Published in final edited form as:

J Low Genit Tract Dis. 2016 April ; 20(2): 154–161. doi:10.1097/LGT.0000000000000194.

Comparison of colposcopic impression based on live colposcopy and evaluation of static digital images

Angela H. Liu, B.S.¹, Michael A. Gold, M.D.², Mark Schiffman, M.D.¹, Katie Smith, M.D.³, Rosemary Zuna, M.D.³, S. Terence Dunn, Ph.D.³, Julia C. Gage, Ph.D.¹, Joan Walker, M.D.³, and Nicolas Wentzensen, M.D.¹

1

2

3

Abstract

Objective—The aim of the study was to evaluate the agreement and compare diagnostic accuracy of colposcopic impressions from live colposcopy versus evaluation of static digital images.

Methods—Live impressions and corresponding static images obtained during colposcopy of 690 women were independently compared. Diagnostic accuracy was calculated for colposcopic impressions from both methods, varying hypothetical thresholds for colposcopically-directed cervical biopsies (acetowhitening or worse, low-grade or worse, high-grade or worse). Stratified analyses investigated the impact of referral cytology, HPV16 infection, and age on colposcopic impression.

Results—Overall agreement between live and static colposcopic visualization was 43.0% ($\kappa = 0.20$; 95%CI: 0.14-0.26) over normal, acetowhitening, low-grade, and high-grade impressions. Classification of acetowhitening or worse impressions showed the highest agreement (92.2%; $\kappa = 0.39$, 95%CI: 0.21-0.57); both methods achieved >95% sensitivity for CIN2⁺. Agreement between live and static colposcopic visualization was 69.3% for rating low-grade or worse impressions ($\kappa = 0.23$, 95%CI: 0.14-0.33) and 71% when rating high-grade impressions ($\kappa = 0.33$, 95%CI: 0.24-0.42). Live colposcopic impressions were more likely to be rated low-grade or worse (p-value < 0.01, OR = 3.5, 95%CI: 2.4-5.0), yielding higher sensitivity for CIN2⁺ at this threshold than static image assessment (95.4% vs. 79.8%, p-value < 0.01). Overall, colposcopic impressions were more likely rated high-grade on live assessment among women referred with high-grade cytology (OR = 3.3, 95% CI = 1.8-6.4), significantly improving the sensitivity for CIN2⁺ (66.3% vs. 48.5%, p-value < 0.01).

Conclusion—Colposcopic impressions of acetowhitening or worse are highly sensitive for identifying cervical precancers and reproducible on static image-based pattern recognition.

To whom correspondence should be addressed: Angela Hui-Chia Liu B.S., Division of Cancer Epidemiology and Genetics, National Cancer Institute, 9609 Medical Center Drive, Room 7E-216, Telephone #: (240) 276-6970, angela.liu@nih.gov.

Conflict of interest statement: The authors declare no conflict of interest.

IRB status: Institutional Review Boards at both the OUHSC and the National Cancer Institute approved the study.

Keywords

Colposcopy; static digital image of the cervix; diagnostic accuracies; cervical intra-epithelial neoplasias; colposcopic impressions

Introduction

High-resolution static digital images of the cervix are being widely used for training of colposcopy, both in developed countries with established cervical cancer screening programs, as well as for web-based, distance-learning to increase expertise in low-resource areas with the heaviest cervical disease burden.^{1,2,9,20}

However, the use of static digital images of the cervix for colposcopic assessment remains controversial. It has been argued that static digital images do not capture the dynamic acetowhitening changes in the cervical epithelium.⁴ The diagnostic reliability and accuracy of static images in teaching colposcopy is an important topic. Colposcopy expertise remains a critical part of most fully developed cervical cancer screening programs. Accurate visual identification of colposcopic abnormalities is crucial for directing placement of diagnostic biopsies. Previous research on colposcopic impression showed 93% of women with high-grade cervical intraepithelial neoplasia histology had at least one acetowhite lesion on colposcopy. While acetowhite lesions are sensitive for detection of high-grade cervical intra-epithelial neoplasia, acetowhitening alone is a common and unspecific finding.¹⁵

We performed ancillary analyses in the Biopsy Study²⁵, a study with a highly-sensitive colposcopy-biopsy protocol for detecting high-grade cervical precancers, to evaluate colposcopic impression based on static images. We compared static images to live colposcopic impression and histologic outcomes to define the utility of static image of the cervix for training, education, case review, and the development of more accurate image-based diagnostic colposcopy technologies.

Methods

Study design and population

Women 18 years and older referred for colposcopy after abnormal cervical screening results at the University of Oklahoma Health Sciences Center (OUHSC) were enrolled between February 2009 and September 2012 as previously described²⁴. Institutional Review Boards at both the OUHSC and the National Cancer Institute approved the study.

Colposcopy and biopsy protocol

Referral for colposcopy was community-based, using the 2007 American Society for Colposcopy and Cervical Pathology (ASCCP) guidelines²⁶. Six colposcopists each performed between 60-179 colposcopic exams. Using a Wallach broom, cervical samples were collected in PreservCyt solution (Hologic, Marlborough, MA, USA) for cytology and HPV testing. Using acetowhitening changes as biopsy threshold, a lesion-directed biopsy sampled each distinct area of aceto-white cervical epithelium in the transformation zone

after application of 5% acetic acid, regardless of the lesion location. If lesion-directed biopsies were fewer than four, a biopsy of the normal-appearing cervical transformation zone epithelium in any quadrant was added. Hematoxylin and eosin-stained tissue sections were obtained for each biopsy and evaluated by a study pathologist for clinical management. In addition, p16-staining was performed to establish biomarker-adjudicated cervical precancer endpoints³ (CINtec, Roche mtm Laboratories, Mannheim, Germany)

Cytology and HPV testing

Community-based referral cytology was reported using the Bethesda nomenclature, including the categories “Normal for intraepithelial lesion or malignancy” (NILM), “atypical squamous cells of undermined significance” (ASC-US), “low-grade squamous intraepithelial lesion” (LSIL), “atypical squamous cells, favor high-grade” (ASC-H), and “high-grade squamous intraepithelial lesion” (HSIL)¹⁹. HPV genotyping was performed using the Linear Array (LA) HPV Genotyping Test (Roche Molecular Diagnostics, Branchburg, NJ) according to the manufacturer’s instructions, with slight modifications^{23,24}.

Image documentation and evaluation

The biopsy protocol also involved taking digital photographs to document colposcopic impressions and biopsy sites. A Nikon D700 SLR digital camera with resolution of 12.1 megapixels (Nikon D700. Nikon Corporation. 2008. Tokyo, Japan) was attached to a Zeiss 150FC colposcope (300mm focal length, 5 standard magnifications: 0.4, 0.6, 1.0, 1.6, 2.4) via a Zeiss Beam splitter 50 and T2 Camera Adaptor F=340 (Kolposkop 150FC from Zeiss. Zeiss. Carl Zeiss Meditec, Inc. Oberkochen, Germany). Initial visual examination was performed after one minute of acetic acid application. To image a representative colposcopic view, colposcopists could re-apply the acetic acid every 3-5 minutes, allowing the color of columnar epithelium to return from blanch-white to red between each application. Using the Boundary Marking Tool 2 (BMT2) software developed by the National Library of Medicine (Boundary Marking Tool 2. National Library of Medicine, Bethesda, MD, U.S.A), colposcopists marked on the digital static images boundaries of distinct aceto-white lesions, biopsy sites, and ranked each lesion/biopsy by order of severity at the time of colposcopy. An overall colposcopic impression was made for each procedure based on evaluation of lesion punctuation, mosaicism, borders, and colors of the whole cervix using parameters similar to the Reid index¹⁷. Colposcopic impressions were rated normal, acetowhitening, low-grade, or high-grade (included invasive cancer). Patients’ referral cytology and age were known at the time of live colposcopic evaluation. The digital images of the cervixes were stored and accessible through a web-based database. Using the same BMT software and criteria for rating colposcopic impression, an expert colposcopist independently evaluated all static images from study participants. The expert colposcopist was not aware of any clinical data of the patients, or the live colposcopic impressions. A separate evaluator determined whether the image quality was satisfactory for diagnosis or unsatisfactory due to blood obscuring the cervical os, poor focus, vaginal wall prolapse, or other factors. Both colposcopists who performed the procedure and the expert colposcopist who reviewed static images were not aware of the results of HPV testing and cytology samples collected at the colposcopy visit.

Statistical Analysis

Agreement between live and static colposcopic visualization was calculated using percent agreement, Kappa statistics, and McNemar test of marginal proportions. Colposcopic impressions and ratings were dichotomized at three levels – acetowhitening or worse impressions versus normal cervix, low-grade lesions or worse versus less severe impressions, and high-grade lesions versus less severe impressions. To assess whether cervical cancer risk indicators affected the rating of colposcopic impressions, the population was stratified by referral cytology results (ASC-US/LSIL, ASC-H/HSIL), HPV16 infection status (HPV-positive but no HPV16, HPV16-positive), and age (21-24, 25-29, 30 years and older).

To evaluate the accuracy of colposcopy for clinically relevant cervical cancer treatment endpoint (CIN2⁺), sensitivities and specificities (including 95% confidence intervals) were calculated using the following – acetowhitening, low-grade impression, and high-grade impression – as hypothetical thresholds for colposcopically-directed cervical biopsies (with the assumption of very high sensitivity of histology given the multiple-biopsy strategy). Analyses were stratified by referral cytology, HPV16 status, and age, to identify cervical cancer risk indicators that might improve the diagnostic accuracy of colposcopy.

The overall analyses included images of high and satisfactory quality. To delineate the effect of image quality, all analyses were repeated for high quality images. All analyses and graphics were performed with the statistical programming language R-3.1.2., in the open-source integrated development environment of RStudio (RStudio: Integrated Development for R. RStudio Team. RStudio, Inc. 2015. Boston, MA, USA).

Results

Study population, colposcopy-biopsy procedures, and image characteristics

Women enrolled in the study had a median age of 26 years old (age range: 18-76 years). As previously described, the number of cervical biopsies performed increased with increasing severity of colposcopic impression.²⁵ The number of lesion-directed biopsies did not differ significantly between those referred with low-grade versus high-grade cytology ($\chi^2=5.42$, p-value = 0.14; Supplemental Table 1). The distribution of referral cytology showed 150 women (21.7%) with ASC-US, 280 (40.6%) with LSIL, and 292 (42.3%) with HSIL or ASC-H. HPV testing showed 58 (8.4%) HPV-negative women, 436 (63.2%) were positive for non-HPV16 types, and 192 (27.8%) tested HPV16 positive. The worst histology diagnosis from up to four lesion-directed colposcopy biopsies showed 409 (59.3%) women with CIN1, 198 (28.7%) with CIN2, and 75 with CIN3⁺ (10.9%; include 4 invasive cancers).

Six hundred and ninety static digital images of the cervix were used in the study, one for each participant. Three hundred and seventeen images were rated high-quality, 241 rated satisfactory, 33 rated not usable, and 99 images were missing ratings of image quality. No significant difference was noted for the distributions of patient characteristics and histology findings between groups of different quality images. Further analyses were restricted to images of high or satisfactory qualities (table 1).

Analyses of agreement between ratings of colposcopic impression by live and static visualizations

The exact overall agreement between live and static colposcopic visualization was 43.0% ($\kappa = 0.20$, 95% CI: 0.14-0.26) across all four categories of colposcopic impressions (table 2). Live colposcopy and static image assessment had similar proportions of normal and high-grade impressions (Normal impression: 7.2% vs. 6.7%; High-grade impressions: 32.4% vs. 28.9%, respectively). In contrast, we observed a pronounced difference between live and static colposcopic visualization in the frequencies of rating acetowhitening and low-grade lesions. A low-grade impression, rather than acetowhitening, was more likely to be reported on live colposcopy compared to static image assessment (acetowhitening impression: 9.9% vs. 27.6%; low-grade impression: 47.1% vs. 36.4%, respectively).

Agreement between live and static colposcopic ratings of acetowhitening or worse impressions versus normal cervix was 92.2% ($\kappa = 0.39$, 95% CI: 0.21-0.57, table 3). Proportions of discordant ratings did not differ significantly between live and static colposcopic distinction of acetowhitening or worse abnormalities from normal cervix. (McNemar test of correlated proportions p-value 0.44), a consistent finding also noted in stratified analyses by referral cytology, HPV16 status, and age (table 4).

Dichotomizing colposcopic impression categories, as low-grade or worse versus less severe impressions, showed an overall agreement of 69.3% ($\kappa = 0.23$, 95% CI: 0.14-0.33, table 3) between live and static visualization. Analyses of discordant ratings showed impressions were more likely to be rated low-grade or worse on live than static colposcopic visualization (p-value < 0.01, OR = 3.5; 95% CI = 2.4-5.0), a trend noted across all stratifications by referral cytology, HPV16 status, and age (table 4).

Dichotomization of colposcopic rating categories, at the level of high-grade versus less severe impressions, showed an overall agreement of 71.0% ($\kappa = 0.33$, 95% CI: 0.24-0.42, table 3) between live and static modes of visualization, and difference in proportions of discordant rating reached borderline significance (p-value = 0.07, table 4). Colposcopic impressions were more likely to be rated high-grade on live than static mode of visualization among women referred with high-grade cytology (OR = 3.3, 95% CI = 1.8-6.4), but not in other stratified analyses (table 4). This finding suggested that the discordant rating of high-grade versus less severe impressions was influenced less by the mode of live versus static colposcopic visualization than by prior knowledge of patients' high-grade referral cytology.

Restricting the analyses of agreement between live colposcopy and static image assessment to high-quality images only showed results consistent with those discussed above.

Comparison of diagnostic accuracies for CIN2⁺ by live versus static colposcopic visualization

Based on the endpoints from the highly-sensitive multiple-biopsy colposcopy protocol, we further compared the accuracy of static image assessment versus live colposcopy for diagnosing CIN2⁺ at each threshold of colposcopic impression (table 5).

Comparison between live colposcopies and static images at the threshold of acetowhitening or worse showed comparable overall diagnostic accuracies for CIN2⁺ (sensitivity: 99.5% vs. 97.3%, p-value = 0.06; specificity: 12.1% vs. 9.0%, p-value = 0.10; table 5)

At the threshold of low-grade impression or worse, ratings on live colposcopy were significantly more sensitive, albeit less specific than those by static image assessment, for the overall detections of CIN2⁺ (sensitivity: 95.4% vs. 79.8%, p-value < 0.01; specificity: 26.7% vs. 43.8%, p-value < 0.01, table 5), a finding that persisted after stratification by referral cytology, HPV16 status, and age (table 6).

At the threshold of high-grade impression, live colposcopy yielded comparable overall CIN2⁺ diagnostic accuracies versus static colposcopic visualization (sensitivity: 52.3% vs. 44.4%, p-value = 0.08; specificity: 79.0% vs. 81.5%, p-value = 0.48, table 5). However, among women with high-grade referral cytology, the greater likelihood to rate the impression as high-grade on live than on static colposcopic visualization led to significantly higher detection sensitivity for CIN2⁺ (66.3% vs. 48.5%, p-value < 0.01), at the trade-off of significantly lower specificity (63.3% vs. 80.0%, p-value = 0.01) (table 6).

Discussion

Previous studies have demonstrated the subjectivity and limited reproducibility of colposcopic impressions; and the limited sensitivity of a single biopsy to find cervical precancers^{6,8,10-16,18,22}. We designed the Biopsy Study to evaluate current colposcopy-biopsy practice and the diagnostic accuracy of static digital imaging of the cervix. We have previously evaluated the accuracy of the multiple-biopsy colposcopy protocol used in the study.²⁵ Here we present findings regarding the clinical utility of static colposcopic images.

As part of the study, static images from all women were assessed by an expert colposcopist without knowledge of patients' referral cytology and age, which were known on live colposcopic evaluation. The exact overall agreement for colposcopic impressions on live versus static visualization was 43.0%, consistent with findings from prior studies¹⁸. Among women with high-grade cytology and HPV16-positive infections, agreement improved slightly to 48.8% and 46.2%, respectively.

Disagreement between live and static colposcopic visualization was significant for the evidently difficult distinction of acetowhitening and low-grade impressions, regardless of referral cytology, HPV16 infection status, or age. While the proportions of colposcopic impressions rated normal and high-grade were comparable, ratings were consistently shifted to low-grade impression from acetowhitening impression on live colposcopy. Higher likelihood to rate high-grade than less severe colposcopic impressions on live colposcopy was observed among women with high-grade referral cytology; while no significant difference in discordant ratings of high-grade versus less severe impressions was noted in other stratified analyses for the agreement between live and static colposcopic visualization. This finding implicates prior knowledge of patients' high-grade referral cytology as the main factor contributing to difference in discordant rating of high-grade versus less severe impressions on live colposcopy and static images. Together, these results suggest that

distinction of acetowhitening from low-grade impression can at times be challenging on static images, likely because dynamic changes in acetowhitening are better captured by live visualization. In contrast, high-grade changes appear to be clear enough for reliable recognition on static images.

Overall CIN2⁺ detection sensitivities increased with ratings of decreasing severity of colposcopic abnormality. Ratings of acetowhitening or worse impressions by live colposcopy and static image assessments both achieved comparable, >95% overall diagnostic sensitivities for CIN2⁺. Regardless of referral cytology or age, ratings of low-grade or worse impressions on live colposcopy were statistically more sensitive, albeit less specific than static images rated low-grade or worse impressions, for CIN2⁺ detection. Overall diagnostic accuracies for CIN2⁺ by ratings of high-grade impression were independent of the mode of live versus static colposcopic visualization. Higher likelihood to rate high-grade than less severe impressions, based on knowledge of patients' high-grade referral cytology, achieved higher sensitivity for CIN2⁺.

In summary, our study identified two factors contributing to differences in rating of live colposcopy and static images – (1) distinguishing acetowhitening and low-grade abnormalities may be more challenging in static images, likely because dynamic changes of acetowhitening are not captured by the latter method; and (2) prior knowledge of patients' high-grade referral cytology led to increases in rating colposcopic impression as high-grade. Both factors affected the theoretical accuracy of colposcopic impression at higher thresholds of abnormality. However, distinguishing minimally abnormal acetowhitening changes from normal cervix was reproducible on static images and achieved high diagnostic sensitivity for cervical precancer, by taking biopsies from the majority of women referred to colposcopy in this population.

Our study population was a typical cytology referral population for colposcopy in the U.S. The digital documentation process of colposcopic impression was well standardized, capturing high-resolution static digital images of the cervix. A single expert colposcopist reviewed all static images for comparison with live colposcopic impressions, minimizing the individual rater bias in evaluating static images. Future studies should extend the analyses to include more colposcopists evaluating static images, particularly to evaluate whether the difficulty in distinguishing acetowhitening from low-grade impression is rater-independent. Rating criteria for both live colposcopy and static image assessment were highly uniform. Multiple colposcopists contributed to live colposcopic impressions, making the results of live colposcopic rating more generalizable. While investigators in the ASCUS-LSIL Triage Study (ALTS) also noted the lack of meaningful difference between the diagnostic accuracies of live colposcopy and static cervigrams^{7,15,21}, our results extended beyond patients with low-grade referral cytology; and the comprehensive colposcopy-biopsy protocol used in this study minimized missed CIN2⁺ cases in our performance evaluation.

One limitation in the analyses arose from using overall colposcopic impressions as hypothetical thresholds for colposcopically-directed cervical biopsies; however, this is an inherent problem of image-based study. Digital image acquisition occurred immediately after acetic acid application per study protocol, with some discretion for the clinician to

capture a representative colposcopic view at the optimal timing relative to acetic acid application. The quality of static colposcopic images was related to operator's familiarity and experience with the image-capture software used in the study. Rating of image quality was influenced by the subjectivity of the independent reviewer, and can improve with objective metrics of image analyses. Masking patients' clinical history in static image evaluation also conflated the impact on colposcopic rating from prior knowledge of patients' referral cytology and age with that from the mode of dynamic versus static colposcopic visualization. Sample size was sufficient for analyses of overall agreement and diagnostic accuracies, but may have limited power to detect statistical significance in stratified analyses. Also, we cannot estimate the impact on colposcopy of the increasing trend of referring women due to persistent HPV infection even in the absence of abnormal cytology, which may make colposcopic diagnosis even more challenging.

Static digital imaging of the cervix is a valuable technology that can meet the demand of training and directly facilitating cervical precancer diagnosis by colposcopy, a procedure performed on 3 million women in the U.S. yearly and many more worldwide.^{5,27} By addressing issues in the current colposcopy-biopsy protocol, our findings validated and more clearly defined the utility of high-resolution, high-quality static digital images of the cervix for CIN2⁺ diagnosis. Our results support current efforts to disseminate, monitor, and maintain the quality of colposcopy expertise by teaching image-based pattern recognition of high-grade lesions, informed by patients' referral cytology whenever available. Using static digital images of high-grade cervical disease in the web-based colposcopy learning model is particularly impactful in low-resource, developing countries, where cervical cancer is disproportionately prevalent and referred women are more likely to present with high-grade disease. While high-quality static images could be pre-selected for training colposcopy, static images with limited quality may pose a challenge for the use of static images in direct, real-time colposcopic diagnosis facilitated by remote colposcopy expertise. Employing static colposcopic images for this purpose of telemedicine in low-resource setting should heed stringent quality control of image acquisition.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The study was supported by the Intramural Research Program of the National Cancer Institute. The funding source has no role in design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript.

References

1. [11 June, 2015] American Society for Colposcopy and Cervical Pathology (ASCCP) Online CME. <http://www.asccp.org/Education/Online-CME/Online-CME-Series>
2. British Society for Colposcopy and Cervical Pathology (BSCCP). [11 June, 2015] <https://www.bsccp.org.uk>
3. Bergeron C, Ordi J, Schmidt D, et al. Conjunctive p16INK4a testing significantly increases accuracy in diagnosing high-grade cervical intraepithelial neoplasia. *Am J Clin Pathol*. 2010; 133:395–406. [PubMed: 20154278]

4. Diggle, P.; Liang, K.; Zeger, SL. *Analyses of Longitudinal Data*. New York (NY): Oxford University Press; 1994. p. 253
5. Davey DD, Neal MH, Wilbur DC, et al. Bethesda 2001 implementation and reporting rates: 2003 practices of participants in the College of American Pathologists Interlaboratory Comparison Program in Cervicovaginal Cytology. *Arch Pathol Lab Med*. 2004; 128:1224–1229. [PubMed: 15504056]
6. Ferris DG, Litaker MS. Prediction of cervical histologic results using an abbreviated Reid colposcopic index during ALTS. *Am J Obstet Gynecol*. 2006; 194:704–10. [PubMed: 16522401]
7. Ferris DG, Schiffman Mark, Litaker Mark S. Cervicography for Triage of Women with Mildly Abnormal Cervical Cytology Results. *Am J Obstet Gynecol*. 2001; 185(4):939–43. [PubMed: 11641682]
8. Gage JC, Hanson VW, Abbey K, et al. Number of cervical biopsies and sensitivity of colposcopy. *Obstet Gynecol*. 2006; 108:264–272. [PubMed: 16880294]
9. International Federation for Cervical Pathology and Colposcopy (IFCPC). [11 June, 2015] <http://www.ifcpc.org/en/education>
10. Jeronimo J, Schiffman M. Colposcopy at a crossroads. *Am J Obstet Gynecol*. 2006; 195:349–53. [PubMed: 16677597]
11. Jeronimo J, Massad LS, Schiffman M. Visual appearance of the uterine cervix: correlation with human papillomavirus detection and type. *Am J Obstet Gynecol*. 2007; 197:47.e1–47.e8. [PubMed: 17618753]
12. Massad LS, Collins YC. Strength of correlation between colposcopic impression and biopsy histology. *Gynecol Oncol*. 2003; 89:424–8. [PubMed: 12798706]
13. Massad LS, Halperin CJ, Bitterman P. Correlation between colposcopically directed biopsy and cervical loop excision. *Gynecol Oncol*. 1996; 60:400–3. [PubMed: 8774646]
14. Massad LS, Jeronimo J, Schiffman M. Interobserver agreement in the assessment of components of colposcopic grading. *Obstet Gynecol*. 2008; 111:1279–1284. [PubMed: 18515509]
15. Massad LS, Jeronimo J, et al. The accuracy of colposcopic grading for detection of high grade cervical intraepithelial neoplasia. *J Low Genit Tract Dis*. 2009; 13:137–144. [PubMed: 19550210]
16. Milne DS, Wadehra V, Mennim D, Wagstaff TI. A prospective follow-up study of women with colposcopically unconfirmed positive cervical smears. *Br J Obstet Gynaecol*. 1999; 106:38–41. [PubMed: 10426257]
17. Saslow D, Solomon D, Lawson HW, et al. American Cancer Society, American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology screening guidelines for the prevention and early detection of cervical cancer. *CA Cancer J Clin*. 2012; 62:147–172. [PubMed: 22422631]
18. Stoler MH, Vichnin MD, Ferenczy A, et al. The accuracy of colposcopic biopsy: analyses from the placebo arm of the Gardasil clinical trials. *Int J Cancer*. 2011; 128:1354–1362. [PubMed: 20506504]
19. Solomon D, Davey D, Kurman R, et al. The 2001 Bethesda System: terminology for reporting results of cervical cytology. *JAMA*. 2002; 287:2114–2119. [PubMed: 11966386]
20. Tan, Jeffrey H; Wrede, C David H. *New Technologies and Advances in Colposcopic Assessment*. Best Practice & Research Clinical Obstetrics & Gynaecology. 2011; 25(5):667–77. [PubMed: 21664876]
21. The ASCUS-LSIL Triage Study (ALTS) Group. Results of a randomized trial on the management of cytology interpretations of atypical squamous cells of unknown significance. *Am J Obstet Gynecol*. 2003; 188:1383–92. [PubMed: 12824967]
22. Wang SS, Zuna RE, et al. Human papillomavirus (HPV) cofactors by disease progression and HPV types in the Study to Understand Cervical Cancer Early Endpoints and Determinants (SUCCEED). *Cancer Epidemiol Biomarkers Prev*. 2009; 18:113–20. [PubMed: 19124488]
23. Wentzensen N, Schiffman M, Dunn ST, et al. Grading the severity of cervical neoplasia based on combined histopathology, cytopathology, and HPV genotype distribution among 1,700 women referred to colposcopy in Oklahoma. *Int J Cancer*. 2009; 124:964–969. [PubMed: 19030188]

24. Wentzensen N, Schiffman M, Dunn T, et al. Multiple human papillomavirus genotype infections in cervical cancer progression in the study to understand cervical cancer early endpoints and determinants. *Int J Cancer*. 2009; 125:2151–2158. [PubMed: 19585494]
25. Wentzensen N, Walker J, et al. Multiple biopsies and detection of cervical cancer precursors at colposcopy. *J Clin Oncol*. 2015; 33:83–9. [PubMed: 25422481]
26. Wright TC Jr, Cox JT, Massad LS, et al. 2001 Consensus Guidelines for the management of women with cervical cytological abnormalities. *JAMA*. 2002; 287:2120–2129. [PubMed: 11966387]
27. Wright TC Jr, Massad LS, Dunton CJ, et al. 2006 consensus guidelines for the management of women with abnormal cervical cancer screening tests. *Am J Obstet Gynecol*. 2007; 197:346–355. [PubMed: 17904957]

List of Abbreviation

ASC-US	Atypical squamous cell of unknown significance
BMT	Boundary Marking Tool
ASC-H	Atypical squamous cells favor high-grade
CIN	Cervical intra-epithelial neoplasia
HPV	Human papillomavirus
HSIL	High-grade squamous intra-epithelial lesion
LSIL	Low-grade squamous intra-epithelial lesion
NILM	Normal for intra-epithelial lesion or malignancy

Table 1

Image quality and patient characteristics

Total (N)	Image Quality			P-value
	Overall 690	High or Satisfactory 558	Not usable or missing 132	
Referral Cytology (N/Column%)				0.99
NILM/Other	8 (1.2)	7 (1.3)	1 (0.8)	
ASC-US	150 (21.7)	120 (21.5)	30 (22.7)	
ASC-H	86 (12.5)	68 (12.2)	18 (13.6)	
LSIL	280 (40.6)	231 (41.4)	49 (37.1)	
HSIL ⁺	206 (29.9)	164 (29.4)	42 (31.8)	
NA	46 (6.7)	36 (6.5)	10 (7.6)	
HPV Status (N/Column%)				0.98
HPV Negative	58 (8.4)	46 (8.2)	12 (9.1)	
HPV Positive, no HPV16	436 (63.2)	351 (62.9)	85 (64.4)	
HPV16 Positive	192 (27.8)	158 (28.3)	34 (25.8)	
NA	4 (0.6)	3 (0.5)	1 (0.8)	
Age (N/Column%)				0.90
18-20 years old	24 (3.5)	20 (3.6)	4 (3.0)	
21-24 years old	248 (35.9)	212 (30.7)	36 (27.2)	
25-29 years old	205 (29.7)	157 (22.8)	48 (36.4)	
30 years and older	204 (29.6)	162 (23.5)	42 (31.8)	
Missing	9 (1.3)	7 (1.3)	2 (1.5)	
Worst Biopsy Diagnosis (N/Column%)				0.98
CIN 1	409 (59.3)	327 (58.6)	82 (62.1)	
CIN 2	198 (28.7)	167 (29.9)	31 (23.5)	
CIN 3 ⁺	75 (10.9)	57 (10.2)	18 (13.6)	
AIS	1 (0.1)	1 (0.2)	0 (0)	
NA	7 (1.0)	6 (1.1)	1 (0.8)	

NILM: Negative for intraepithelial lesion or malignancy; ASC-US: Atypical squamous cells of undetermined significance; LSIL: Low-grade squamous intraepithelial lesion; HSIL: High-grade squamous intraepithelial lesion; HPV: Human papillomavirus. There is no statistically significant difference in the distributions of patient characteristics and histology outcomes between the three groups of different quality images.

Table 2

Agreement between live colposcopic impression and static image assessment

Colposcopic Impressions by Static Digital Image Assessment						
Live Colposcopic Impressions	Normal	Acetowhitening	Low-grade	High-grade	NA	Total (column %)
	16	16	6	2	0	40 (7.2)
	5	21	24	5	0	55 (9.9)
	12	79	113	59	0	263 (47.1)
	1	36	53	90	1	181 (32.4)
NA	3	2	7	5	2	19 (3.4)
Total (row %)		154 (27.6)	203 (36.4)	161 (28.9)	3 (0.5)	558

■ % Agreement = 43.0% ($\kappa = 0.20$, 95% CI: 0.14-0.26)

Live colposcopic impressions were rated with knowledge of patients' referral cytology and age, which were not known on static image assessment.

Overall agreement of rating dichotomized colposcopic impressions between live and static visualization

	N	Threshold for positive colposcopic rating								
		Acetowhitening or worse			Low-grade or worse			High-grade or worse		
		%	κ	95% CI	%	κ	95% CI	%	κ	95% CI
Overall	558	92.2	0.39	0.21-0.57	69.3	0.23	0.14-0.33	71.0	0.33	0.24-0.42

Table 4
Comparing discordant ratings of dichotomized colposcopic impressions using McNemar test of correlated proportions

Stratification	N	Threshold of positive colposcopic rating	Frequency of Positive Colposcopic Rating (%)		McNemar statistics		
			Live Colposcopy	Static Image Assessment	Difference (%)	P-value	Odds Ratio 95% CI
Overall	558	Acetowhitening or worse	92.6	93.7	1.1	0.44	1.3 0.7-2.5
		Low-grade or worse	82.3	65.4	16.9	<0.01	3.5 2.4-5.0
		High-grade or worse	33.5	29.0	4.5	0.07	1.4 0.99-1.9
Referral Cytology	351	Acetowhitening or worse	90.6	91.1	0.5	0.86	1.1 0.57-2.3
		Low-grade or worse	79.4	62.0	17.4	<0.01	3.4 2.2-5.2
		High-grade or worse	23.7	22.9	0.8	0.83	1.1 0.71-1.6
ASC-H/HSIL ⁺	164	Acetowhitening or worse	98.8	98.1	0.7	1.0	1.5 0.25-9.0
		Low-grade or worse	89.5	69.1	20.4	<0.01	5.1 2.4-11
		High-grade or worse	54.9	37.7	17.2	<0.01	3.3 1.8-6.4
HPV Status	351	Acetowhitening or worse	90.0	92.9	2.9	0.10	2.0 0.94-4.3
		Low-grade or worse	79.6	61.4	18.2	<0.01	3.8 2.4-6.1
		High-grade or worse	29.5	26.0	3.5	0.25	1.3 0.86-2.0
HPV 16 Positive	158	Acetowhitening or worse	97.4	95.4	2.0	0.45	2.5 0.49-13
		Low-grade or worse	88.9	75.2	13.7	<0.01	4.0 1.8-9.2
		High-grade or worse	45.7	39.2	6.5	0.22	1.45 0.85-2.50
Age Group	21-24 yo	Acetowhitening or worse	94.1	96.1	2.0	0.39	2.0 0.60-6.6
		Low-grade or worse	83.7	67.5	16.2	<0.01	4.0 2.1-7.7
		High-grade or worse	28.1	30.0	1.9	0.71	1.1 0.70-1.8
25-29 yo	157	Acetowhitening or worse	97.3	97.3	0	1.0	1.0 0.20-5.0
		Low-grade or worse	88.0	68.7	19.3	<0.01	4.2 2.0-8.7
		High-grade or worse	44.0	34.7	9.3	0.06	1.8 1.0-3.3
30 yo	162	Acetowhitening or worse	85.6	87.5	1.9	0.66	1.3 0.56-3.2
		Low-grade or worse	75.6	60.0	15.6	<0.01	2.7 1.5-4.8
		High-grade or worse	30.0	22.5	7.5	0.07	2.0 1.0-4.0

Table 5
Colposcopy accuracy for CIN2⁺ by live and static visualization

	N	Live Colposcopy						Static Image Assessment						P-value	
		Total	CIN 2+	Sensitivity (%)	95% CI (%)	Specificity (%)	95% CI (%)	Youden Index	Sensitivity (%)	95% CI (%)	Specificity (%)	95% CI (%)	Youden Index	Sensitivity	Specificity
Impression or worse	533	218	99.5	97.1-100	12.1	8.8-16.3	0.12	97.3	94.0-98.9	9.0	6.2-12.7	0.06	0.063	0.10	
Grade or worse	533	218	95.4	91.5-97.7	26.7	21.9-32.0	0.22	79.8	73.8-84.8	43.8	38.4-49.4	0.24	<0.01	<0.01	
Grade or worse	533	218	52.3	45.5-59.1	79.0	74.0-83.3	0.31	44.4	37.8-51.2	81.5	76.7-85.5	0.26	0.08	0.48	

Table 6

Comparison of colposcopy accuracy for CIN2⁺ stratified by referral cytology, HPV16 status, and age

Colposcopic Impression	Stratification	N	Live Colposcopy				Static Image Assessment				P-value	
			Total	CIN2 ⁺	Sensitivity	95% CI	Specificity	95% CI	Sensitivity	95% CI	Sensitivity	Specificity
Acetowhitening or worse	Overall	533	218	99.5	97.1-100	12.1	8.8-16.3	97.3	94.0-98.9	9.0	6.2-12.7	0.063
	Referral Cytology	347	104	99.0	93.9-100	12.8	9.0-17.8	95.2	88.6-98.2	10.2	6.9-14.9	0.38
	ASC-US/LSIL	161	101	100	95.4-100	3.3	0.6-12.5	99.0	93.8-99.9	3.3	0.6-12.5	1.0
	ASC-H/HSIL	337	119	99.2	94.7-100	15.1	10.8-20.8	97.5	92.4-99.4	10.2	6.7-15.1	0.04
	HPV Status	151	95	100	95.2-100	5.4	1.4-15.8	96.9	90.6-99.2	5.4	1.4-15.8	0.25
	HPV16 ⁺	203	86	100	94.7-100	10.3	5.6-17.6	97.7	91.2-99.6	5.7	2.5-11.8	0.11
	Age Group	149	69	100	93.4-100	5.0	1.6-13.0	98.6	91.3-99.9	4.8	1.6-12.5	1.0
	25-29 yo	158	52	98.1	88.4-99.9	19.8	12.9-28.9	94.3	83.4-98.5	15.1	9.1-23.7	0.36
	30 yo	533	218	95.4	91.5-97.7	26.7	21.9-32.0	79.8	73.8-84.8	43.8	38.4-49.4	<0.01
	Overall	347	104	95.2	88.6-98.2	27.2	21.8-33.3	78.8	69.5-86.0	44.7	38.4-51.1	<0.01
Low-grade or worse	Referral Cytology	161	101	95.0	88.3-98.2	20.0	11.2-32.7	77.2	67.6-84.7	43.3	30.8-56.7	<0.01
	ASC-H/HSIL	337	119	94.1	87.8-97.4	28.4	22.7-35.0	77.0	68.4-84.0	47.1	40.5-53.8	<0.01
	HPV Status	151	95	96.8	90.4-99.2	23.2	13.4-36.7	83.5	74.3-90.0	35.7	23.7-49.7	0.14
	HPV16 ⁺	203	86	95.3	87.9-98.5	24.8	17.5-33.8	85.1	75.4-91.5	43.9	35.1-53.1	<0.01
	Age Group	149	69	94.2	85.1-98.1	17.5	10.2-28.0	77.5	65.7-86.2	39.8	29.4-51.1	<0.01
	25-29 yo	158	52	96.2	85.7-99.3	34.0	25.2-43.9	77.4	63.5-87.3	47.2	37.5-57.1	<0.01
	30 yo	533	218	52.3	45.5-59.1	79.0	74.0-83.3	44.4	37.8-51.2	81.5	76.7-85.5	0.08
	Overall	347	104	39.4	30.1-49.5	82.7	77.2-87.1	36.5	27.5-46.6	82.8	77.3-87.2	0.74
	Referral Cytology	161	101	66.3	56.2-75.2	63.3	49.8-75.1	48.5	38.5-58.6	80.0	67.3-88.8	<0.01
	ASC-H/HSIL	337	119	47.1	37.9-56.4	79.8	73.7-84.8	45.9	36.9-55.1	84.9	79.4-89.2	0.87
High-grade or worse	HPV Status	151	95	58.9	48.4-68.8	75.0	61.4-85.2	43.3	33.4-53.7	67.9	53.9-79.4	0.04
	HPV16 ⁺	203	86	40.7	30.4-51.8	80.3	71.8-86.9	39.1	29.0-50.2	76.4	67.8-83.4	1.0
	Age Group	149	69	63.8	51.2-74.7	72.5	61.2-81.6	52.1	40.0-64.0	80.7	70.3-88.3	0.18
	25-29 yo	158	52	55.8	41.4-69.3	82.1	73.2-88.6	47.2	33.5-61.2	88.7	80.7-93.8	0.17
	30 yo	533	218	52.3	45.5-59.1	79.0	74.0-83.3	44.4	37.8-51.2	81.5	76.7-85.5	0.08
	Overall	347	104	39.4	30.1-49.5	82.7	77.2-87.1	36.5	27.5-46.6	82.8	77.3-87.2	0.74
	Referral Cytology	161	101	66.3	56.2-75.2	63.3	49.8-75.1	48.5	38.5-58.6	80.0	67.3-88.8	<0.01
	ASC-H/HSIL	337	119	47.1	37.9-56.4	79.8	73.7-84.8	45.9	36.9-55.1	84.9	79.4-89.2	0.87
	HPV Status	151	95	58.9	48.4-68.8	75.0	61.4-85.2	43.3	33.4-53.7	67.9	53.9-79.4	0.04
	HPV16 ⁺	203	86	40.7	30.4-51.8	80.3	71.8-86.9	39.1	29.0-50.2	76.4	67.8-83.4	1.0