Stafleu van Loghum, onderdeel van Springer Media BV. 2014 15-28.
30. Branca M, Longatto-Filho A, Recommendations on quality control and quality assurancy in cervical cytology. Acta Cytol 2015;59:361-9.
31. Gutiérrez-Enríquez SO, Chávez-Hernández L, Terán-Figueroa Y, Gaytán-Hernández D, Oros-Ovalle C, Gallegos-García V, *et al*. Concordance in the interpretation of cervical cytology for the early diagnosis of cervical cancer. OJOG 2016;6:714-24.
32. Bigras G, Wilson J, Russell L, Johnson G, Morel D, Saddik M. Interobserver concordance in the assessment of features used for diagnosis of cervical atypical squamous cells and squamous intraepithelial lesions (ASC-US, ASC-H, LSIL and HSIL). Cytopathol. 2013;24:44-51.
33. Wiener HG, Klinkhamer P, Schenck U, Arbyn M, Bulten J, Bergeron C, *et al*. European guidelines for quality assurance in cervical cancer screening: recommendations for cytology laboratories. Cytopathology 2007;18:67-78.