Original Article

Inter- and intraobserver agreement in whole-slide digital ThinPrep samples of low-grade squamous lesions of the cervix uteri with known high-risk HPV status: A multicentric international study

Ivana Kholová, MD, PhD, MIAC [1,2]; Giovanni Negri, MD [3]; Maria Nasioutziki, MD, PhD, MIAC[4]; Laura Ventura, PhD[5]; Arrigo Capitanio, MD[6]; Massimo Bongiovanni, MD, FIAC [7]; Paul A. Cross, MD[8]; Claire Bourgain, MD, PhD, MIAC[9]; Henrik Edvardsson, MD, PhD[10]; Rosario Granados, MD, PhD, FIAC [11]; Artur Lipiński, MD, PhD[12]; Ellen Christina Obermann, MD[13]; Maurizio Pinamonti, MD[14]; Henrieta Sidlova, MD, PhD[15]; Margareta Strojan Fležar, MD, PhD, MIAC[16]; Folkert J. van Kemenade, MD, PhD[17]; Danijela Vrdoljak-Mozetic, MD, PhD, MIAC[18]; Ambrogio Fassina, MD, PhD [19]; and Beatrix Cochand-Priollet, MD, PhD, MIAC[20]

BACKGROUND: High-risk human papilloma virus (HR HPV) testing and liquid-based cytology are used for primary cervical screening. Digital cytology, based on whole-slide scanned samples, is a promising technique for teaching and diagnostic purposes. The aim of our study was to evaluate the interobserver and intraobserver variation in low-grade squamous lesions, HR HPV status bias, and the use of whole-slide scanned digital cervical cytology slides. **METHODS:** Fifteen expert cytopathologists evaluated 71 digitalized ThinPrep slides (31 atypical squamous cells of undetermined significance [ASC-US], 21 negative for intraepithelial lesion or malignancy, and 19 low-grade squamous intraepithelial lesion cases). HR HPV data were accessible only in the second round. **RESULTS:** In interobserver analysis, Kendall's coefficient of concordance was 0.52 in the first round and 0.58 in the second round. Fleiss' kappa values were 0.29 in the first round and 0.31 in the second round. In the ASC-US category, Fleiss kappa increased from 0.19 to 0.22 in the second round and the increase was even higher expressed by Kendall's coefficient: from 0.42 to 0.52. In intraobserver analysis, personal scores were higher in the second round. **CONCLUSIONS:** The interobserver and intraobserver variability in low-grade squamous lesions was within fair agreement values in the present study, in line with previous works. The comparison of two rounds showed that expert cytopathologists are generally unbiased by the knowledge of HR HPV data, but that being informed of the HR HPV status leads to a better agreement. Stain quality and back discomfort were highlighted as factors affecting digital cytopathology use. ***Cancer Cytopathol* 2022;130:939-948**.

KEY WORDS: ASC-US; cervical cancer; digital cytopathology; HPV; interobserver agreement; intraobserver agreement.

Corresponding Author: Ivana Kholová, MD, PhD, MIAC, Pathology, Fimlab Laboratories Oy, Arvo Ylpön katu 4, 33520 Tampere, Finland (ivana.kholova@tuni.fi).

[1]Tampere University, Faculty of Medicine and Health Technology, Tampere, Finland; [2]Pathology, Fimlab Laboratories, Tampere, Finland; [3]Pathology Unit, Central Hospital Bolzano, Italy; [4]Faculty of Medicine, Aristotle University of Thessaloniki and Molecular Cytopathology Department, Hippokration General Hospital, Thessaloniki, Greece; [5]Department of Statistical Sciences, University of Padova, Padova, Italy; [6]Department of Pathology, University Hospital Linköping, Linköping, Sweden; [7]Synlab Pathology, Lausanne, Switzerland; [8]South of Tyne Pathology Department, Queen Elizabeth Hospital, Gateshead, England; [9]Imelda Hospital, Bonheiden, Belgium; [10]Department of Pathology and Cytology, Karolinska University Laboratory, Karolinska Hospital Huddinge, Stockholm, Sweden; [11]University Hospital de Getafe, Madrid, Spain; [12]Department of Pathology and Oncological Cytology, Wroclaw Medical University, Wroclaw, Poland; [13]University Basel, Basel, Switzerland; [14]University Hospital Cattinara, Trieste, Italy; [15]St. Elisabeth Cancer Institute Hospital, Bratislava, Slovak Republic; [16]Faculty of Medicine, University of Ljubljana, Institute of Pathology, Ljubljana, Slovenia; [17]Erasmus MC University Medical Center, Department of Pathology, Rotterdam, the Netherlands; [18]Department of Pathology and Cytology, Clinical Hospital Centre Rijeka, Rijeka, Croatia; [19]Department of Medicine (DIMED), Surgical Pathology & Cytopathology Unit, University of Padova, Padova, Italy; [20]Cochin Hospital, University of Paris, Paris, France

Part of the study was presented as a platform presentation at European Congress of Cytology October 3–11, 2021 and was presented as a poster at 69th Annual Scientific Meeting of the American Society of Cytopathology November 11–14, 2021.

Received: April 7, 2022; **Revised:** June 7, 2022; **Accepted:** June 20, 2022

Published online July 14, 2022 in Wiley Online Library (wileyonlinelibrary.com)

DOI: 10.1002/cncy.22624, wileyonlinelibrary.com

INTRODUCTION

Since George Papanicolaou established the role of the Papanicolaou (Pap) smear in cervical cancer detection and prevention,[1] several milestones have happened in cervical cytology. In particular, the introduction of liquid-based cytology (LBC) and more recently high-risk human papilloma virus (HR HPV) molecular tests have had a decisive impact on cervical cancer screening.[2] In Europe, many countries have implemented HR HPV primary screening programs either regionally or nationally.[3–8] In the era of HR HPV primary screening, the number of cervical cytology slides has dramatically decreased, but there is a relative increase in atypical cytological findings at least in the first screening round.[6,7]

Low-grade changes of the squamous epithelia include both atypical squamous cells of undetermined significance (ASC-US) and low-grade squamous intraepithelial lesion (LSIL). The ASC-US category is defined by The Bethesda System for Reporting Cervical Cytology (TBS)[9] as changes suggestive but not sufficient for an LSIL interpretation. Despite defined criteria of ASC-US and particularly LSIL, several studies have shown that the intra- and interobserver agreement of these categories is low.[10,11] In fact, all aspects of cervical cancer diagnostic workup (i.e., cytology, histology, and colposcopy) show interpretive variability.[10–13]

Because most screening detected lesions are low-grade lesions, particularly after the first HR HPV screening round, and low-grade lesions are among those with the poorest agreement, the issue is one of high importance for cytologists and pathologists involved in screening programs.

Digital cytology has been widely used in cytology external quality assurance programs[14–18] and cytology education and training,[19–21] but less so in routine cytology practice. Despite revolutionary developments such as the creation of large digital data sets coupled with deep learning methods and artificial intelligence, the use of digital cytology has lagged behind its use in surgical pathology. There are several technical challenges related to cytology image acquisition such as focusing on three-dimensional cell groups and navigating slides at high magnification when screening impairing cytology workflow, the large file size and scanning times for whole-slide images in cytology.[22–25] Last, but not least, education in the use of digital cytology and ergonomics issues[26] should also be taken into the account.

The aims of the present study were (1) to evaluate the interobserver and intraobserver agreement in low-grade squamous lesions and (2) to evaluate HR HPV status biases among expert cytopathologists with HR HPV screening and/or triaging experience, and (3) to assess the ease of use and compliance with the digital platform.

MATERIALS AND METHODS

Selection and characterization of cases

Three expert pathologists (B.C.P., M.N., G.N.) retrieved 71 cervical LBCs (ThinPrep, Hologic, Marlborough, Massachusetts, USA) slides from the archives of their laboratories. The original cytomorphological diagnosis according to TBS were ASC-US in 31 cases, NILM in 21, and LSIL in 19. According to HR HPV status, there were 20 HR HPV–positive and 11 HR HPV–negative ASC-US cases, 16 HR HPV–positive and 3 HR HPV–negative LSIL cases, and 11 HR HPV–positive and 10 HR HPV–negative NILM cases, respectively. All enrolled cases had known HR HPV status and either histology or cytology follow-up or 3 years of clinical follow-up. All LSIL cases had histology follow-up. Follow-up details are in Figure 1.

Scanning and digital platform

The slides were anonymized, assigned a unique study number, scanned by a Hamamatsu scanner at 40× magnification without any depth of focus (Hamamatsu Photonics, Hertfordshire, UK), and posted on internet platform (Eurocytology, www.eurocytology.cloud) accessible by individual password.

All slides were annotated with regions of interest with a total of 576 annotations (mean, 6.12 [range, 2–16] annotations per case).

The evaluation form included patient age and brief clinical data. Dedicated medical grade computer screen was used only by two participants (13%). HR HPV status was available for the observers only in the second round. The participants coded the cases according to TBS categories for squamous epithelia (0 negative, 1 ASC-US, 2 LSIL, 3 atypical squamous cells, cannot exclude HSIL (ASC-H), 4 high-grade squamous intraepithelial lesion [HSIL]).

Rounds and participants

The first round was opened in January 2021 and the second round in March 2021. Each round was open for 6 weeks. Participants were 15 expert cytopathologists

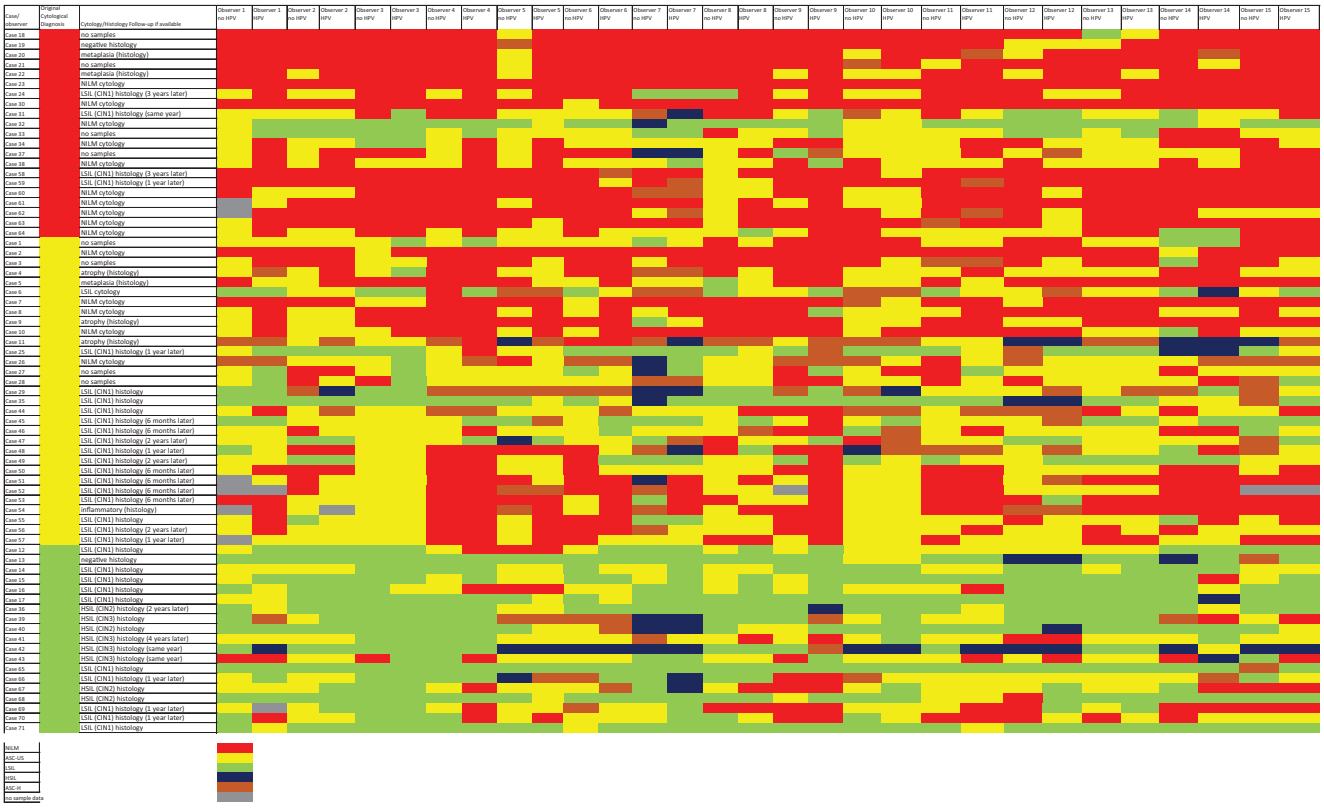

FIGURE 1. Heat map summarizing all answers in both rounds according to diagnoses. ASC-H indicates atypical squamous cells – cannot exclude high-grade squamous intraepithelial lesion; ASC-US, atypical squamous cells – undetermined significance; HR HPV, high-risk human papilloma virus; HSIL, high-grade squamous intraepithelial lesion; LSIL, low-grade squamous intraepithelial lesion; NILM, negative for intraepithelial lesion or malignancy.

(eight females, seven males) from 13 European countries. The cytopathology, digital pathology, HR HPV screening/triaging experience, and compliance data were collected with questionnaires.

Statistical methods

Categorical variables were expressed as frequencies (and percentages) and numerical variables as means (and SD). Categorical variables were expressed as frequencies (and percentages) and numerical variables as means (and SD). Interobserver agreement was assessed by Fleiss' kappa and Kendall's coefficient of concordance between all raters in two rounds. Bootstrap resampling was used to compute the SDs of the indices. To compare the indices in the two rounds (i.e., intraobserver variation for individual raters), the methods by Feldt, Woodruff, and Salih[27] were used to provide p values. The free software R (http://www.r-project.org/) was used for statistical analyses. The level of significance was set at $p < .05$.

The strength of association is defined for Fleiss' kappa as follows: 1 for perfect agreement, 0.81–0.99 for almost perfect agreement, 0.61–0.80 for substantial agreement, 0.41–0.60 for moderate agreement, 0.21–0.40 for fair agreement, 0–0.20 slight agreement, and < 0 poor agreement. The strength of association is defined for Kendall's coefficient of concordance as follows: ± 0.30 or above strong, ± 0.20 to 0.29 moderate, ± 0.10 to 0.19 weak, and less than ± 0.10 very weak.

RESULTS

Participants

Eleven participants (73.3%) were from an academic hospital, three (20.0%) from other hospitals, and one (6.7%) from a private laboratory. The cytopathology practice of the participants was 24.20 ± 7.95 years on average ranging from 8 to 38 years. LBC experience was on average 13.47 ± 8.28 years with the longest period being 30 years.

TABLE 1. Fleiss' kappas/Kendall's coefficients in total and in all three TBS categories according to the first and second rounds

Diagnostic category	First round (unknown HR HPV status) Fleiss' kappa	First round (unknown HR HPV status) Kendall's coefficient	Second round (known HR HPV status) Fleiss' kappa	Second round (known HR HPV status) Kendall's coefficient
All	0.29 ± 0.03	0.52 ± 0.04	0.31 ± 0.03	0.58 ± 0.04
NILM	0.32 ± 0.05	0.62 ± 0.09	0.28 ± 0.07	0.46 ± 0.08*
ASC-US	0.19 ± 0.04	0.42 ± 0.05	0.22 ± 0.04	0.52 ± 0.05
LSIL	0.28 ± 0.06	0.47 ± 0.07	0.16 ± 0.04	0.38 ± 0.07

Abbreviations: ASC-US, atypical squamous cells – undetermined significance; HR HPV, high-risk human papilloma virus; LSIL, low-grade squamous intraepithelial lesion; NILM, negative for intraepithelial lesion or malignancy; TBS, The Bethesda System for Reporting Cervical Cytology.
*$p = .02$.

Cervical cytopathology on average comprised 33% of participants´ workload (with up to 70%) with an annual average of 5967 ± 4870.84 reviewed slides (range, 300–15000 slides).

The average experience time with the use of HR HPV in a screening context was 14.93 ± 5.91 years, with a range of 5 to 30 years. In particular, HR HPV primary screening practical experience was on average 1.50 ± 1.54 years, with the longest experience being 5 years. HR HPV triage experience was on average 11.07 ± 6.83 years with the longest experience of 25 years.

The washout period between the two rounds was 5.1 weeks on average.

Interobserver agreement

Kendall's coefficient of concordance was 0.52 ± 0.04 in the first round and 0.58 ± 0.04 in the second round. Fleiss´ kappa values were 0.29 ± 0.03 in the first round and 0.31 ± 0.03 in the second round (Table 1; Figures 2 and 3).

Senior observers, characterized as those with a working experience of more than 20 years, rated with Kendall's coefficient of concordance 0.54 in the first round and 0.53 in the second round. Fleiss´ kappa values for senior experts were 0.30 in the first round and 0.31 in the second round.

The agreement evaluated as Fleiss´ kappa values/Kendall's coefficients for all three Bethesda categories separately (i.e., for NILM, ASC-US and LSIL in the first and second rounds) is shown in Table 1. Importantly, the highest agreement was seen for the NILM category followed by low-grade dysplasia. In numbers, Fleiss´ kappa values ranged from 0.28 ± 0.07 to 0.32 ± 0.05 and Kendall's coefficient ranged from 0.46 ± 0.08 to 0.62 ± 0.09 in NILM category. Fleiss´ kappa ranged from 0.19 ± 0.04 to 0.22 ± 0.04 and Kendall's coefficient ranged from 0.42 ± 0.05 to 0.52 ± 0.05 for the ASC-US category. Finally, Fleiss´ kappa values ranged from 0.16 ± 0.04 to 0.28 ± 0.06 and Kendall's coefficient ranged from 0.38 ± 0.07 to 0.47 ± 0.07 for the LSIL category. A heat map summarizing all the respondent's answers in both rounds according to diagnoses is shown in Figure 1. There is less uniformity in ASC-US cases with more diverse answers in comparison to NILM and LSIL cases. Single cases with lower agreement were nevertheless present in all categories: case 32 in the NILM category with predominantly LSIL answers, case 2 in the ASC-US category with mainly NILM answers, and

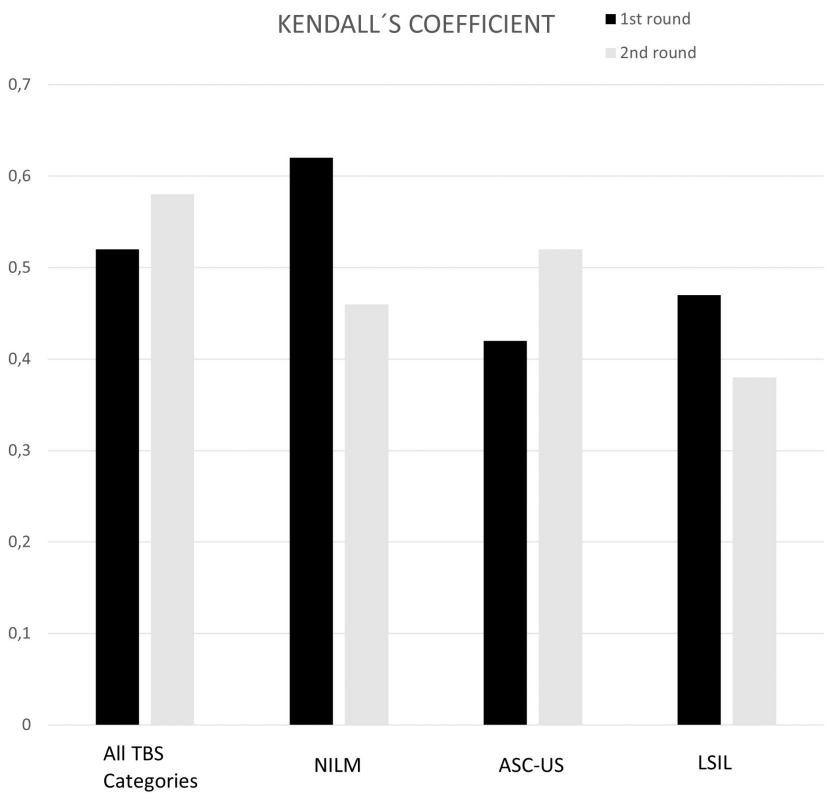

FIGURE 2. Kendall's coefficients in the first round without HR HPV status and in the second round with HR HPV status knowledge in all cases and according to TBS categories. ASC-US indicates atypical squamous cells – undetermined significance; HR HPV, high-risk human papilloma virus; LSIL, low-grade squamous intraepithelial lesion; NILM, negative for intraepithelial lesion or malignancy; TBS, The Bethesda System for Reporting Cervical Cytology.

case 42 in the LSIL category with some HSIL answers. Few ASC-US cases had also trended to be answered as ASC-H category.

Observers 7, 10, and 14 tended to more overdiagnosing and observers 4 and 9 to underdiagnosing. Nine (60%) participants admit overdiagnosing in uncertain cases.

Intraobserver agreement: HR HPV bias

In the NILM category, the knowledge of HR HPV status statistically decreased the Kendall's coefficients (0.62 ± 0.09 in the first round vs. 0.46 ± 0.08 in the second round, $p = .02$). HR HPV status also influenced categorization in mild lesions because the agreement in the ASC-US category improved with the knowledge of the HR HPV status: Fleiss kappa increased from 0.19 ± 0.04 without the HR HPV status knowledge to 0.22 ± 0.04 with a known HR HPV result and the increase was even higher with Kendall's coefficient: from 0.42 ± 0.05 to 0.52 ± 0.05. On the other hand, the LSIL category decreased in agreement with HR HPV status knowledge: Fleiss´ kappa decreased from 0.28 ± 0.06 without HR HPV information to 0.16 ± 0.04 with the known status and Kendall's coefficient decreased from 0.47 ± 0.07 without knowledge of the HR HPV status to 0.38 ± 0.07 with it. Overall, personal scores were higher (in better agreement) in the second round, when the HR HPV status was known: the statistical significance was $p = .004753$ for Fleiss´ kappa and $p = .0005778$ for Kendall's coefficient.

Digital experience and compliance with digital platform

The majority of participants already had some experience with digital pathology ($n = 12$, 80%), mainly from research (histopathology, 80%; cytopathology, 73%), and, to a lesser extent, from routine work (histopathology, 27%; cytopathology, 13%).

The digital cytology experience and compliance in the present study was evaluated on average (scale 1–5, with 1 being the best) and the results are as follows: scores 2.8 for time consuming in comparison with

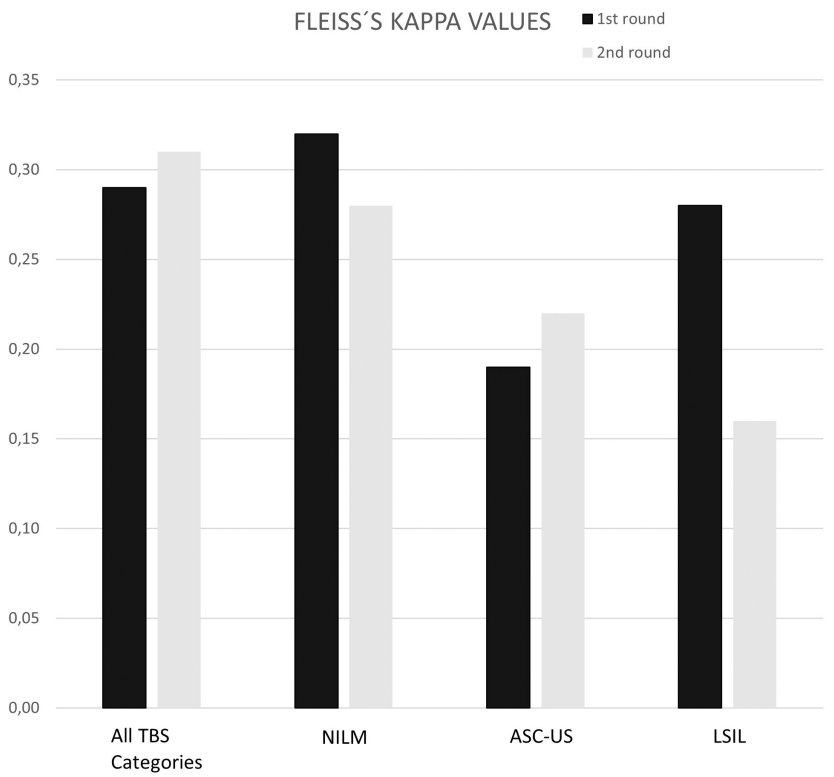

FIGURE 3. Fleiss´ kappa values in the first round without HR HPV status and in the second round with HR HPV status knowledge in all cases and according to TBS categories. ASC-US indicates atypical squamous cells – undetermined significance; HR HPV, high-risk human papilloma virus; LSIL, low-grade squamous intraepithelial lesion; NILM, negative for intraepithelial lesion or malignancy; TBS, The Bethesda System for Reporting Cervical Cytology.

conventional microscopy, 3.0 for eye discomfort, and 3.6 for back discomfort. Concerning technical aspects (scale 1–5, 1 every case, 5 no case), problems with image sharpness/focusing were rated as 3.7, problems with stain quality as 4.4, and problems with scan quality as 3.9. Overall, 47% of participants would recommend digital cytology in their department based on the experience in the present study.

DISCUSSION

Since George Papanicolaou's era, cervical cancer screening has changed massively: conventional smears have been widely replaced by HR HPV testing and liquid-based cytology, whereas digitalization is regarded as a very promising technique for the future. In the modern setting of an organized, HR HPV-based cervical screening, most screening detected lesions are incidental low-grade lesions, particularly after the first HR HPV-screening round.[28] Because low-grade lesions are among those with the poorest agreement, this is a potential issue for cytologists and pathologists involved in screening programs. The knowledge of the HR HPV status has been reported as beneficial for the cytological interpretation by some authors,[29] whereas others have reported a potential issue with that.[30]

In the present study, the interobserver and intraobserver variability in low-grade squamous lesions was within fair agreement (Fleiss´ kappas 0.29–0.31 and Kendall's coefficient of concordance 0.52–0.58). These findings are in agreement with previously published results.[31] Several studies with quality control and assurance scope showed ranges for kappa value of 0.54–0.69[32] and even 0.66–0.77 with improvement to 0.86–0.93.[33]

Nevertheless, the ASC-US category failed in the majority of studies obtaining a kappa value of 0.40[32] or even 0.16.[31] When the ASC-US category was pooled into ASC-US favor reactive versus ASC-US favor SIL, the value was 0.23 and 0.36, respectively, and despite subcategorization of the values, the agreement remained poor to fair.[31] Pioneering BIRST-1 and BIRST-2 studies showed timeline improvement even in borderline categories including the ASC-US category with an increase from

39.9% to 61.7% agreement in the BIRST-2 study;[34,35] however, direct comparison to our results is not possible as different values of concordance have been used. The use of various concordance values makes comparison of various studies challenging. In our series, Fleiss´ kappa values ranged from 0.19 to 0.22 and Kendall's coefficient ranged from 0.42 to 0.52 in the ASC-US category. In contrast, Stoler and Schiffman showed a kappa of 0.64 in the ASC-US category, superior to 0.51 in the LSIL and 0.56 in the NILM categories.[10] In the same study, colposcopic biopsy and loop electrosurgical excision procedure (LEEP) revealed comparable kappa values to liquid-based cytology.[10] Variability is a problem in all diagnostic aspects of cervical cancer, involving cytology, histology, and colposcopy. All three methods are subjected to interpretative variability.[10–13] Practically, the management of a particular TBS category may vary in different countries, so there may be a tendency to categorize according to local clinical management and guidelines.

Bigras et al. analyzed ASC-US–related cytomorphological features in detail and nuclear features had significantly lower reproducibility (kappa 0.46) with the lowest agreement for chromatin texture in comparison to other cytological findings.[36] Interestingly, fellows´ conditional kappa value were 0.70 for ASC-US in comparison to 0.60 for cytopathologists in another study showing the role of active education.[37] From that perspective, it will be valuable in the near feature to analyze our samples for particular features related to the highest and lowest agreement. Endocervical atypical lesions with less agreement are often impaired by degenerative changes and lack of nuclear enlargement.[38]

When the two rounds were compared, expert cytopathologists were not biased by HR HPV data generally. Nevertheless, personal scores showed better agreement in the second round with known HR HPV status with statistical significance ($p = .004753$ for Fleiss´ kappa and $p = .0005778$ for Kendall's coefficient). HR HPV known status influenced assessment in the ASC-US category in the present study: the agreement increased from 0.19 to 0.22 by Fleiss´ kappa and grew even higher by Kendall's coefficient from 0.42 to 0.52.

Several studies have reported HR HPV–positive cases to be highly likely reported as abnormal, so there is a bias of HR HPV knowledge in the interpretation of Pap smears or liquid-based preparations.[30,39–41] The phenomenon was noticed both in cytopathologist and cytotechnologist studies.[30,39–41] The Tahoe study showed a significant difference in the ASC-US, LSIL, and HSIL categories,[41] whereas we identified a p value significance only in the NILM category and personal scores. Generally, if the observer knows the HR HPV status, he or she tends to interpret minor cellular changes as abnormal findings upgrading the TBS category. Nevertheless, also downgrading of HR HPV–negative ASC-US has also been observed.[39] The HR HPV/Pap triage algorithm differs from a Pap smear–alone algorithm, so the diagnostic accuracy needs to be reevaluated in the HR HPV/Pap smear/liquid-based preparation triage and Pap smear/liquid-based preparation alone.[30] In the analysis of the College of American Pathologists PAP Education program LSIL was misclassified in 2.4% of cases and ThinPrep LSIL slides were more likely to be misclassified than SurePath LSIL slides.[42]

HR HPV primary screening programs have shown an increase in colposcopy referral rates[6,7] and the role of cytology triage is pivotal.[6] Importantly, prior knowledge of HR HPV status improved CIN2+ detection[43] in agreement with higher detection rates in HR HPV primary screening programs.[6]

Although cytology still has a main role in cervical screening, different challenges need to be faced by the cytological community in the future. In the spite of efforts to promote the training of cytotechnicians,[44] experienced cytologists are getting older and it will be increasingly difficult to replace them, particularly considering the suboptimal attention devoted to cytopathology education in most academic settings. Thus, the development of new digital systems based on whole-slide technology seems very promising.

Digital platforms have been gradually used in research, quality assurance programs[14–18] and routine diagnostics despite technical challenges in focusing on three-dimensional cell groups, large file size, and increased acquisition times for whole-slide images in cytology.[22–25] In our study, technical problems with image sharpness/focusing and scan quality were less common than with stain quality, suggesting a bright future for digital cytopathology. Notably, dedicated medical-grade computer screens were used only on a few occasions, and their use would probably increase both ease of use and diagnoses. High staining quality and uniformity in routine stains represents the basis of digitalization. Indeed, pathologists are good in getting used to new systems and tools, and developing

confidence, and 47% of participants would recommend digital cytology in their department based on their experience with the present study. Further standardization of digital cytopathology will reduce the problems and pitfalls.

Skills with digital pathology have been gained mainly from histopathology and research in the present study accompanying survey. The time needed to examine virtual gynecological cytology slides was higher than conventional microscopy slides in previous study,[21] our participants also subjectively scaled time demands at 2.8 in the 1 to 5 range, with 1 being best in mild preference of conventional microscopy. Ergonomic issues such as eye and back discomfort were graded 3 and 3.6, respectively. Pathology routine reporting is repetitive and continuous, so the risk of musculoskeletal injuries namely back and carpal tunnel syndrome is significant. In addition to visual fatigue, work environment, noise, and temperature should also be considered.[26]

In fact, whole-slide digital cytology will profit enormously from artificial intelligence–based diagnostic algorithms. In the past, several attempts were made with automated screening systems,[45–47] which were based on glass slides. Applying artificial intelligence to virtual slides may open new perspectives, reducing the strain of the screening activity and helping in the interpretation of slides in the context of large population-based screening programs. Artificial intelligence–driven digital pathology has been raising several ethical challenges, namely privacy of data and form of consent when sharing of data for research, commercial benefit of data and related public trust, and biases of data series. Involvement of nonmedical experts in diagnostic workup may also awaken ethical and responsibility issues.[48,49]

In conclusion, the interobserver and intraobserver variability in low grade squamous lesions was within fair agreement values in the present study in agreement with previous studies. The comparison of two rounds showed that expert cytopathologists are not generally biased by HR HPV data, but knowledge of HR HPV status leads to a better agreement. Stain quality and back discomfort were the main complaints in using digital cytopathology in our compliance survey. Following a good compliance with digital cytopathology, 47% of participants would recommend digital cytopathology in their department based on their experience in the present study.

AUTHOR CONTRIBUTIONS

Ivana Kholová: Conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, visualization, supervision, validation, writing-original draft, and writing-editing. **Giovanni Negri:** Conceptualization, data curation, resources, validation, and writing-editing. **Maria Nasioutziki:** Conceptualization, data curation, resources, validation, and writing-editing. **Laura Ventura:** Formal analysis, methodology, validation, writing-original draft, and writing-editing. **Arrigo Capitanio:** Methodology, project administration, resources, and writing-editing. **Massimo Bongiovanni:** Conceptualization, data curation, validation, and writing-editing. **Paul A. Cross:** Conceptualization, data curation, writing-original draft, and writing-editing. **Claire Bourgain:** Data curation and writing-editing. **Henrik Edvardsson:** Data curation and writing-editing. **Rosario Granados:** Data curation and writing-editing. **Artur Lipiński:** Data curation and writing-editing. **Ellen Christina Obermann:** Data curation and writing-editing. **Maurizio Pinamonti:** Data curation and writing-editing. **Henrieta Sidlova:** Data curation and writing-editing. **Margareta Strojan Fležar:** Data curation and writing-editing. **Folkert J. van Kemenade:** Data curation and writing-editing. **Danijela Vrdoljak-Mozetic:** Data curation and writing-editing. **Ambrogio Fassina:** Conceptualization, methodology, project administration, resources, writing-original draft, and writing-editing. **Beatrix Cochand-Priollet:** Conceptualization, data curation, resources, validation, supervision, and writing-editing.

ACKNOWLEDGMENTS

The study was initiated, organized, financed (sample shipping), and endorsed by European Federation of Cytology Societies.

FUNDING INFORMATION

The study was initiated, organized, financed (sample shipping), and endorsed by European Federation of Cytology Societies.

CONFLICTS OF INTEREST

The authors made no disclosures.

REFERENCES

1. Papanicolaou GN, Traut HF. The diagnostic value of vaginal smears in carcinoma of the uterus. *Am J Obst Gynecol.* 1941;42:193–206.
2. Bergeron C, von Knebel Doeberitz M. The role of cytology in the 21st century: the integration of cells and molecules. *Acta Cytol.* 2016;60(6):540–542. doi:https://doi.org/10.1159/000449402
3. Maver PJ, Poljak M. Primary HPV-based cervical cancer screening in Europe: implementation status, challenges, and future plans. *Clin Microbiol Infect.* 2020;26(5):579–583. doi:https://doi.org/10.1016/j.cmi.2019.09.006
4. Ibáñez R, Mareque M, Granados R, et al. Comparative cost analysis of cervical cancer screening programme based on molecular detection of HPV in Spain. *BMC Womens Health.* 2021;21(1):178. doi:https://doi.org/10.1186/s12905-021-01310-8
5. Aitken CA, van Agt HME, Siebers AG, et al. Introduction of primary screening using high-risk HPV DNA detection in the Dutch cervical cancer screening programme: a population-based cohort study. *BMC Med.* 2019;17(1):228. doi:10.1186/s12916-019-1460-0
6. Veijalainen O, Kares S, Kotaniemi-Talonen L, et al. Primary HPV screening for cervical cancer: results after two screening rounds in a regional screening program in Finland. *Acta Obstet Gynecol Scand.* 2021;100(3):403–409. doi:https://doi.org/10.1111/aogs.14021

7. Kaljouw S, Jansen EEL, Aitken CA, Harrijvan LM, Naber SK, de Kok IMCM. Reducing unnecessary referrals for colposcopy in hr HPV-positive women within the Dutch cervical cancer screening programme: a modelling study. *Gynecol Oncol.* 2021;160(3):713–720. doi:https://doi.org/10.1016/j.ygyno.2020.12.038
8. Kares S, Veijalainen O, Kholová I, et al. HIGH-RISK HPV testing as the primary screening method in an organized regional screening program for cervical cancer: the value of HPV16 and HPV18 genotyping? *APMIS.* 2019;127(11):710–716. doi:https://doi.org/10.1111/apm.12990
9. Abdul-Karim FW, Powers CN, Bererk JS, et al. Atypical aquamous cells. In: Nayar R, Wilbur DC, editors. Bethesda System for Reporting Cervical Cytology, Third Edition. *Springer*, 2015:103–134.
10. Stoler MH, Schiffman M; Atypical Squamous Cells of Undetermined Significance-Low-grade Squamous Intraepithelial Lesion Triage Study (ALTS) Group. Interobserver reproducibility of cervical cytologic and histologic interpretations: realistic estimates from the ASCUS-LSIL Triage Study. *JAMA.* 2001;285(11):1500–1505. doi:https://doi.org/10.1001/jama.285.11.1500
11. Grenko RT, Abendroth CS, Frauenhoffer EE, Ruggiero FM, Zaino RJ. Variance in the interpretation of cervical biopsy specimens obtained for atypical squamous cells of undetermined significance. *Am J Clin Pathol.* 2000;114(5):735–740. doi:https://doi.org/10.1309/K7C9-X5P0-001B-2HK5
12. Sellors JW, Nieminen P, Vesterinen E, Paavonen J. Observer variability in the scoring of colpophotographs. *Obstet Gynecol.* 1990;76(6):1006–1008.
13. Vallikad E, Siddartha PT, Kulkarni KA, et al. Intra and inter-observer variability of transformation zone assessment in colposcopy: a qualitative and quantitative study. *J Clin Diagn Res.* 2017;11(1):XC04–XC06. doi:https://doi.org/10.7860/JCDR/2017/21943.9168
14. Ross J, Greaves J, Earls P, Shulruf B, Van Es SL. Digital vs traditional: are diagnostic accuracy rates similar for glass slides vs whole slide images in a non-gynaecological external quality assurance setting? *Cytopathology.* 2018;29(4):326–334. doi:https://doi.org/10.1111/cyt.12552
15. Cross PA, Hodgson C, Crossley J, Crossley B. Development of a technical external quality assurance scheme in non-gynaecological cytology in UK. *Cytopathology.* 2015;26(2):71–74. doi:https://doi.org/10.1111/cyt.12245
16. UK NEQAS International Quality Expertise. Accessed February 3, 2022. https://ukneqas.org.uk/
17. Labquality External Quality Assessment. Accessed February 3, 2022. https://www.labquality.fi/en/external-quality-assessment/
18. College of American Pathologists. Proficiency Testing (PT)/External Quality Assessment (EQA) Programs Process. Accessed February 3, 2022. https://www.cap.org/laboratory-improvement/laboratories-outside-the-usa/proficiency-testing-pt-external-quality-assessment-eqa-programs/proficiency-testing-pt-external-quality-assessment-eqa-programs-process
19. Stewart J 3rd, Bevans-Wilkins K, Bhattacharya A, Ye C, Miyazaki K, Kurtycz DF. Virtual microscopy: an educator's tool for the enhancement of cytotechnology students' locator skills. *Diagn Cytopathol.* 2008;36(6):363–368. doi:https://doi.org/10.1002/dc.20821
20. Van Es SL, Kumar RK, Pryor WM, Salisbury EL, Velan GM. Cytopathology whole slide images and adaptive tutorials for postgraduate pathology trainees: a randomized crossover trial. *Hum Pathol.* 2015;46(9):1297–1305. doi:https://doi.org/10.1016/j.humpath.2015.05.009
21. Dee FR, Donnelly A, Radio S, Leaven T, Zaleski MS, Kreiter C. Utility of 2-D and 3-D virtual microscopy in cervical cytology education and testing. *Acta Cytol.* 2007;51(4):523–529. doi:https://doi.org/10.1159/000325788
22. McAlpine ED, Pantanowitz L, Michelow PM. Challenges developing deep learning algorithms in cytology. *Acta Cytol.* 2021;65(4):301–309. doi:https://doi.org/10.1159/000510991
23. Pantanowitz L. Digital cytology: look how much has been achieved. *Cytopathology.* 2020;31(5):370–371. doi:https://doi.org/10.1111/cyt.12866
24. Chantziantoniou N, Mukherjee M, Donnelly AD, Pantanowitz L, Austin RM. Digital applications in cytopathology: problems, rationalizations, and alternative approaches. *Acta Cytol.* 2018;62(1):68–76. doi:https://doi.org/10.1159/000484434
25. Pantanowitz L, Parwani AV, Khalbuss WE. Digital imaging for cytopathology: are we there yet? *Cytopathology.* 2011;22(2):73–74. doi:https://doi.org/10.1111/j.1365-2303.2011.00852.x
26. Krupinski EA. Virtual slide telepathology workstation of the future: lessons learned from teleradiology. *Hum Pathol.* 2009;40(8):1100–1111. doi:https://doi.org/10.1016/j.humpath.2009.04.011
27. Feldt LS, Woodruff DJ, Salih FA. Statistical inference for coefficient alpha. *Applied Psychol Measure.* 1987;11:93–103.
28. Ronco G, Dillner J, Elfström KM, et al; International HPV screening working group. Efficacy of HPV-based screening for prevention of invasive cervical cancer: follow-up of four European randomised controlled trials. *Lancet.* 2014;383(9916):524–532. doi:https://doi.org/10.1016/S0140-6736(13)62218-7
29. Bergeron C, Giorgi-Rossi P, Cas F, et al. Informed cytology for triaging HPV-positive women: substudy nested in the NTCC randomized controlled trial. *J Natl Cancer Inst.* 2015;107(2):dju423. doi:https://doi.org/10.1093/jnci/dju423
30. Richardson LA, El-Zein M, Ramanakumar AV, et al; PEACHS (Pap Efficacy After Cervical HPV Status) Study Consortium. HPV DNA testing with cytology triage in cervical cancer screening: influence of revealing HPV infection status. *Cancer Cytopathol.* 2015;123(12):745–754. doi:https://doi.org/10.1002/cncy.21596
31. Crum CP, Genest DR, Krane JF, et al. Subclassifying atypical squamous cells in Thin-Prep cervical cytology correlates with detection of high-risk human papillomavirus DNA. *Am J Clin Pathol.* 1999;112(3):384–390. doi:https://doi.org/10.1093/ajcp/112.3.384
32. Confortini M, Di Stefano C, Biggeri A, et al. Daily peer review of abnormal cervical smears in the assessment of individual practice as an additional method of internal quality control. *Cytopathology.* 2016;27(1):35–42. doi:https://doi.org/10.1111/cyt.12195
33. de Morais LSF, Magalhães JC, Braga IDS, Marega LA, Tavares SBDN, Amaral RG. Performance of laboratories after 10 years of participating in external quality monitoring in cervical cytology. *Acta Cytol.* 2020;64(3):224–231. doi:https://doi.org/10.1159/000502433
34. Sherman ME, Dasgupta A, Schiffman M, Nayar R, Solomon D. The Bethesda Interobserver Reproducibility Study (BIRST): a web-based assessment of the Bethesda 2001 System for classifying cervical cytology. *Cancer.* 2007;111(1):15–25. doi:https://doi.org/10.1002/cncr.22423
35. Kurtycz DFI, Staats PN, Chute DJ, et al. Bethesda Interobserver Reproducibility Study-2 (BIRST-2): Bethesda System 2014. *J Am Soc Cytopathol.* 2017;6(4):131–144. doi:https://doi.org/10.1016/j.jasc.2017.03.003
36. Bigras G, Wilson J, Russell L, Johnson G, Morel D, Saddik M. Interobserver concordance in the assessment of features used for the diagnosis of cervical atypical squamous cells and squamous intraepithelial lesions (ASC-US, ASC-H, LSIL and HSIL). *Cytopathology.* 2013;24(1):44–51. doi:https://doi.org/10.1111/j.1365-2303.2011.00930.x
37. Chebib I, Rao RA, Wilbur DC, Tambouret RH. Using the ASC:SIL ratio, human papillomavirus, and interobserver variability to assess and monitor cytopathology fellow training performance. *Cancer Cytopathol.* 2013;121(11):638–643. doi:https://doi.org/10.1002/cncy.21328
38. Pulkkinen J, Huhtala H, Krogerus L, Hollmén S, Laurila M, Kholová I. Endocervical cytology: inter- and intra-observer variability in conventional Pap smears. *Acta Cytol.* 2022;66(3):206–215. doi:https://doi.org/10.1159/000522212
39. Aitken CA, Holtzer-Goor KM, Uyterlinde A, et al. The impact of knowledge of HPV positivity on cytology triage in primary high-risk HPV screening. *J Med Screen.* 2019;26(4):221–224. doi:https://doi.org/10.1177/0969141319864991

40. Doxtader EE, Brainard JA, Underwood D, Chute DJ. Knowledge of the HPV status biases cytotechnologists' interpretation of Pap tests originally diagnosed as negative for intraepithelial lesion or malignancy. *Cancer Cytopathol.* 2017;125(1):60–69. doi:https://doi.org/10.1002/cncy.21783
41. Moriarty AT, Nayar R, Arnold T, et al. The Tahoe Study: bias in the interpretation of Papanicolaou test results when human papillomavirus status is known. *Arch Pathol Lab Med.* 2014;138(9):1182–1185. doi:https://doi.org/10.5858/arpa.2012-0115-CP
42. Crothers BA, Ghofrani M, Zhao C, et al. Low-grade squamous intraepithelial lesion or high-grade squamous intraepithelial lesion? Concordance between the interpretation of low-grade squamous intraepithelial lesion and high-grade squamous intraepithelial lesion in Papanicolaou tests: results from the College of American Pathologists PAP Education Program. *Arch Pathol Lab Med.* 2019;143(1):81–85. doi:https://doi.org/10.5858/arpa.2018-0003-CP
43. Benoy IH, Vanden Broeck D, Ruymbeke MJ, et al. Prior knowledge of HPV status improves detection of CIN2+ by cytology screening. *Am J Obstet Gynecol.* 2011;205(6):569.e1–e7. doi:https://doi.org/10.1016/j.ajog.2011.06.101
44. Anic V, Eide ML, Cochand-Priollet B, Vrdoljak Mozetic D, Negri G, Vielh P. Recommendations of the European Advisory Committee of Cytotechnology and European Federation of Cytology Societies for Training and Education of Cytotechnologists in Europe. *Acta Cytol.* 2021;65(3):199–204. doi:https://doi.org/10.1159/000513899
45. O'Leary TJ, Tellado M, Buckner SB, Ali IS, Stevens A, Ollayos CW. PAPNET-assisted rescreening of cervical smears: cost and accuracy compared with a 100% manual rescreening strategy. *JAMA.* 1998;279(3):235–237. doi:https://doi.org/10.1001/jama.279.3.235
46. Wilbur DC, Prey MU, Miller WM, Pawlick GF, Colgan TJ. The AutoPap system for primary screening in cervical cytology. Comparing the results of a prospective, intended-use study with routine manual practice. *Acta Cytol.* 1998;42(1):214–220. doi:https://doi.org/10.1159/000331549
47. Boon ME, Ouwerkerk-Noordam E, Meijer-Marres EM, Bontekoe TR. Switching from neural networks (PAPNET) to the Imager (Hologic) for computer-assisted screening. *Acta Cytol.* 2011;55(2):163–166. doi:https://doi.org/10.1159/000323310
48. McKay F, Williams BJ, Prestwich G, Bansal D, Hallowell N, Treanor D. The ethical challenges of artificial intelligence-driven digital pathology. *J Pathol Clin Res.* 2022;8(3):209–216. doi:https://doi.org/10.1002/cjp2.263
49. Cui M, Zhang DY. Artificial intelligence and computational pathology. *Lab Invest.* 2021;101(4):412–422. doi:https://doi.org/10.1038/s41374-020-00514-0