Clinical Epidemiology and Global Health 7 (2019) 500–503

Contents lists available at ScienceDirect

Clinical Epidemiology and Global Health

journal homepage: www.elsevier.com/locate/cegh

Inter observer variability among gynecologists in manual cervix image analysis for detection of cervical epithelial abnormalities

Vidya Kudva[a,b], Shyamala Guruvare[c,*], Keerthana Prasad[a], Kiran Abhijit Kulkarni[d], Premalatha T S[d], Asha Kamath[e], Suma Nair[f], Chythra R. Rao[f]

[a] School of Information Sciences, Manipal Academy of Higher Education, Manipal, 576104, India
[b] Department of Electronics and Communications Engineering, NMAMIT, Nitte, 574110, India
[c] Department of Obstetrics and Gynecology, Kasturba Medical College, Manipal Academy of Higher Education, Manipal, 576104, India
[d] Department of Gynecologic Oncology, St. John's Medical College, Sarjapur Road, Bengaluru, Karnataka, 560034, India
[e] Department of Biostatistics, Prasanna School of Public Health, Manipal Academy of Higher Education, Manipal, 576104, India
[f] Department of Community Medicine, Kasturba Medical College, Manipal Academy of Higher Education, Manipal, 576104, India

1. Introduction

Cancer cervix is amenable to screening and early detection. Although Pap smear is the most widely used tool for cervical cancer screening, it has its own drawbacks in terms of low sensitivity (52%)[1] and the need for efficient networking between smear collection and cytology laboratories.[2] Besides, varying levels of sensitivity and specificity has been reported based on the grade of intraepithelial neoplasia [(sensitivity and specificity of 93%and 73% in low grade intraepithelial neoplasia (LSIL) versus 64% and 84% in high grade intraepithelial neoplasia (HSIL)][3]

Visual Inspection with Acetic acid (VIA) is a simple, cost effective test with a sensitivity that varies from 62.5 to 80%, and a specificity of 80–98.8% and is most suited for resource poor settings.[4–6] Depending on the skills of the staff who perform the test, accuracy of VIA varies widely; the screening performance of physician's assessment was significantly better than the nurse's.[7,8]

As none of the screening tools have been considered as optimal, researchers have even suggested a combination of VIA/Pap as it increased the sensitivity and specificity for detection of cervical cancer.[9,10]

A few researchers have tried the feasibility of using commercially available smartphones for acquisition and evaluation of VIA images. Their results have been similar to direct VIA, which shows a high inter observer variability.[11–14]

With this background in mind, the current study was designed to evaluate the inter observer variability among gynecologists in assessing cervical images acquired using a specified device.

2. Materials and methods

Institutional Ethics Committee approval was obtained before the conduct of the study. Written informed consent was obtained from the women participating in the study for image capturing and sharing among the Gynecologists. The study was conducted prospectively at two University affiliated healthcare settings in Southern India. The study participants were recruited at the rural health care centers where regular cancer screening camps are conducted. Married women aged above 25 years were invited to participate in the study and a total of 233 women consented. At first a Pap test was done using Ayres spatula for all the women and the sample was smeared over the glass slide and fixed with 95% alcohol solution. The slide was sent to cytology laboratory for Papanicoloau staining and cytological analysis. After clearing the cervical mucus using a wet cotton swab, cervix was inspected for abnormal areas. 5% acetic acid was applied on to the cervix and left for 1 min. Images of the cervix before and 1 min after the application of acetic acid were obtained using a device with 13Mp camera under uniform lighting condition with dual tone LED flash. An android application which was developed for image acquisition was used. The application had the options of acquiring and storing the images of participants' identification numbers, and images of cervix before and after the application of acetic acid.

The 233 VIA images were coded and shared with a team of seven gynecologists with different levels of expertise. An excel sheet that specified the Image code and the image interpretation as either VIA positive/negative or normal/abnormal and any additional remarks that the gynecologists offered to share was also shared for the purpose of reporting. Absolute confidentiality was ensured throughout the process of image capturing, sharing and reporting.

The gynecologists interpreted the image so captured and reported

* Corresponding author.
E-mail addresses: vidyakamathkudva@gmail.com (V. Kudva), shyamala.g@manipal.edu (S. Guruvare), keerthana.prasad@manipal.edu (K. Prasad), kiranabhijit@gmail.com (K.A. Kulkarni), premalatha_siddartha@yahoo.co.in (P. T S), asha.kamath@manipal.edu (A. Kamath), suma.nair@manipal.edu (S. Nair), chythra.raj@manipal.edu (C.R. Rao).

https://doi.org/10.1016/j.cegh.2019.05.009
Received 7 May 2018; Accepted 30 May 2019
Available online 30 May 2019

accordingly. Readings of one among the seven, a Gynaecologic Oncology expert, was considered as the gold standard. The expert has been working exclusively in this field for more than 10 years and is actively involved in cervical cancer screening camps and workshops. The expert reported the images as per the International Agency for Research on Cancer (IARC) practical manual on visual screening for cervical neoplasia.[15] Subsequently, these reports were grouped based on similar lines of International Federation for Cervical Pathology and Colposcopy (IFCPC) 2011 categorisation.[16]

Remaining six gynecologists were divided into three different groups. First group (G1 and G2) consisted of two general gynaecology experts at tertiary level with nearly 15 years of experience in general gynecology including gynecologic oncology, but not exclusively involved in oncology work. These practitioners are also involved in undergraduate and postgraduate (Obstetrics and Gynecology) training. The second group (G3 and G4) had two general gynecologists at secondary level with nearly 15 years of clinical experience in general gynecology; however, their encounter with women having gynecology/oncology problems was less in comparison to the first category. The third group (G5 and G6) consisted of two gynaecology residents who had exposure only in terms of assisting colposcopies at the tertiary hospital.

Cohen's kappa statistics was used to determine the inter observer variability in interpreting the positive and negative images. In addition to kappa, sensitivity and specificity of individual interpretation and intra class correlation coefficient were also calculated using SPSS statistical software version 15. Weighted kappa was estimated using the R software to determine the degree of agreement among the seven gynecologists. Kappa is interpreted as: < 0 Less than chance agreement; 0.01–0.20 Slight agreement; 0.21–0.40 Fair agreement; 0.41–0.60 Moderate agreement; 0.61–0.80 Substantial agreement and 0.81–0.99 Almost perfect agreement.

As per the standard protocol in the settings, if the Pap test reported a Low Grade Squamous Intraepithelial Neoplasia (LSIL), High Grade Squamous Intraepithelial Neoplasia (HSIL) or malignancy, women were further subjected to colposcopy directed biopsy. Based on the biopsy report further management was planned.

3. Results

Of the 233 cervix images that were shared for image interpretation, three were not reported and therefore 230 cases were considered for final analysis. Of these, majority of the study participants 219 (95%) were parous, while a good number 167 (72.6%) were aged less than 50 years and premenopausal 191 (83%).

Of the 230 images, 13 (5.6%) were VIA positive as per the gold standard (expert gynae-oncologist). These were categorised as: 5 major lesions, 6 minor lesions and 2 miscellaneous (one flat wart and the other a small aceto white area with peeling epithelium) lesions. Among the VIA negative images, 120 were interpreted as normal, 51 as squamous metaplasia, 24 as atrophic cervix, five as ectopy, and 17 as miscellaneous categories.

Two women in the group had abnormal Pap test (LSIL), of which one was diagnosed with malignancy and both of them were managed at the tertiary care centre as per standard protocols.

There was a wide range of sensitivity (57.1%–92.9%) and specificity (54.3%–94.5%) among individual gynecologists in image interpretation (Table 1 and Table 2).

Kappa value varied between 0.11 and 0.493 indicating slight to moderate agreement and the weighted Kappa value was 0.2 indicating an overall slight agreement among the specialists. Table 2 illustrates slight to fair agreement (kappa value 0.1, 0.16 and 0.3) between the 3 groups of specialists in comparison with the gold standard (expert gynae-oncologist) as evidenced by the difference in their experience and knowledge. Intra class correlation coefficient among the three groups were 0.397, 0.54, and 0.911 respectively.

Table 3 shows that a high false positive rate was observed while reporting 'normal' cervix and 'squamous metaplasia'. On reviewing the remarks provided by the gold standard it was found that among the 'normal' cervix images reported as VIA positive by the three groups of gynecologists, in 22% presence of mucus could have aided the false positivity. Likewise, 'squamous metaplasia' was the other condition with a higher false positivity, where acetowhiteness of gland openings and Nabothian cysts (64.4%) were the likely reasons behind false interpretation as VIA positive.

4. Discussion

The study was carried out to evaluate the inter observer variability in interpretation of cervix images among seven gynecologists of varying clinical expertise and experience. The results indicate that there was a wide variation in sensitivity (57.1%–92.9%) and specificity (54.3%–94.5%) among individual specialists in image interpretation. Kappa value varied between 0.11 and 0.493 indicating slight to moderate agreement among them.

Shankaranarayanan R et al. reported pooled sensitivity, specificity, positive and negative predictive values of 80%, 92%, 10% and 99%, respectively with VIA, for detecting cervical intraepithelial neoplasia grade 2 or worse lesions.[17] Ghosh et al. found visual inspection after Lugol's Iodine (VILI) (expansion) to be a useful test in screening for cervical neoplasia either as a single tool or in combination with Pap test. They found a high sensitivity of 100% and a specificity of 93.3% with VILI.[18]

In the present study evaluation was just restricted to VIA and VILI was not considered. Pap smear with its low sensitivity could not be considered as the gold standard and it would have been ideal to have colposcopy guided biopsy. However, the purpose of the analysis was to document the degree of variability in reporting cervix images and not testing the accuracy of reporting; hence the consideration of an expert as the gold standard. The gold standard had used strict criteria while reporting the images, while the others reported based on their learning and experience. Prior to reporting, none of the specialists in any of the three groups had any focused training regarding the interpretation of cervical findings. Our observations indicate the need for periodic training and skill updating of the health care providers in simple techniques such as the interpretation of VIA images.

However, from the reported literature it appears that training the raters also would not minimise the inter observer variability. In their study Vedantham et al. found the highest inter observer variability among the two Gynecologists who were the trainers for other four raters.[2] In their study there was a marked variation in the rates of VIA positivity among the six gynecologists who conducted the examinations. VIA positivity ranged from a low of 4% for two gynecologists to a high of 18% for one gynecologist and 31% for another gynecologist. The greatest variation was between providers 1 and 5, the two providers who had received the most training. Similarly in the present study, the group 1 specialists showed highest intragroup variation.

When the gold standard based on similar lines of IFCPC categorisation was compared with the false positive reporting of the specialists, the discrepancies noted were observed in categories, 'normal' and 'squamous metaplasia'. Gland openings, Nabothian follicles in the metaplastic zone were the common reasons for false positivity. In a small number even mucus had compromised the image quality and made the field appear 'white'.

Vedantham et al. found that among control group VIA positivity was 15.5% in women with inflammation and 6.1% in women without inflammation and the association was found to be statistically significant. (p value < 0.001).[2]

In their commentary on the issue of inter observer variability, Parashari A and Singh V attributed it to variation in man power training, light source used for visualization, and preparation of diluted (4–5%) acetic acid and its storage.[19] In the present study the issue of

Table 1
Statistics of individual raters from different groups (n = 228).

Parameters	General Gynecology Experts at Tertiary Level		General Gynecologists at secondary level		Gynecology residents	
	G1	G2	G3	G4	G5	G6
Sensitivity (%)	57.1	92.9	71.4	64.3	78.6	85.7
Specificity (%)	85.8	54.3	84.3	94.5	69.9	69.9
Negative Predictive Value (%)	96.9	45.7	84.3	97.6	98.1	98.7
Positive Predictive Value (%)	20.5	7.1	22.7	45.0	14.3	15.4
Kappa Value	0.234	0.110	0.278	0.493	0.156	0.177
Weighted Kappa Subjects = 228; Raters = 7; Kappa = 0.2						

Table 2
Agreement across various groups of gynecologists in comparison with the gold standard (expert gynae-oncologist).

Parameters	General Gynecology Expert at Tertiary Level (Group 1)	General Gynecologist at secondary level (Group 2)	Gynecology residents (Group 3)
Sensitivity (%)	75.0	67.9	82.2
Specificity (%)	70.1	89.4	69.9
Negative Predictive Value (%)	71.3	90.9	98.4
Positive Predictive Value (%)	13.8	33.9	14.9
Kappa Value	0.172	0.385	0.167
Intra class correlation coefficient (ICC)	0.397	0.54	0.911

Table 3
Cervix conditions with predominant false positive rate.

Evaluation Groups	Normal N (%)	Squamous Metaplasia N (%)	Atrophy N (%)	Ectopy N (%)	Miscellaneous N (%)
G1&G2 (n = 114)	47 (41.2)	52 (45.6)	07 (06.1)	05 (04.3)	3 (02.6)
G3&G4 (n = 38)	17 (44.7)	11 (28.9)	03 (07.8)	04 (10.5)	03 (07.8)
G5&G6 (n = 124)	50 (40.3)	42 (33.8)	14 (11.3)	08 (06.4)	10 (08.0)

light variation was overcome by the use of a single device with the same light source.

The challenges in introducing high-quality cytology screening in Low and Middle Income countries (LMIC) have led to the WHO recommendation of screen-and-treat programs using HPV tests and/or VIA followed by cryotherapy for comprehensive cervical cancer prevention and control programs in LMICs.[20,21] With such high degree of interobserver variability, using only VIA as screening tool may result in performing an unacceptable number of otherwise unindicated cervical ablative procedures in young women. Hence it is worth exploring the possibility of making the reporting more accurate and precise.

To minimise the interobserver variability, it is important to train the person doing VIA/image acquisition regarding the procedure, clearing of mucus completely without causing abrasions over the surface, specifications during image capturing and about the various categories when there is high possibility of false positive/negative reporting. To bring down the subjectivity, the need for a device with a decision support system that can report the cervix images, is emphasized which probably would provide a more objective and precise report.

5. Conclusion

There was significant inter observer variability among gynecologists with varying levels of expertise while reporting manual cervix image analysis for detection of cervical epithelial abnormalities. An automated system for cervix image analysis might be able to overcome inter observer variability, and provide for a more reliable and valid interpretation.

Conflicts of interest

The authors declare that they have no conflict of interest.

Ethical approval

Institutional Ethics Committee approval was obtained for this study and written informed consent was obtained from the women participating in the study.

Acknowledgements

This publication is made possible by a sub agreement from the Consortium for Affordable Medical Technologies (CAMTech) at Massachusetts General Hospital with funds provided by the generous support of the American people through the United States Agency for International Development. The contents are the responsibility of the University where the study is conducted and do not necessarily reflect the views of Massachusetts General Hospital, USAID or the United States Government.

We would like to acknowledge the support of Dr. Annekathryn Goodman, Associate professor of Obstetrics, Gynecology and reproductive Biology at Harvard Medical School, Boston, USA for the valuable inputs while designing the image capturing device.

References

1. Bhattacharya AK, Nath JD, Deka H. Comparative study between pap smear and visual inspection with acetic acid (VIA) in screening of CIN and early cervical cancer. *Midlife Health.* 2015;6(2):53–58.
2. Vedantham H, Silver MI, Kalpana B, et al. Determinants of VIA (visual inspection of the cervix after acetic acid application) positivity in cervical cancer screening of women in a peri-urban area in Andhra Pradesh, India. *Cancer Epidemiol Biomark Prev.* 2010;19(5):1373–1380.
3. Cobucci R, Maisonnette M, Macêdo E, et al. Pap test accuracy and severity of squamous intraepithelial lesion. *Indian J Cancer.* 2016;53:74–76.
4. Khodakarami N, Farzaneh F, Aslani F, Alizadeh K. Comparison of Pap smear, visual inspection with acetic acid, and digital cervicography as cervical screening strategies. *Arch Gynecol Obstet.* 2011;284(5):1247–1252.
5. Shastri SS, Mittra I, Mishra GA, et al. Effect of VIA screening by primary health

workers: randomized controlled study in Mumbai, India. *JNCI J. Nat. Cancer Institute.* 2014;106(3):dju009. https://doi.org/10.1093/jnci/dju009.
6. Veena R. VIA in cervical cancer screening. *IOSRJDMS.* 2012;1(1):1–4.
7. Raifu AO, El-Zein M, Sangwa LG, et al. Determinants of cervical cancer screening accuracy for visual inspection with acetic acid (VIA) and Lugol's iodine (VILI) performed by nurse and physician. *PLoS One.* 2017;12(1):e0170631.
8. Sherigar B, Dalal A, Durdi G, Pujar Y, Dhumale H. Cervical cancer screening by visual inspection with acetic acid-interobserver variability between nurse and physician. *Asian Pac J Cancer Prev APJCP.* 2010;11:323–326.
9. Ibrahim A, Aro AR, Rasch V, Pukkala E. Cervical cancer screening in primary health care setting in Sudan: a comparative study of visual inspection with acetic acid and Pap smear. *Int J Womens Health.* 2012;4:67–73.
10. Sherpa P, Pant AD. Correlation of visual inspection with cytological and histological findings in cervical neoplasia. *J Pathol Nepal.* 2017;7:1070–1073.
11. Quinley KE, Gormley RH, Ratcliffe SJ, et al. Use of mobile telemedicine for cervical cancer screening. *J Telemed Telecare.* 2011;17:203–209.
12. Catarino R, Vassilakos P, Scaringella S, et al. Smartphone use for cervical cancer screening in low-resource countries: a pilot study conducted in Madagascar. *PLoS One.* 2015;10(7).
13. Ricar GD, Wisniak A, Catarino R, et al. Use of smartphones as adjuvant tools for cervical cancer screening in low-resource settings. *J Low Genit Tract Dis.* 2015;19:295–300.
14. Bagga R, Suri V, Srinivasan R, et al. Feasibility of using smart phone camera as imaging device for screening cervical cancer in a low resource setting. *J Postgrad Med Educ Res.* 2016;50(2):69–74.
15. Sankaranarayanan R, Wesley RS. *A Practical Manual on Visual Screening for Cervical Neoplasia. IARC Technical Publication No 41.* IARC Press Lyon; 2003.
16. Bornstein J, Bentley J, Bosze P, et al. Colposcopic terminology of the international federation for cervical Pathology and colposcopy. *Obstet Gynecol.* 2011;120(1):166–172 2012.
17. Sankaranarayanan R, Nessa A, Esmy PO, Dangou JM. Visual inspection methods for cervical cancer prevention. *Best Pract Res Clin Obstet Gynaecol.* 2012;26(2):221–232.
18. Ghosh P, Gandhi G, Kochhar PK, Zutshi V, Batra S. Visual inspection of cervix with Lugol's iodine for early detection of premalignant & malignant lesions of cervix. *Indian J Med Res.* 2012;136(2):265–271.
19. Aditya P, Veena S. Reasons for variation in sensitivity and specificity of visual inspection with acetic acid (VIA) for the detection of pre- cancer and cancer lesions of uterine cervix. *Asian Pac J Cancer Prev APJCP.* 2013;14(12):7761–7762.
20. World Health Organization. *WHO Guidelines: WHO Guidelines for Screening and Treatment of Precancerous Lesions for Cervical Cancer Prevention.* Geneva, Switzerland: WHO; 2013.
21. Sankaranarayanan R. Screening for cancer in low- and middle-income countries. *Annals of Global Health.* 2014;80:412–417.