Agreement rates between AI classification of cervical images and clinical expert visual impressions in high- and low-resource settings

D. Levitz[1]

MobileODT Ltd., Ben Avigdor 8, Tel Aviv 6721832 Israel

ABSTRACT

Methods based on artificial intelligence (AI) have recently attracted a lot of attention, as they offer a mechanism to overcome the shortage of highly qualified expert in making clinical decisions on diagnosis and treatment. Many diagnostic tests based on AI have been validated against panels of experts. Within the field of cervical cancer screening and management, the combining AI with digital cervicography (DC) has appeal, as it can provide a rapid, accurate result with minimal inter-operator bias. Recently, a cloud-based classifier – VisualCheck™ – was integrated into a commercial DC system. To validate that VisualCheck™ works in different types of health systems, local expert panels consisting of 3-5 reviewers were recruited in South Korea, Poland, and India to conduct a retrospective review of local images and their VisualCheck™ scores. Each expert reviewed images from N=600 patients, half of which tested positive with VisualCheck™, and half that tested negative. Reviewers noted whether they agreed with the VisualCheck™ result for each patient exam, as well as provided annotations on visual features in the images. Agreement rates and concordance were calculated for between the majority panel opinion (MPO) and the VisualCheck™ results, as well as between pairs of individual reviewers. An agreement rate of 70% or above was set as the acceptance criteria. The final agreement rates between the MPO and VisualCheck™ results were 84%, 76%, and 74% for South Korea, Poland, and India, respectively, while the final concordance between the MPO and VisualCheck™ results were 73%, 51%, and 54% for South Korea, Poland, and India, respectively. Similar agreement rates (0.65-0.95) were observed between individual expert pairs. Concordance values ranged between 0.05-0.48. Both agreement rates and concordance between experts were consistent with values reported in the scientific literature. Collectively, these results demonstrate robust performance of VisualCheck™, with results that are comparable to local expert opinion.

Keywords: cervical cancer screening, deep learning, automated visual evaluation

1. INTRODUCTION

Global epidemiological research shows that there are over 300,000 annual deaths caused by cervical cancer each year, and that this rate has not decreased in decades despite sustained screening and vaccination efforts [1]. The main burden of cervical cancer disproportionately affects women in low and middle-income countries (LMICs), where the health systems do not have sufficient resources to contain the disease. As a result, cervical cancer is still the second most common cause of female cancer deaths in LMICs [2]. In high income countries such as the United States, cervical cancer primarily affects women in underserved communities that are missed by the health system [3]. Most women in under-resourced health systems are not regularly screened for cervical cancer. And of those who do actually get screened and test positive, only a small fraction make it to the follow up procedure, as high loss to follow-up (LTFU) rates prevent patients from receiving the care they need [4].

Screening and management of cervical cancer patients have several challenges. The most common screening test

[1]All correspondence should be sent to liorshwarzwald@mobileodt.com

worldwide – cervical cytology – requires the scraping of cells off the surface of the cervix and sending them to a remote lab for further processing [5]. Such a process can take a couple of weeks before there is an answer. Even routine sample processing and slide reading could lead to quality issues in resource constrained health systems, where there are few central labs and no expertise in reading and interpreting cytopathology slides. Another way to screen for cervical cancer is to test for the presence of human papillomavirus (HPV), the virus that causes over 99% of cervical cancers [6]. Like cytology, HPV tests also require physical sampling of the cervix and/or vagina. Due to better reliability (namely, a very high negative predictive value), HPV tests are slowly becoming the primary screening test in developed health systems. However, HPV testing also required specialized equipment and personnel to operate the machine reading the assay [7], and the result still can take days. New HPV tests are now available that can run at the point of care (PoC) [8], yet operational issues persist. And importantly, both cervical cytology and HPV testing rely on a secondary procedure to complete a diagnosis - colposcopy - the magnified examination of the cervix by an expert clinician. In health systems with less resources, visual inspection with acetic acid (VIA) is still very commonplace [9]. In VIA, acetic acid is applied on the cervix using an elongated swab, and the provider assesses whether or not the cervix turned opaque white, in a process known as acetowhitening. VIA has the advantage of providing an immediate answer, which allows treating patients in the same visit as the diagnosis, and thus eliminating loss to follow up. However, VIA is not reproducible, as the performance is heavily reliant on the quality of the procedure and expertise level of the provider [10]. Digital cervicography (DC), a variation of VIA that involves capturing digital documentation using a camera, and external quality review with local experts. Although DC has reduced operator variability, and led to scaled implementation in Zambia [11], as well as also in South Korea [12].

A result of all these diagnostic tests is that the most common screening methods in both high- and low-resource settings (cytology and VIA, respectively) all depend on subjective visual interpretation. Automation would help reduce inter-operator variability and help improve test accuracy and reproducibility. Artificial intelligence (AI) methods have recently been proposed to mitigate such performance variations when visual interpretation is required [13]. In a number of medical fields, there has been significant developments of late testing AI to address the challenges inherent with visual screening. These image classifiers are based on deep learning, a subset of AI, have succeeded in many different medical applications, including detecting skin cancer [14] and diabetic retinopathy [15], predicting stroke [16], and assessing bone health [17] and diagnostic mammograms [18]. All of these studies relied on image interpretation and review by a panel of experts as their ground truth.

Recently, several studies have attempted to assess the use of AI in cervical cancer care. Automated visual evaluation – using AI to interpret DC images – has been recently compared to cytology, VIA, and HPV tests [19] using data from a natural history study back in the 1990s [20, 21]. Using cervical intraepithelial neoplasia (CIN) grade 2 or worse (CIN 2+) histopathology as an endpoint, the results showed 13-20% improvement in the area under the ROC (receiver operating characteristic) curve (AUC), thus serving as a valuable proof of concept for AVE as a technology that can be used to detect cervical precancerous lesions. However, the data for this study came from a single clinic, and the images were digitized version of images captured on film with a cerviscope – a device consisting of an SLR camera and an auxiliary lens that has been discontinued almost 20 years ago. A follow up AVE study [22] built a classifier from digital images captured by the Enhanced Visual Assessment (EVA) System (MobileODT), a commercially available mobile colposcope built around a smartphone platform, using visual impressions by an expert panel as ground truth. The images came from 17 countries where EVA was used in routine clinical care. Using these data, the authors achieved performance that was almost as good as the original AVE study (AUC>0.9). The classifier was then retrained using a different neural network architecture in order to optimize it for running on the cloud, and then was linked up with the EVA System image portal, a digital documentation tool that is part of the EVA System. The resulting classifier and workflow – VisualCheck™ – were piloted in several sites across the globe, including South Korea, Poland, and India. Despite these small pilots, no wide-scale studies comparing the output of VisualCheck™ to local experts were done.

In this study, we retrospectively compare VisualCheck™ results from three sets of cervical images captured during routine care in India, South Korea, and Poland to the majority decision of panels consisting of local experts. These three health systems represent different levels of resources and clinical workflows. Each national data set contained images from N=600 patients, half of which were VisualCheck™-positive (VC+), for a total of 1800 patients. In all

three national data sets, the agreement rates were higher than 70%, and were comparable to the agreement rates between individual expert reviewers (65-94%). Concordance values between the panel majority and VisualCheck™ were higher than 50%. Key features of interest for DC were tabulated and were consistent with the VisualCheck™ outputs. These data show that the VisualCheck™ classifier is robust enough to have similar performance in very different clinical settings, suggesting that AI holds promise as a decision support tool for clinicians performing visual cervical evaluations, either colposcopy or VIA.

2. METHODS

2.A. Image classifier components

The classifier used in this analysis was built in partnership with the National Cancer Institute (NCI) and National Library of Medicine (NLM). Details are provided in Ref. 22. Briefly, a set of 7585 images was reviewed by three world-renowned colposcopists, labeling each image as one of the following: (1) Probably High-grade, defined by cervical intraepithelial neoplasia grade 2 or worse (CIN 2+); (2) Probably Low-grade or normal; (3) Possibly High-grade; (4) Uncertain (image quality was deemed insufficient to allow for accurate categorization); or (5) Postcryotherapy, where the cervix had previously been ablated, and accurate categorization was not possible. Using these images, a classifier was trained and validated that achieved impressive performance, including an ROC AUC above 0.90 for identifying the likely cervical precancer cases from controls [22].

Expanding on that initial study, a new clinical workflow was developed that included factoring in additional parameters for image classification. These include a secondary classifier that filters out poor quality images [23, 24], which runs prior to any assessment of tissue pathology. This qualifying classifier ensured that only high-quality images most suited for machine learning classification were used in the subsequent automated analysis. The main cervical image classifier - VisualCheck™ – was retrained and optimized to run on Tensorflow framework in a more time-efficient and scalable architecture.

Another key addition to the classifier presented in Ref. 22 is that the original classifier gave a predicted value per patient (as only one image per patient was used), while the algorithm developed for this study was designed to classify multiple cervical images in a single patient exam. Here, each image in the patient exam was processed separately, resulting in one prediction score value per image. Scores above 0.5 were considered positive. A weighted average was then used to combine prediction values from multiple images, resulting in one aggregate score per exam. The weight of each image came from the quality score provided by the qualifying classifier. Cross-validation of this process end to end on all images from the original set produced an AUC of 0.89.

2.B. Images

All images used in this retrospective analysis were collected by the EVA System during routine clinical use. The EVA System combines a mobile digital colposcope with an integrated online documentation portal, uploading images to a secure cloud server for each patient automatically. This made image selection straightforward. In each of the countries from which images were collected – Poland, South Korea, and India – clinicians were using EVA either as a colposcope, or as a DC image capture device. Because of differences in these health systems, Polish providers primarily used the device as a colposcope while Korean providers primarily used it for DC at screening, whereas Indian providers were mixed in their use of the device.

Altogether, anonymized, deidentified EVA images from N=1800 patients were randomly selected from the 3 health systems, 600 from each. For every patient, all non-deleted images from the patient exam were analyzed with VisualCheck™ and aggregated into a single prediction score for the exam, as described in Section 2.A. In each country, 300 patients in the pool were VC+ while the other 300 were VC-. The distributions of VisualCheck™ scores (probability density functions, or PDFs) of VisualCheck™ scores for the images used in the three data sets study are shown in Fig. 1.

2.C. Image review panels

Three separate panels of clinical experts were used to review and evaluate images from the 3 countries. Each "national panel" only evaluated images from that specific country. In each country, experts were chosen for review. Experts that did not fully review the images themselves were excluded. The final number of reviewers was 5 in Poland, 3 in South Korea, and 4 in India. In the analyses below, expert answers from within a national panel were compared against one another, with both agreement rates and inter-observer concordance coefficient (□) calculated [25].

During image review and evaluation, experts evaluated images on a per patient basis, while also providing general annotations on the images and procedure. Specifically, the panels were shown all the images from a single patient, along with the aggregate VisualCheck™ results, and they noted whether or not they agreed with VisualCheck™ output. Additionally, they provided annotations on the patient's image set that included cervix visibility, full squamo-columnar junction (SCJ) visibility, the presence of acetowhitening, the presence of vascular features (punctation, mosaicism, atypical vessels), colposcopic impression (normal, low grade lesion, high grade lesion, or cancer), whether or not it was a satisfactory image, whether or not Lugol's images were included, and the availability of results from endocervical curettage (ECC) test. The key differences between colposcopic impression label and agreement with VisualCheck™ label is that colposcopic impression labels permitted the reviewer to not answer, while the agreement data forced them to make a decision for each patient. The analysis presented here focused on the agreement question in order to avoid any bias caused by variations in how experts define the various colposcopic impression choices, specifically low-grade lesions.

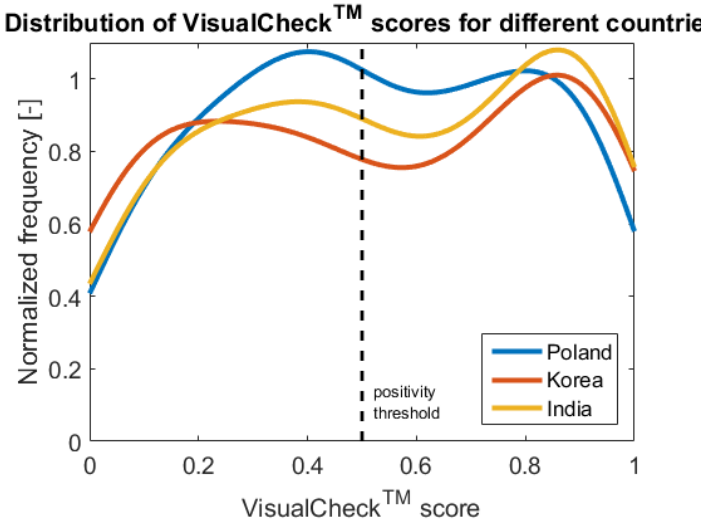


Fig. 1: Distribution of VisualCheck™ prediction scores for images reviewed by panel.

2.D. Data analysis

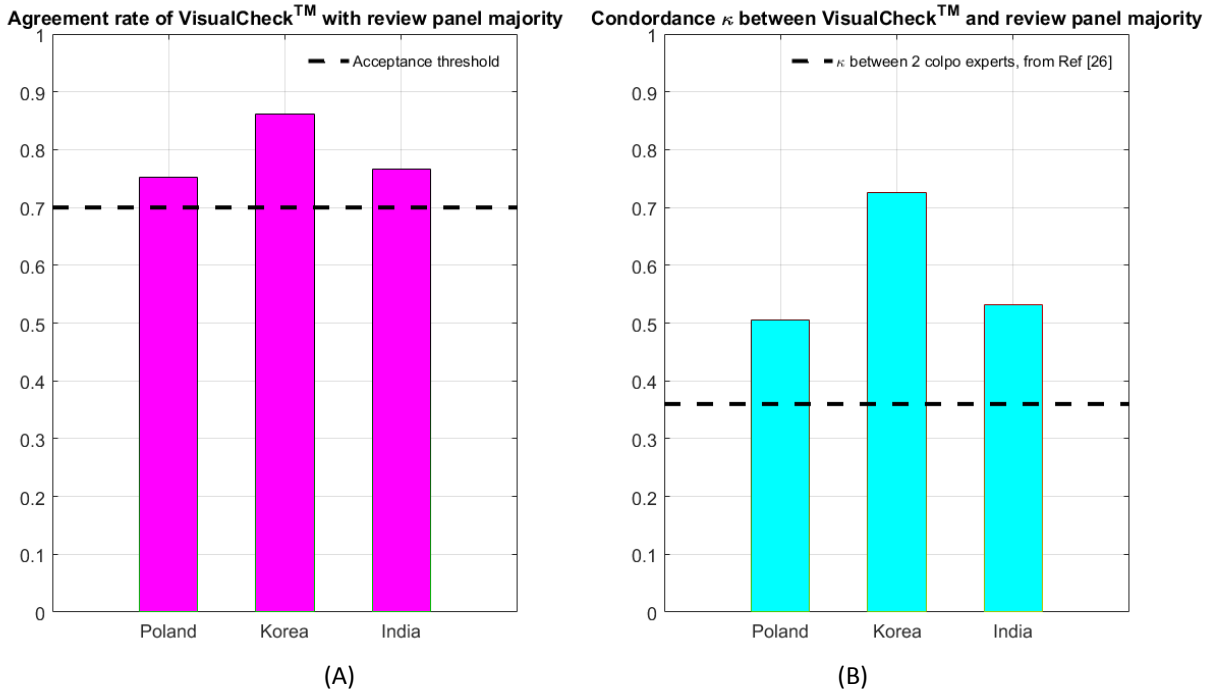
Image reviews and annotations from the 3 panels were analyzed separately for each country. Analyses were done using both Excel and Matlab. For each country, the majority panel opinion (MPO) on agreement rate and other parameters were calculated on a per exam basis. To calculate the MPO for a given parameter, positives and negatives are assigned ones and zeros, respectively, and averaged arithmetically. Thus if 3 reviewers agreed and 2 disagreed, the rate was 0.6. If 2 reviewers agreed, 2 disagreed, and another did not provide an answer, the agreement rate would be 0. 5. Both VisualCheck™ scores and MPO scores ranged between zero and one. MPO values above a threshold of 0.5 were considered positive.

The primary endpoint of this analysis is agreement rate, with the acceptance criteria defined as agreement rates between the VisualCheck™ score and MPO being above 0.7. Several secondary endpoints were also considered, including the agreement rates between reviewers which assesses how many patients the reviewers classified identically out of the total pool, together with their concordance, which assesses inter-reviewer agreement while also accounting for reviewers guessing some of their answers, given by Cohen's □ statistic [25]. And in addition to

these primary and secondary endpoint data, the additional parameter labels provided by the panel were also analyzed. For these features, the MPO was one of 4 classes: “Yes”, “No”, “no answer”, and “Unclear”, with the last class representing exams without agreement by the majority of the panel.

3. RESULTS

A comparison of the MPO agreement rate to the VisualCheck™ score for all 3 countries is shown in Fig. 2. The dashed line in Fig. 2A denotes the minimum acceptance criteria for agreement rate of 70%; the dashed line in Fig. 2B denotes the literature value of concordance between 2 colposcopy experts [26]. results can be interpreted as follows: values ≤ 0 as indicating no agreement and 0.01–0.20 as none to slight, 0.21–0.40 as fair, 0.41– 0.60 as moderate, 0.61–0.80 as substantial, and 0.81–1.00 as almost perfect agreement [25]. It can be seen that in each of the 3 national image sets analyzed, the 70% agreement threshold was met, with only minor variations observed in agreement rates between the 3 countries. Moreover, in each of the 3 countries the concordance between the panel majority and VisualCheck™ was notably higher than the concordance between 2 colposcopy experts [26].



(A) (B)

Fig. 2: Agreement rate (A) and concordance (B) comparing panel majority to VisualCheck™ score for 3 countries.

As a secondary analysis, the pairwise agreement and concordance rates were calculated for all reviewer pair combinations within the 3 panels (Table 1). It can be seen that the ranges of agreement and concordance rates varied between the 3 countries, with ranges of 0.75-0.88, 0.65-0.90, 0.82-0.95 for Poland, South Korea, and India, respectively. Concordance values were lower than the agreement rates, and in general were much more uniform across the 3 national data sets, ranging from 0.05–0.36, 0.09–0.24, and 0.06–0.42 in Poland, South Korea, and India, respectively.

Table 1: Agreement and concordance values calculated between reviewer pairs on each of the three panels.

Reviewer ID	Reviewer ID	Agreement	Concordance	Reviewer ID	Reviewer ID	Agreement	Concordance
Poland				South Korea			
1	2	0.7982	0.2807	6	7	0.8975	0.1693
1	3	0.8021	0.1300	6	8	0.6959	0.2423
1	4	0.7893	0.3206	7	8	0.6503	0.0899
1	5	0.7542	0.2898	India			
2	3	0.8789	0.0489	9	10	0.8201	0.1145
2	4	0.8333	0.3035	9	11	0.8231	0.2047
2	5	0.7648	0.1906	9	12	0.8200	0.1276
3	4	0.8377	0.1262	10	11	0.9492	0.4229
3	5	0.7962	0.1638	10	12	0.9461	0.1601
4	5	0.7987	0.3568	11	12	0.9256	0.0609

In order to assess the inherent variation in image quality and ability to discern features of interest, the prevalence of features (represented by MPO) across the 3 national data sets were also compared. Three cervical features of interest were chosen and are presented in Fig. 3: SCJ visibility, acetowhitening, and vascular features such as punctation, mosaicism, and atypical vessels. It appears as if the SCJ (Fig. 3A) was more clearly visible in images in the Korean data set than in the Polish or Indian images. Acetowhitening (Fig. 3B) was much more balanced across the three national sets. Vascular features (Fig. 3C) were absent in the Polish and Korean images, while there were mixed reviews regarding the Indian images.

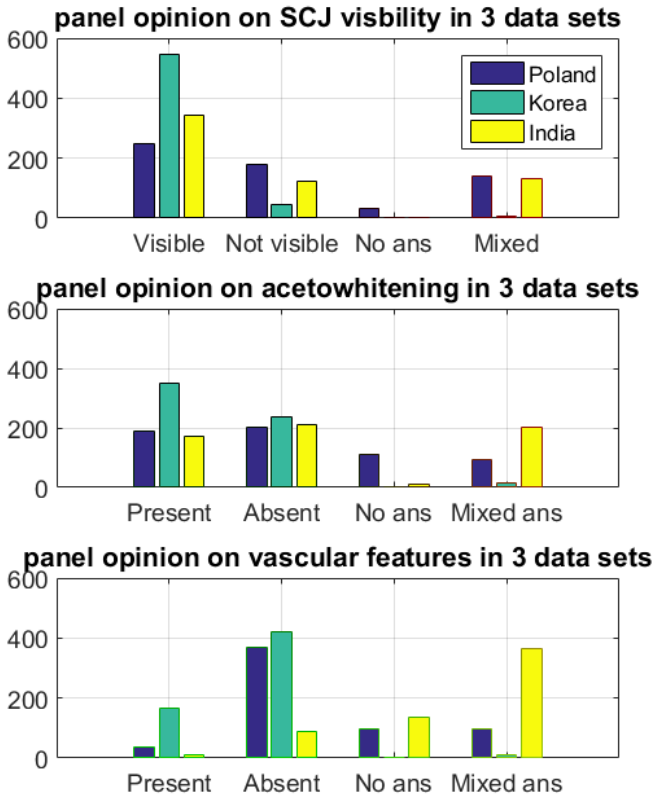


Fig. 3: Analysis of the presence of features in the three national data sets.

4. DISCUSSION

In this study, three sets of 600 anonymized, deidentified EVA images (300 VC+, 300 VC-) were selected at random from a large existing database. Image sets were collected from routine clinical use of the EVA System in 3 countries: Poland, South Korea, and India. Images from a particular set were reviewed by panels of local experts, consisting of 3-5 reviewers. Reviewers were presented with images and the VisualCheck™ diagnosis, and they then noted whether or not they agreed with the diagnosis. They also annotated the images for common visual colposcopic features. Agreement rates between majority panel opinion and the VisualCheck™ score showed that in each national data set, agreement rates were above 70%, meeting the predefined acceptance criteria (Fig. 2A), and were also higher than the 36% concordance between 2 colposcopy experts [26] (Fig. 2B). Agreement rates between individual reviewers ranged between 0.65-0.94 in all 3 national data sets, while concordance rates ranged between 0.04-0.42 (Table 1). In comparing visual features across the 3 data sets, SCJ visibility was best in Korean images, acetowhitening was split evenly across the three data sets, and some vascular features were present in the Indian images (Fig. 3). Overall, these results show consistent classifier performance across 3 geographies when compared to local expert opinion.

The most significant result of this study is the consistent agreement rates between VisualCheck™ score and expert review panels in three separate geographies, with very different patient populations, health systems, and access to resources (Fig. 2). The agreement rates between MPO and VisualCheck™ is comparable to the agreement rates among experts themselves, as shown in our data (Table 1), and is consistent with values that have been reported in the literature [26]. Similarly, the comparable concordance values between MPO and VisualCheck™ score (Fig. 2B), as well as between experts on the panels, also fit within previous values reported in the literature (0.36) [26]. Such values are not uncommon in medicine, and can be expected when the reviewers have different levels of training, different sub-specialization, or when their routine practice is on different segments of the population. For example, one recent study in evaluating bone density X-Rays produced concordance values that ranged from 0.15 to 0.90 [27]. The consistency of these numbers with others reported in the literature confirm the primary outcome of the study, and demonstrate the overall robustness of the VisualCheck™ classifier.

Several insights can be obtained in examining the visual colposcopic features identified by the MPO in the images across the 3 data sets (Fig. 3). For example, acetowhitening is relatively balanced between present and absent in the three data sets (Fig. 3B), demonstrating that it is indeed a determining factor in how VisualCheck™ identifies precancerous lesions. Because the data was partitioned by VisualCheck™ scores (50% above 0.5, 50% below 0.5, see Fig. 1), the comparable frequency makes intuitive sense. SCJ visibility was best in the Korean images, far better than in Poland and India (Fig. 3A). Images from South Korea come from a screening population of private doctors in high resource clinics. Moreover, DC followed by expert review is routinely used alongside cytology in Korean private clinics [12], thus providers use EVA more routinely than providers in Poland (where some or all of the images come from colposcopy clinic that uses EVA one day a week), as well as some of those in India. Finally, vascular features (punctation, mosaicism, atypical vessels) appeared to be present more frequently in India than in South Korea and Poland, as indicated by the large fraction of mixed reviews (Fig. 3C). In previous studies, Indian experts identified high rates of cervicitis and other inflammatory conditions in EVA data sets [22, 28, 29]. Thus, the high amount of "potential" vascular features also makes intuitive sense, as some experts disagree on whether or not to attribute a vascular feature to cervicitis or precancerous lesions [30].

It's worth noting that the VisualCheck™ classifier was trained on images from 17 different countries, including India, but not Poland and South Korea, yet performance in the three data sets was similar (Fig. 2A-B). In other words, South Korea and Poland represented external test sets. However, in general EVA images from India are quite heterogeneous – acquired in different parts of the country with different EVA phone models, by providers with vastly different levels of training, and involving a mixture of screening and colposcopy patient populations. The Indian images used to train the classifier represent a subset of this larger data set, as they came from old EVA images [28, 29] (captured mostly in 2017). This could potentially explain why performance was not better in India than in Poland and South Korea.

In comparing the analyses performed on the three national data sets, several factors could have affected the results. One group of factors involves the images, the patients, and the PoC providers that captured them. All providers capturing images with EVA undergo the same training modules initially, but their experience in image capture is very personal and affected by their surroundings, as well as the clinical workflow they follow. For example, in some Indian clinics, there were multiple people in the room during the exam – nurses, physician assistants, and at times even other patients (as in screening camps). In all 3 countries, there is ambiguity in the guidelines regarding age (postmenopausal women were examined) and pregnancy [31-33], which likely added some variation, although it is unclear how much. Additionally, there were also differences in disease prevalence for both CIN and other diseases. Korean screening patients had lower disease prevalence, while Polish colposcopy patients had higher disease prevalence. Indian data was much more heterogeneous – both screening, triage, and colposcopy patients that come from private clinics, hospitals, and screening camps. These variations all affected the images, both in terms of general image quality (sharpness, framing), and the quality of the clinical procedure (full visibility of the transformation zone, correct timing of image capture relative to acetowhitening, etc.). Some of these effects contributed to the feature frequency data shown in Figs. 3A-C.

The other group of factors that contributed to the heterogeneity of the data was the reviewers on the three panels. While only well-known local experts were recruited to join the panels, they have different backgrounds and expertise levels. For example, Polish reviewers ranged from 5+ years' experience to 20+ years' experience. These differences led to disagreements among members of the panel, who were blinded to the others' answers. And because image interpretation is subjective, it is impossible to objectively assess when the experts were right and when they were wrong. This issue is a common challenge in radiological applications when reader panels are used [27, 34].

In looking at the distribution of VisualCheck™ scores (Fig. 1), it can be seen that the images in the three national data sets had different probabilities for "disease". A very high VisualCheck™ score means the classifier is confident of the presence of disease, while very low scores mean the classifiers is confident of the absence of disease, and there are many shades of gray in between. All three data sets had a peak above 0.8, meaning that the VC+ images were clearly positive. However, the negatives were different, as Korean data had a peak around 0.2-0.3, while India and Poland had peaks around 0.4, which is close to the disease threshold. This suggests that it should be "easier" for a review panel to classify the Korean images, in comparison to the Polish and Indian images. This may explain the slightly higher agreement rate and concordance seen in the Korean review (Fig. 2A-B). Further research is required to test this assertion.

Considering the huge sources of variation in the data – patient populations, PoC provider training in EVA, and review panels' experience – the consistent agreement rates between local experts and VisualCheck™ is notable. In all three national data sets, the agreement rates between VisualCheck™ and the MPO was comparable to agreement rates among the experts themselves. Concordance results were similar. The appeal of VisualCheck™ as a classifier is that it can minimize the effects of inter-operator bias. The results in Fig. 2A-B verifies that it can act as a virtual replacement for a local panel of experts

The study described here had several limitations. The current analysis did not include histopathology-correlated images but instead relied on reader panels. Thus, there was no ability to compare the numbers directly against ground truth and calculate sensitivity, specificity, ROC AUC, and other predictive metrics. Including histopathology results, even on a fraction of the data sets, would have enabled assigning weights to reviewers and arriving at a more robust MPO. Another major limitation was that the reviewers from all sites were not compared against one another on a set of tests. Reviewers only received images from their national data set, but no objective assessment of reviewer expertise was done across the panels using a standardized set of images. As a result, the panels and MPO from each country were only compared against VisualCheck™, but not directly against one another, and this limited the scope of the conclusions that could be drawn. Verification bias [35] was not properly accounted for, which results from priming the reviewer with the VisualCheck™ diagnosis. Similarly, assessments of intra-reviewer agreement / concordance [36] calculated on a subset of the data could provide additional insight about individual reviewers. These parameters should be quantified and corrected for any offsetting effect they may have had on the results. Future studies will build on the current study by looking at countries in additional continents, including

South Africa and Brazil. A slightly refined design will address the limitations described above by including a standardized subset of biopsy correlated images, and a second subset to account for verification bias.

5. CONCLUSION

In conclusion, in three separate national data sets of cervical EVA images, there was agreement between the VisualCheck™ score and majority of an external review panel of local experts in over 70% of cases, and concordance in 50% of cases. Agreement rates and concordance among the reviewers on the panel were comparable to the agreement rates with the VisualCheck™ scores. The consistently high agreement rates between the VisualCheck™ score and local panel of experts suggests that VisualCheck™ could potentially act as an alternative for local expert opinion.

ACKNOWLEDGEMENTS

The authors would like to thank image reviewers who analyzed images and provided clinical impressions that served as the basis of this paper. In Poland, the reviewers were: Dr. Grzegorz Glab of GMW Spółka Partnerska, Opole; Dr. Zbigniew Wróbel of Medi-Lab S.C., Świdnica; Dr. Bogdan Kałużewski of NZOZ Genos S.C., Łódź; Dr. Wojtek Homola of FemiMea Klinik, Wrocław; and Dr. Bartłomiej Barczyński of the Clinical Gyencology-Oncology Department of the University Hospital in Lublin, Lublin. In Korea, the reviewers were: Dr. Yeo Kyeong of Y Women's Clinic, Seoul; Dr. Kim Jeongyuen of Jiin Clinic, Seoul; and Dr. Kwon Soyoung of Riz Ob/Gyn Clinic, Seoul. In India, the reviewers were: Dr. Leela Digumarti of Homi Baba Cancer Hospital and Research Centre, Visakhapatnam, Dr. Mallika Samuel of Apollo Hospital and Apollo Research Center, Chennai; Dr. Jaideep Malhotra of Rainbow IVF, Agra; Dr. Ramesh Maturi of Parijatham women wellness clinics, Hyderabad; and Dr. Renuka Matti of Hiranandani Hospital, Mumbai. This study was funded by MobileODT.

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018 Nov;68:(6):394-424. (http://www.ncbi.nlm.nih.gov/pubmed/30207593)

2. de Martel C, Plummer M, Vignat J, Franceschi S. Worldwide burden of cancer attributable to HPV by site, country and HPV type. International journal of cancer. 2017 Aug 15;141(4):664-70. (https://www.ncbi.nlm.nih.gov/pubmed/28369882)

3. https://www.cancer.org/content/dam/cancer-org/research/can-cer-facts-and-statistics/cancer-facts-and-gures-for-african-americans/cancer-facts-and-gures-for-african-americans-2016-2018.pdf [retrieved August 2018]

4. McKee MD, Lurio J, Marantz P, Burton W, Mulvihill M. Barriers to follow-up of abnormal Papanicolaou smears in an urban community health center. Archives of Family Medicine. 1999 Mar 1;8(2):129.

5. Solomon D, Davey D, Kurman R, Moriarty A, O'Connor D, Prey M, Raab S, Sherman M, Wilbur D, Wright Jr T, Young N. The 2001 Bethesda System: terminology for reporting results of cervical cytology. Jama. 2002 Apr 24;287(16):2114-9.

6. Ronco G, Dillner J, Elfström KM, Tunesi S, Snijders PJ, Arbyn M, Kitchener H, Segnan N, Gilham C, Giorgi-Rossi P, Berkhof J. Efficacy of HPV-based screening for prevention of invasive cervical cancer: follow-up of four European randomised controlled trials. The lancet. 2014 Feb 8;383(9916):524-32.

7. Jeronimo J, Holme F, Slavkovsky R, Camel C. Implementation of HPV testing in Latin America. Journal of Clinical Virology. 2016 Mar 1;76:S69-73.

8. Kelly H, Mayaud P, Segondy M, Pai NP, Peeling RW. A systematic review and meta-analysis of studies evaluating the performance of point-of-care tests for human papillomavirus screening. Sexually transmitted infections. 2017 Dec 1;93(S4):S36-45.

9. Denny L, Kuhn L, De Souza M, Pollack AE, Dupree W, Wright TC. Screen-and-treat approaches for cervical cancer prevention in low-resource settings: a randomized controlled trial. Jama. 2005 Nov 2;294(17):2173-81. (PMID: 16264158)

10. Mwanahamuntu MH, Sahasrabuddhe VV, Kapambwe S, Pfaendler KS, Chibwesha C, Mkumba G, Mudenda V, Hicks ML, Vermund SH, Stringer JS, Parham GP. Advancing cervical cancer prevention initiatives in resource-constrained settings: insights from the Cervical Cancer Prevention Program in Zambia. PLoS medicine. 2011 May 17;8(5):e1001032. (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3096609/)

11. Bateman AC, Parham GP, Sahasrabuddhe VV, Mwanahamuntu MH, Kapambwe S, Katundu K, Nkole T, Mulundika J, Pfaendler KS, Hicks ML, Shibemba A. Clinical performance of digital cervicography and cytology for cervical cancer screening in HIV-infected women in Lusaka, Zambia. Journal of acquired immune deficiency syndromes (1999). 2014 Oct 1;67(2):212.

12. Bae SN, Kim JH, Lee CW, Song MJ, Park EK, Lee YS, Lee KH, Hur SY, Yoon JH, Lee SJ. Correlation between the digital cervicography and pathological diagnosis performed at private clinics in Korea. International journal of medical sciences. 2012;9(8):698.

13. Miller DD, Brown EW. “Artificial intelligence in medical practice: the question to the answer?”. The American journal of medicine. 2018 Feb 1;131(2):129-33.

14. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017 Feb 2;542:(7639):115-18. (http://www.ncbi.nlm.nih.gov/pubmed/28117445).

15. Abràmoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous ai-based diagnostic system for detection of diabetic retinopathy in primary care offices. NPJ Digital Medicine. 2018 2018/08/28;1:(1):39. (https://doi.org/10.1038/s41746-018-0040-6).

16. Gupta S, Sharma DK, Gupta MK. Artificial intelligence in diagnosis and management of ischemic stroke. Biomed J Sci Tech Res [serial on the Internet]. 2019; 13(3). (https://biomedres.us/fulltexts/BJSTR.MS.ID.002398.php).

17. Krishnaraj A, Barrett S, Bregman-Amitai O, Cohen-Sfady M, Bar A, Chettrit D, Orlovsky M, Elnekave E. Simulating dual-energy X-ray absorptiometry in CT using deep-learning segmentation cascade. Journal of the American College of Radiology. 2019 Oct 1;16(10):1473-9. (https://doi.org/10.1016/j.jacr.2019.02.033).

18. Ribli D, Horváth A, Unger Z, Pollner P, Csabai I. Detecting and classifying lesions in mammograms with deep learning. Scientific Reports. 2018 2018/03/15;8:(1):4165. (https://doi.org/10.1038/s41598-018-22437-z).

19. Hu L, Bell D, Antani S, Xue Z, Yu K, Horning MP, et al. An observational study of deep learning and automated evaluation of cervical images for cancer screening. JNCI: Journal of the National Cancer Institute [serial on the Internet]. 2019. (https://doi.org/10.1093/jnci/djy225).

20. Herrero R, Schiffman MH, Bratti C, Hildesheim A, Balmaceda I, Sherman ME, Greenberg M, Cárdenas F, Gómez V, Helgesen K, Morales J. Design and methods of a population-based natural history study of cervical neoplasia in a rural province of Costa Rica: the Guanacaste Project. Revista Panamericana de Salud Pública. 1997;1:362-75. (PMID: 9180057)

21. Schneider DL, Herrero R, Bratti C, Greenberg MD, Hildesheim A, Sherman ME, Morales J, Hutchinson ML, Sedlacek TV, Lorincz A, Mango L. Cervicography screening for cervical cancer among 8460 women in a high-risk population. American journal of obstetrics and gynecology. 1999 Feb 1;180(2):290-8. (https://doi.org/10.1016/S0002-9378(99)70202-4)

22. Xue Z, Novetsky AP, Einstein MH, Marcus JZ, Befano B, Guo P, Demarco M, Wentzensen N, Long LR, Schiffman M, Antani S. A demonstration of automated visual evaluation (AVE) of cervical images taken with a smartphone camera. International Journal of Cancer. 2020 Apr 30.

23. Automated Cervicography Using a Machine Learning Classifier, J.Kim, SR. Kim, SY Kwon, R. Nissim, J. Mink, D. Levitz, Poster presented at ACSSP 2019

24. Zall Y, Aljawad NA, Tsah D, Djaoui RY, Hillel S, Nissim R, Levitz D, "Live assessment of colposcopic image quality at the point of care using a machine learning classifier running on Android," 2020: Proc. SPIE 11230-12, Optics and Biophotonics in Low-Resource Settings VI, 1086911.

25. Cohen J. A coefficient of agreement for nominal scales. Educational and psychological measurement. 1960 Apr;20(1):37-46.

26. Ferris DG, Litaker M, ALTS Group. Interobserver agreement for colposcopy quality control using digitized colposcopic images during the ALTS trial. Journal of lower genital tract disease. 2005 Jan 1;9(1):29-35.

27. Bonnyman AM, Webber CE, Stratford PW, MacIntyre NJ. Intrarater reliability of dual-energy X-Ray absorptiometry–based measures of vertebral height in postmenopausal women. Journal of Clinical Densitometry. 2012 Oct 1;15(4):405-12.

28. Samuel M, Swaminathan J, Rajagopal S, Ramamurthy C, Jagan M, Sangamithirae D, Thalachinthala A, Kumari K, Ganta K, Hemalekha K, Solomon H, Kondati S, Sebag C, Peterson CW, Levitz D. Prevalence of Cervical Dysplasia and Cervicitis in South India Comparing Standard Cytology and Mobile Colposcopy. Pan Asian J Obs Gyn 2019; 2: 110-117.

29. Matti R, Gupta V, D'sa D, Sebag C, Peterson CW, Levitz D. Introduction of Mobile Colposcopy as a Primary Screening Tool for Different Socioeconomic Populations in Urban India. Pan Asian J Obs Gyn 2019, 2: 4-11.

30. Singer A, Khan A. Singer and Monaghan's Cervical and Lower Genital Tract Precancer: Diagnosis and Treatment. John Wiley & Sons; 2014 Jun 3.

31. Spaczynski M, Kotarski J, Nowak-Markwitz E. Management of abnormal PAP smear – consensus guidelines of the National Cervical Cancer Screening Programme in Poland Coordinating Centre, the Polish Gynaecologic Society, the Polish Society of Pathologists and the Polish Society of Colposcopy and Uterine Cervix Pathology. Ginekol Pol. 2009; 80(2): 129–133.

32. Chopra SJ, Mathew A, Maheshwari A, Bhatla N, Singh S, Rai B, Surappa ST, Ghosh J, Sharma D, Bhaumik J, Biswas M. National Cancer Grid of India consensus guidelines on the management of cervical cancer. Journal of global oncology. 2018 Jun;4:1-5.

33. Min KJ, Lee YJ, Suh M, Yoo CW, Lim MC, Choi J, Ki M, Kim YM, Kim JW, Kim JH, Park EW. The Korean guideline for cervical cancer screening. Journal of gynecologic oncology. 2015 Jul 1;26(3):232-9.

34. Armato III SG, Roberts RY, Kocherginsky M, Aberle DR, Kazerooni EA, MacMahon H, van Beek EJ, Yankelevitz D, McLennan G, McNitt-Gray MF, Meyer CR. Assessment of radiologist performance in the detection of lung nodules: dependence on the definition of "truth". Academic radiology. 2009 Jan 1;16(1):28-38.

35. Punglia RS, D'Amico AV, Catalona WJ, Roehl KA, Kuntz KM. Effect of verification bias on screening for prostate cancer by measurement of prostate-specific antigen. New England Journal of Medicine. 2003 Jul 24;349(4):335-42.

36. McHugh ML. Interrater reliability: the kappa statistic. Biochemia medica: Biochemia medica. 2012 Oct 15;22(3):276-82.

Original Article

To cite: Matti R, Gupta V, DK D'Sa, Sebag C, Peterson CW, Levitz D. Introduction of Mobile Colposcopy as a Primary Screening Tool for Different Socioeconomic Populations in Urban India. Pan Asian J Obs Gyn 2019;2(1):4-11.

Received on:
13/12/2018

Accepted on:
30/01/2019

Introduction of Mobile Colposcopy as a Primary Screening Tool for Different Socioeconomic Populations in Urban India

[1]Renuka Matti, [2]V Gupta, [3]D'Sa DK, [4]C Sebag, [5]CW Peterson, [6]D Levitz

[1]Consultant, [2]Center Head, [3]Manager, [4]Product Manager, [5]Vice President, [6]Chief Technology Officer
[1]Department of Obstetrics and Gynecology, Dr LH Hiranandani Hospital, Mumbai, India
[2,3]Operations and Marketing, International Oncology, Mumbai, India
[4-6]Mobile Optical Detection Technologies (ODT), Tel Aviv, Israel

ABSTRACT

Objective: To assess the feasibility of using a mobile colposcope as a screening tool for underserved urban populations in Mumbai.

Materials and methods: This study was conducted in hospital as well in health camps. First, visual screening was compared to standard of care cytology in a hospital-based setting. Thereafter, the technology was tested in field conditions of urban screening camps, where visual screening (Vis) using the mobile colposcope was used instead of visual inspection with acetic acid (VIA). In the hospital setting, total 321 women underwent routine cytological screening, followed by visual screening using a mobile colposcope. In the Camp, total 150 women were screened with the mobile colposcope. Clinical decisions and socioeconomic information were entered in the mobile colposcope smartphone app. The study duration was 8 months.

Results: Agreements between Vis and cytology was found in 157 cases. For dysplasia, there were 14 women who appeared Vis positive but Papanicolaou test (Pap) negative. Only one women who appeared Vis negative was positive for dysplasia. For cervicitis, 29 women were Vis positive and Papanicolaou test (Pap) negative but 117 women were Vis negative but were positive for inflammation on Papanicolaou test (Pap) testing. Approximately 60% of 23 patients called back for colposcopy and biopsy were lost to follow up. Use of the mobile colposcope in the screening camp allowed for improved workflow and documentation, and the experience was more positive than VIA for both patients and providers. In comparing socioeconomic level to pathology, cervicitis was common for both low and middle income patients, whereas dysplasia was almost entirely observed in low income patients.

Conclusion: Visual cervical cancer screening with a mobile colposcope is a feasible option to screen for cervical cancer in field as well as hospital settings. Additional research is needed to find a way to mitigate the frequency of loss to follow-up, which was significant in this study.

Keywords: Cervical cancer, cervicitis, cytology, cervical dysplasia, screening, low resource settings, colposcopy, digital health.

INTRODUCTION

India is home to one fourth of all deaths from cervical cancer worldwide.[1] Despite 432 million women being of eligible age for cervical cancer screening, only 3.1% of eligible women are screened per 3 years period in India.[2] Consequently, 67,000 women die from this preventable disease each year.[3] Much of this is attributed to loss-to-follow-up (LTFU); among women who undergo primary screening and were referred to a tertiary hospital for confirmation or treatment, up to 80% never returned for a follow-up examination or treatment.[4] To reach all women in India, a method that is low-cost, and high quality, is needed to reach diverse populations.

Methods for cervical cancer screening in India include visual inspection with acetic acid (VIA), cytology, and human papillomavirus (HPV) testing.

Address for Correspondence
Renuka Matti
Senior Gynecologist
Dr LH Hiranandani Hospital
Hillside Avenue, Powai, Mumbai, India
drrenumatti61@hotmail.com

VIA is the least expensive method, but its effectiveness is highly dependent upon provider skill level.[5] Cytology is the standard of care in India, but is expensive and less sensitive.[6] HPV testing is more sensitive than both methods, but still has associated costs and lab processing requirements.[7] Both cytology and HPV testing use significant laboratory and human resources, and require multiple patient visits, resulting in high LTFU rates.[4]

Given these challenges for implementing effective screening programs in India and other Low middle income countries (LMICs), an affordable, mobile cloud-connected colposcope was developed on a smartphone platform, allowing capture of colposcopy-quality images at the time of primary screening. The device - the Enhanced Visual Assessment (EVA) System (Mobile Optical Detection Technologies (ODT)) - is a fraction of the size of a traditional colposcope. To address limitations in visual screening in low-resource settings, software was built into the EVA mobile application for workflow management and remote quality assurance. The EVA System has been successfully deployed in several LMICs, including Kenya,[8,9] Haiti,[10] Mexico,[11] and Cambodia.[12]

In the present study, we assessed the feasibility of using a mobile colposcope as a screening tool for underserved urban populations in the Mumbai region.

MATERIALS AND METHODS

This feasibility study took place over an 8 months period from March-October, 2017. It consisted of two components: a head-to-head comparison between visual screening using a mobile colposcope against standard of care conventional cytology (Pap smears) in an urban hospital setting, and field testing in a screening camps in an urban slum in which visual screening with EVA was compared to VIA. Both components of the feasibility study were designed as cross sectional studies, utilized the same clinical staff (mainly gynecology experts), and had the same inclusion/exclusion criteria for patient recruitment. Specifically, these criteria limited the patient age to 18-65 years, and excluded patients who were pregnant, menstruating, or had a prior hysterectomy.

Prior to screening, all the women in study were educated and sensitized about both screening methods of the study (visual screening and conventional cytology, if applicable). All women were counseled on the next steps following an examination, and providers explained data usage and risk before an informed consent was signed in Hindi and/or English, depending on the request of the patient. Approval for both study components was obtained from the Clinical Trials Registry - India (CTRI registration number–CTRI/2017/03/013660), as well as from Institutional Ethics Committee Review board.

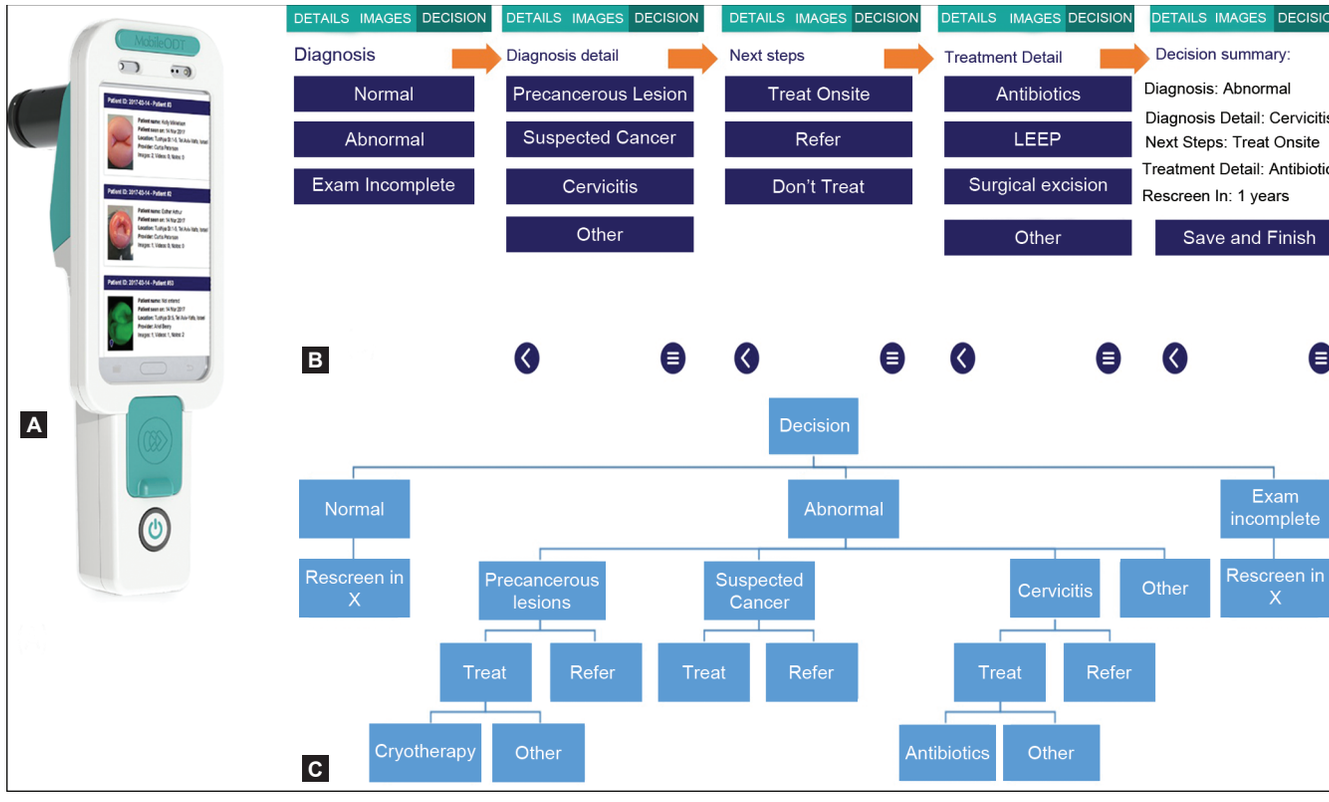
VISUAL SCREENING

All gynecologists who used the mobile colposcope (Fig. 1A) for screening attended a one-day training on use of the EVA System that included modules on the use of the software, job aid, and instructions of operation to capture clinically useful images of the cervix. There are three parts to using the app. First, basic patient information is collected (age, marital and socioeconomic status) directly into the EVA System mobile application that is used to operate the device. During screening, diluted 5% acetic acid was applied to the cervix, and the cervix was visualized. Both white light and green filter images of the cervix were captured for documentation. Expert gynecologists examined the cervix for signs of acetowhitening, inflammation, and other abnormalities. After the examination, the gynecologist recorded the impression from the visualization on the decision support job aid.

The decision support job aid is a unique feature of the EVA System (Fig. 1B). It is a workflow management engine integrated into the device, which was used to document visual impression at the time of screening according to the tree in Figure 1C. Providers recorded any abnormalities, including dysplasia and cervicitis. Patient details, images, annotations, and colposcopic impression by provider were automatically uploaded to the HIPAA-compliant image portal through an integrated SIM card.

HOSPITAL-BASED INVESTIGATION

Initially, the hospital-based setting recruited eligible patients from women who visited the clinic for routine screening over the study period. No additional outreach was conducted for hospital-based screenings. A total of N=321 patients were enrolled, and both cytology and visualization with EVA were provided according to the



Figs 1A to C: (A) EVA System, (B) Illustrated representation of Decision Support Job Aib on the EVA System App., (C) Full decision tree of job aid

study procedure. Visual screening using EVA was offered to the patient at no additional fee.

During screening procedures, the patients first underwent routine, conventional, scrape-based cytology and endocervical sample collection with dedicated brushes. Thereafter, visual screening with EVA was conducted, as described above. All women testing positive, either visually or by cytology, were asked to return for a follow-up colposcopy with biopsy that was offered to the patient free of charge. Patients with cervicitis diagnosed from the visualization were prescribed antibiotics and were asked to return for a follow-up screening in 1-3 months, after the infection would clear following antibiotic treatment.

All cytology and histopathology samples were sent to the hospital for routine processing and review. Pathologists documented sample adequacy presence of inflammation, and presence of abnormal cells.

FIELD TESTING

In addition to the hospital-based setting, EVA was tested under field conditions in urban screening camps. In contrast to the hospital study, patients here were recruited prior to screening camps by partner Non government organizations (NGOs) through community outreach and sensitization. Lead investigators set up eight mobile screening camps based in community centers, places of worship, and schools, where they could reach high-risk, low income women, including commercial sex workers, who face extremely high risk for cervical epithelial abnormalities.[13-15] Some supportive staff were junior clinicians who volunteered their time for the outreach activities within urban slums. Unlike in the hospital-based setting, here conventional cytology was not provided given the operational transport complexities. The comparison to VIA was on the operational level, in terms of ease of use and documentation.

Note that the hospital offered visual screening to its (low income) staff free of charge. However, these patients (N=38) were not offered cytology services, and as such, were grouped with the screening camp population.

DATA ANALYSIS

At the end of both study components, all cytology and histopathology results were collected and compared to visual impressions from the primary screening recorded in the EVA app. Decisions by the job aid were compared to cytology results and biopsy, as well as to age and socioeconomic status (low, middle, high, or prefer not to answer).

RESULTS

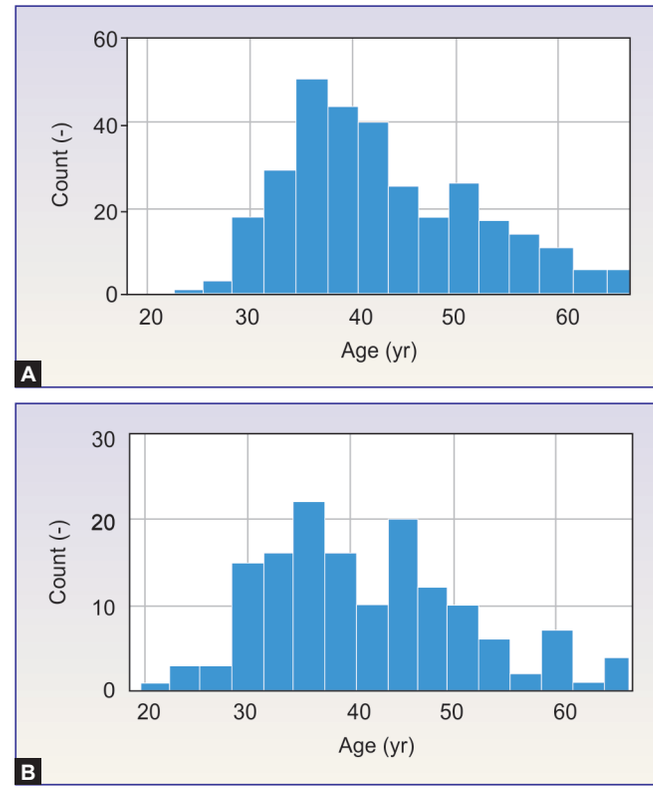
A total of 420 patients were recruited to enroll in the hospital-based setting, of which 99 were excluded due to improper cytological tracking or processing of the sample. The final patient number was N=321. In the screening camp, a total of 150 patients were enrolled. Figure 2 shows the age distribution of the hospital-based setting (Fig. 2A) and the screening camp (Fig. 2B).

Patients in the hospital-based setting were screened both visually and using standard of care cytology. The results comparing visualization to cytology (Table 1).

EVA and cytology agreed on 157 total cases. In terms of disagreement, most of the diagnoses involved misclassification of cervicitis. There were 117 inflammatory smears that were not visually classified as cervicitis, and there were 29 cases of visual inflammation where the cytology was non-inflammatory. In terms of dysplasia, there were 14 cases of visual screening/Pap- cases, but only one visual screening/Pap+ case.

To determine the accuracy of EVA and cytology, both methods were compared against the histopathology golden standard. Total 34 cervical biopsy samples were sent for histopathological examination 24 from the hospital-based setting and 10 from the screening camps. Altogether, 20 biopsies processed did not yield a complete histopathological classification, 15 from the hospital-based setting, and five from the screening camps. It is assumed these were LTFU, lower than the 80% Indian standard when a referral is made to a tertiary hospital for completion of diagnosis and treatment.[16] Within the entire data set, there were three positive biopsy-confirmed cases which were visual screening but lacked cytology assessment, one case caught by both methods, and one case missed by both methods.

Additional data is recorded in the EVA mobile application, both in the job aid that documents the clinical decisions taken at the point of care,



Figs 2A and B: Histogram of patient age in (A) Hospital-based setting and (B) Screening camp

Table 1

Comparison of EVA to cytology results

	Normal Pap	*Inflammatory*	*ASCUS, dysplasia, cancer*	*Pap unknown*
EVA normal	88	96	0	7
EVA cervicitis	24	64	0	5
EVA dysplasia	5	9	4	0
EVA other	5	12	1	1

and in the new patient screen that records basic patient information, including socioeconomic status. Figure 3 shows the socioeconomic makeup of the enrolled patients in both the hospital-based setting and screening camps, in terms of low- and middle-income (only one patient was high income).

At the screening camp, EVA was compared against standard naked eye visualization, though no differences in diagnosis between the two methods were noted. Patients reacted very positively to the use of the mobile colposcope.

In terms of the pathologies encountered in both components of this study, a breakdown of conditions is shown in Figure 4, for both low income and middle income patients from the hospital based-setting, and low income patients from the screening camp (only three middle income patients participated in the camp). It can be seen that disease prevalence was approximately the same between the groups (29-41%). However, dysplasia was more prevalent in the screening camp population, in comparison to patients from the hospital-based setting.

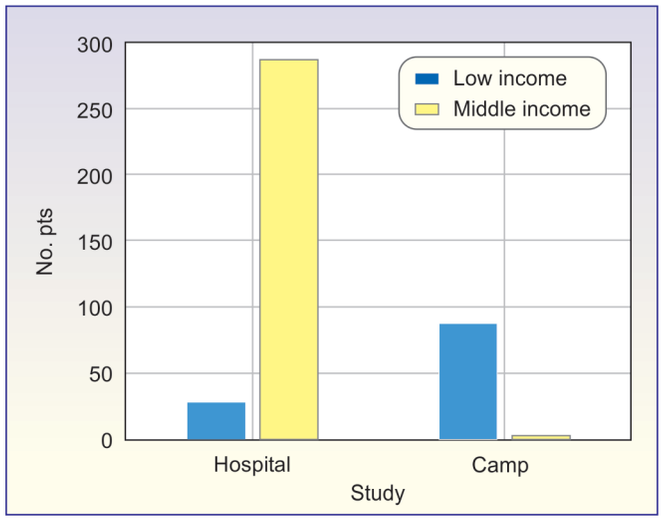


Fig. 3: Socioeconomic makeup of patient populations

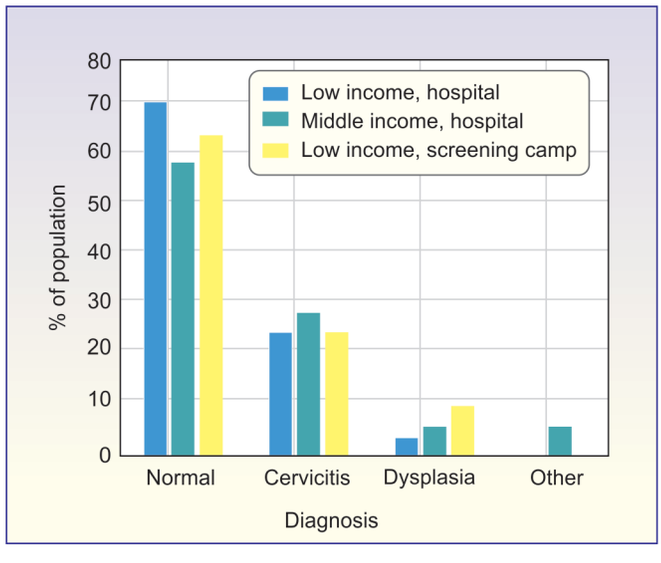


Fig. 4: Comparison of different gynecological pathologies observed with EVA, as a function of socioeconomic level

DISCUSSION

In our study, we assessed feasibility of visual screening using mobile colposcopy in two clinical settings in Mumbai, India. First, we compared visual screening to conventional cytology in a hospital-based setting, where we showed that more patients tested positive for cervical dysplasia visually than using cytology. Cervicitis was much more common than dysplasia (Table 1, Fig. 4). About two thirds of the patients called back for colposcopy and biopsy were LTFU. There were also three biopsy confirmed visual screening positive that lacked cytology results. In a screening camp setting, the EVA System allowed capture of important patient information that are difficult to capture in such settings, including age and socioeconomic status.

Our results from the hospital-based setting showed that rates of dysplasia were much lower than expected. Specifically, positive cytology rates were surprisingly low. A total of 5 patients tested positive for cytology (atypical squamous cells of unknown significant or ASCUS threshold), out of a total 321 patients, which is 1.5% of patients. In a country with such high mortality from cervical cancer,[1] these rates appear to be surprisingly low. However, other studies from the Mumbai region showed 3% Pap screening + rate.[16] In comparison, the total number of visual screening positive patients[18] represented 5.6% of the total patients, which is still low, but within the expected range. The difference between 14 visual screening +/Pap - cases and one visual screening negative/Pap+ case suggests that there are probably cases where cytology is missing. Similar results have been reported previously in India.[17] With only 3.1% of the eligible population actually getting screened, false-negative cytology poses a significant risk to the patient.

The high rates of cervicitis are particularly intriguing. About twice as many patients has in the hospital-based setting had inflammatory smears relative to visually detected cervicitis (56% versus 29%, Table 1). Such large disparities mean detecting dysplasia is much more

Table 2

A summary of the histopathology results for dysplasia compared against visual and cytology screening

Biopsy Analysis	*Biopsy Positive*	*Biopsy Negative*	*Biopsy Incomplete*
vis+/Papanicolaou test (Pap)+	1	2	
vis+/Papanicolaou test (Pap)–		2	11
vis+/No Papanicolaou test (Pap) (hospital)	1	1	
vis-/No Papanicolaou test (Pap) (hospital)		1	1
vis-/Papanicolaou test (Pap)–	1		3
vis+/No Papanicolaou test (Pap) (camp)	2	3	5

vis: visual screening

difficult, since cervicitis can cover up dysplasia and patients need to be rescreened following antibiotics treatment. Illustrating this is the one visual screening–/Pap-, biopsy-confirmed dysplasia case, where the user wrote on the EVA app that both cervicitis and acetowhitening were present, and they were not sure which to mark.

Another striking finding in this study was the histopathology analysis (Table 2) where 20 of 34 patients called back for biopsy were LTFU, and importantly, 15 of those 24 cases were in the hospital-based setting. These rates of LTFU are significant, because the clinical staff reached out multiple times to the relevant patients by phone, email, and SMS, to let them know that biopsy services were offered to them at no cost, yet they were still unable to get them back to the clinic for confirmatory biopsy. These patients are predominantly low income patients (Fig. 4), which suggests that economic factors could have affected the LTFU rates. These high LTFU rates highlight a challenge to cervical cancer care in India, and should be looked into further, given the fact that our study compares favorably with other studies in India which had LTFU rates as high as 80%.[4] Information recorded on the EVA app (including phone number and email) allowed clinicians to persistently call patients back to help them return for colposcopy with biopsy.

The screening camps represents another clinical scenario in which visual screening using EVA was piloted. Such camps serve a different patient population (Fig. 3), and represent different field conditions relative to a stationary hospital clinic. Cytology and biopsy are not as readily available, given the resources and infrastructure they require. Because of the camps' physical condition and overall patient volume, proper record keeping is quite challenging. Anecdotally, it was much was easier to use EVA for documentation of patient information than existing methods (hand-written records or information typed on a laptop). And the information stored by the app allowed for much more rapid data analysis following the deployment.

On a qualitative level, feedback from providers showed they felt the device, in comparison to naked-eye visualization, reduced the time of the exam because less "looking" had to be done. Moreover, the digitized data capture improved documentation of the results at the community level, increased trust among patients due to the ability to see the images from the examination and increased the patient's sense of empowerment and ownership over her body.

The socioeconomic makeup of the patients enrolled in the study (Fig. 3) was similar to expected levels – the hospital clinic had a large majority (89.4%) of middle income women, some (8.8%) low income women. In contrast, the screening camp population was predominantly lower income. Our socioeconomic data shows two key findings: first, that dysplasia was much more common in low income patients than middle income patients, and second, that cervicitis rates are much higher than dysplasia rates. Both of these results are in general agreement with the previous reports in the literature.[18,19] Moreover, the numbers from a similarly designed multi-center trial in six sites in southern India[19] showed very similar results, with recorded cervicitis rates being much higher than dysplasia.

While formal cost effectiveness data was not collected in this study, given the high cost of consumables and laboratory testing costs associated with cytology, it is assumed that mobile colposcopy is more cost effective given the device's 10 years lifespan and ability to screen up to 50 patients per day with the same consumable as VIA, the lowest cost method of screening today. However, like cytology, the examination does require a higher level of expertise to not only collect specimens but also distinguish the difference between cervical infections and cervical dysplasia.

There were several limitations to the current study. First, the screening technologies tested in the hospital-based setting were limited to cytology and EVA; an improved comparison would have included HPV testing as well. HPV testing, with its high negative predictive value, would allow better assessment of false negatives, which was not possible under the current protocol. The screening camp effort should separate the provider using EVA from the provider performing VIA, to better document the differences between the two methods. And finally, a behavioral health economist should be consulted on how to incentivize patients to return for follow-up check-ups, although this a challenging task all on its own.

One possible way to reduce LTFU rates would be to utilize a single visit approach method, where a follow-up colposcopy with biopsy can be performed immediately after screening. Further investigation should be done using this approach to better assess clinical accuracy of the EVA System, utilizing new real-time consultation features, to conduct confirmatory biopsy at the primary screening. No technology available today has both high negative predictive value and fast response time needed for a single-visit approach. Two technologies under development that could potentially enable implementing this are point of care HPV testing and automated visual evaluation (AVE).[20] Until then, there will be a compromise between accuracy and time/cost.

CONCLUSION

In this study we assessed feasibility of visual cervical cancer screening using a mobile colposcope in urban clinical settings in India. In the hospital based settings, visual screening detected less cases of cervicitis than conventional cytology. However, more cases of dysplasia were identified at the primary screening from visualization than conventional cytology. Histopathology results revealed very high LTFU rates. Under field conditions in a screening camp, visual screening using EVA allowed for capturing more pertinent information about the (underserved) patient population, and overall there was a more positive experience for both the provider and the patient.

ABBREVIATIONS

Automated visual evaluation	AVE
Clinical Trials Registry - India	CTRI
Enhanced Visual Assessment System	EVA
Human papillomavirus	HPV
Loss-to-follow-up	LTFU
Low and middle income countries	LMICs
Optical Detection Technologies	ODT
Papanicolaou test	Pap
Visual inspection with acetic acid	VIA
Visual screening	Vis

DISCLOSURE STATEMENT

C Sebag and D Levitz are employees of Mobile ODT and own stock in the company. Additionally D Levitz sits on Mobile ODT Board of Directors.

ROLE OF THE FUNDING SOURCE

This research was financially supported by the World Bank/IFC TechEmerge Program, IFC Project No. 600972. However, the funders did not have a role in study design; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the paper for publication. All authors of this paper confirm final responsibility for the decision to submit for publication.

ACKNOWLEDGMENTS

The authors wish to thank nurses and supporting staff at both Hiranandani Hospital and through screening camp led by NGOs. We would like to thank all NGOs involved in all studies, whose names are not mentioned to protect the communities they serve. This study was funded by the International Finance Corporation and the World Bank through the Tech Emerge program (Award #600972).

REFERENCES

1. Ferlay J, Soerjomataram I, Ervik M. GLOBOCAN 2012, Cancer Incidence and Mortality Worldwide in 2012. Lyon, France: International Agency for Research on Cancer; 2012. Available from: http://www.globocan.iarc.fr.
2. Bruni L, Barrionuevo-Rosas L, Albero G. ICO/IARC Information Centre on HPV and Cancer (HPV Information Centre). Human Papillomavirus and Related Diseases in India. Summary Report 27 July 2017.
3. Sankaranarayanan R, Rajkumar R, Theresa R. Initial results from a randomized trial of cervical visual screening in rural south India. Int J Cancer 2004;109: 461-467.
4. Mittal S, Biswas J, Basu P. Visual inspection with acetic acid as a screening test for cervical cancer. Biomed Res J 2014;1(1):23-33.

5. Denny L, Sankaranarayanan R. Secondary prevention of cervical cancer. Int J Gynaecol Obstet 2006;94 (Suppl 1): S65–S70.
6. Basu P, Meheus F, Chami Y, Hariprasad R, Zhao F, Sankaranarayanan R. Management algorithms for cervical cancer screening and precancer treatment for resource-limited settings. Int J Gynaecol Obstet 2017;138: 26-32.
7. Peterson CW, Rose D, Mink J, Levitz D. Real-time monitoring and evaluation of a visual-based cervical cancer screening program using a decision support job aid. Diagnostics 2016; 6(2).
8. Mink J. Statistics on Clinical Activities and Practices of Cervical Cancer Clinics in Kenya, and beyond. Presented at the American Society for Colposcopy and Cervical Pathology annual meeting, 2017. www.asccp.org/.../x-8-mink-jonah-statistics-on-clinical-activities-in-kenya-pptx-pd
9. Millien C, Jean-Baptiste MC, Manite G, Levitz D. Remote quality assurance in cervical cancer screening in low resource settings using a handheld smartphone-based colposcope. Proc. SPIE 9314, Optics and Biophotonics in Low-Resource Settings, 93140A; March 2015. *https://mobileodt.com/wp-content/uploads/2018/02/SPIE-2015-Remote-quality-assurance-in-cervical-cancer-screening-in-low-resource-settings-using-a-handheld-smartphone-based-colposcope.pdf*
10. Madiedo M, Contreras S, Villalobos O, Kahn BS, Safir A, Levitz D. Mobile colposcopy in urban and underserved suburban areas in Baja California. Proc. SPIE 9699, Optics and Biophotonics in Low Resource Settings II, 96990I; March 2016.
11. Thay S. Determining the Optimal Cervical Carcinoma Screening Method in HIV Positive and HIV Negative Cambodian Women. Presented at the American Society for Colposcopy and Cervical Pathology annual meeting, 2018.
12. Mak R, Van Renterghem L, Cuvelier C. Cervical smears and human papillomavirus typing in sex workers. Sex Transm Infect 2004;80:118-120.
13. Tideman RL, Thompson C, Rose B, Gilmour S, Marks C, van Beek I, et al. Cervical human papillomavirus infections in commercial sex workers-risk factors and behaviours. Int J STD AIDS 2003;14:840-847.
14. Wong WC, Wun YT, Chan KW, Liu Y. Silent killer of the night: A feasibility study of an outreach well-women clinic for cervical cancer screening in female sex workers in Hong Kong. Int J Gynecol Cancer 2008;18:110-115.
15. Sankaranarayanan R, Thara R, Sharma A, Roy C, Shastri S, Mahé C, et al. Accuracy of conventional cytology: results from a multicentre screening study in India. Med Screen2004;11:77-84.
16. Shastri SS, Dinshaw K, Amin G, Goswami S, Patil S, Chinoy R, et al. Concurrent evaluation of visual, cytological and HPV testing as screening methods for the early detection of cervical neoplasia in Mumbai, India. Bull World Health Organ 2005; 83:186-194.
17. Thobbi VA, Khan F. Cervical cytology by pap smear in reproductive population. Int J Reprod Contracept Obstet Gynecol 2018;7:1988-1992.
18. Sreedevi A, Javed R, Dinesh A. Epidemiology of cervical cancer with special focus on India. Int J Women's Health 2015:7 405-414.
19. Samuel J, Swaminathan C, Sebag S, Rajagopal C, Ramamurthy M, Jagan, et al. Rates of cervicitis across multiple sites in India", Presented at the American Society for Colposcopy and Cervical Pathology annual conference (2018), 000436.
20. Demarco M, Hu L, Antani S, Yu K, Gage J, Befano B, et al. Automated visual evaluation for cervical cancer screening and management: promise and limitations", Presented at International Papillomavirus Conference (2018), IPVC8-0278.

CISION
PR Newswire

MobileODT Technology Improves Cervical Cancer Detection Rates - Study

Cervical cancer contributes to approximately 6-29% of all cancers in Indian women

India accounts for one-quarter of cervical cancer globally

17% of all cancer deaths amongst women in India aged 30-69 are due to cervical cancer

EVA System visualisation technology offers potential for early detection of cervical cancers

NEWS PROVIDED BY
MobileODT →
Jan 17, 2018, 09:01 ET

CHENNAI, India, Jan. 17, 2017 /PRNewswire/ -- As per a new World Bank/IFC-funded study by MobileODT and Apollo Hospitals, deploying the EVA (Enhanced Visual Assessment) System at primary cervical cancer screenings leads to a higher detection rate of suspected precancerous and cancerous lesions for women in both urban and rural areas, than a pap smear alone.

The study assessed MobileODT's EVA System for high-quality and cost-effective cervical cancer screenings across India. 567 women were screened for cervical cancer via colposcopy with the EVA System, as well as a Pap smear at the first screening. Expert gynecologists reviewed cases through the EVA System's software to confirm suspicion of precancerous and/or cancerous lesions, and other abnormalities for potential treatment.

The gynecologists shared high scores of feasibility, usability, and patient satisfaction with the EVA System, which proved to be a quick and easy way to evaluate the cervix. All investigators felt it can be extended to peripheral centers with training where Pap smear collection and interpretation is not easily available.

MobileODT's CEO Ariel Beery said: "We're excited to see these results, and are grateful for the opportunity to work with Apollo to extend our life-saving technology to the women of India. Few healthcare institutions in the world have the expertise and reach of Apollo, and we're looking forward to working with Apollo to save more lives in the future."

Dr. Mallika Samuel, Senior Consultant Gynecologist, Apollo Hospitals, Chennai and Lead Investigator of the study said, "The Indian healthcare landscape presents significant challenges to patient access and affordability. The deployment of the EVA System can facilitate taking cervical screening to the people of India, enabling early screening and treatment thus slowing down incidences of cervical cancer. Further, it maximizes cervical cancer screening programs and reduces the time to follow up with patients, especially in rural and remote areas."

See our related news post: http://www.mobileodt.com/news/eva-system-deployment-leads-to-higher-cervical-cancer-detection-rates-in-india-study

About MobileODT

MobileODT is creating the next generation of smart medical solutions. Our EVA (Enhanced Visualisation and Assessment) System combines biomedical optics with the power and connectivity of cellular technology. In 2017, MobileODT piloted the EVA System for Cervical Cancer with Apollo Hospitals and is attending the India-Israel Global Innovation Challenge Initiative in Mumbai, January 15-19, 2018.

For further comment, please contact:
Dharshana Kankaria
dharshanakankaria@mobileodt.com

About Apollo Hospitals

Apollo Hospitals is also the world's largest private cancer care provider and runs the world's leading solid organ transplant program. It is acclaimed for pioneering the private healthcare revolution in the country. Each year, Apollo Hospitals reaches a quarter billion patients in India, providing lifesaving care.

For further comment, please contact:
Kusum Sahijpal
Kusum_s@apollohospitals.com

SOURCE MobileODT

Related Links

http://www.mobileodt.com