# Performance of automated visual evaluation as a triage test for HPV+ patients from a screening camp in rural China

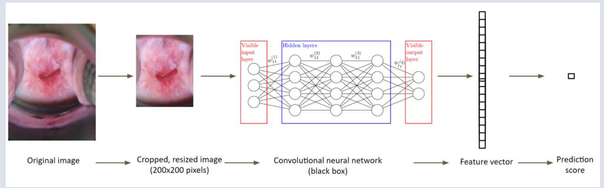
AT Goldstein[1], S Bedell[1], R Lipson[2], CM Sebag[3], L Lobel[3], D Levitz[3]

*1 Center for Vulvovaginal Disorders, Washington, DC, USA; 2 United Family Healthcare Hospitals, Beijing, China; 3 MobileODT Ltd., Tel Aviv, Israel*

TAP TO RETURN TO KIOSK MENU

## Introduction

Automated visual evaluation (AVE) is a promising technology that uses a machine learning (ML) classifier to predict the likelihood of pathology in a cervical image. AVE is accurate, fast, and inexpensive, and thus it has tremendous potential for utilization in low resource settings (LRS), where cytology, and at times HPV testing, are not readily available.


In order to be deployed in LRS, AVE needs to be integrated into an imaging device with computing power that can operate in such a setting. One device that meets this criteria is the Enhanced Visual Assessment (EVA) System (MobileODT), a cloud-connected mobile colposcope used in >40 countries globally.

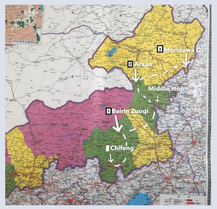
The **goal** of this study is to determine whether a preliminary version of AVE can be used at the point of care (PoC) in LRS, and how it integrates into the clinical workflow. Here, AVE was used to assess patients in a screening camp in Inner Mongolia, China.

## Methods

An existing AVE classifier was integrated into the EVA System as web service. The classifier was deployed on the cloud and is called by the provider from the patient records on the EVA image portal.

The EVA System was used for imaging HPV+ patients, as part of the care provided in a screening camp in Inner Mongolia, China. Suspect regions on the cervix were biopsied and analyzed in United Family Healthcare Hospitals in Beijing. AVE scores were compared to histopathology results. Positives were defined by CIN 2+ histopathology.

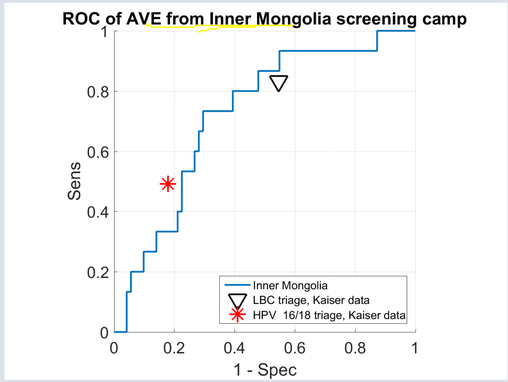
A secondary analysis also assessed test reliability at the PoC. Test inadequacy rates were measured against colposcopic impression.




Source: Wikipedia

## Results

A Receiver Operating Characteristic curve was calculated on AVE scores, showing AVE to be more sensitive than specific on HPV+ patients from the screening camp.




Colposcopic impression vs. AVE

| | AVE+ | AVE- |
|---|---|---|
| Colpo + | 112 | 11 |
| Colpo- | 5 | 0 |

### Test inadequacy rates

- Colpo impression: 17/147 = 11.6%

Defined as “Unable to make colposcopic impression”

- AVE: 2/147 = 1.3%

Defined as “patient case did not sync to portal in reasonable time to provide the patient an answer”

## Discussion

- ***First use of AI for cervical cancer detection at PoC in LRS***
- Low failure rate for AVE, in comparison to colposcopic impression
- 3G connectivity sufficient for running classifier on EVA portal
  - Further testing needed for areas with lower connectivity
- Classifier performance comparable to those of triage technologies reported in the literature from high resource clinic.
- AVE performance was poorer on these images than in global validation set. These differences are explained by model generality, observed between External test set vs. Holdout test set

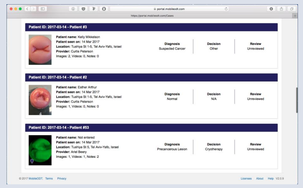
# Performance of automated visual evaluation as a triage test for HPV+ patients from a screening camp in rural China

AT Goldstein[1], S Bedell[1], R Lipson[2], CM Sebag[3], L Lobel[3], D Levitz[3]

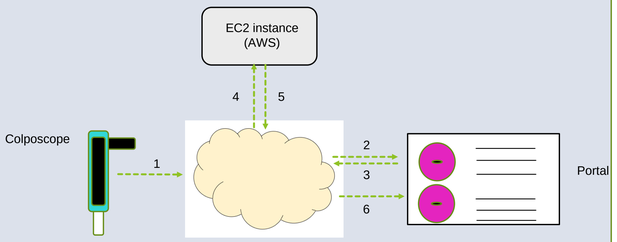
*1 Center for Vulvovaginal Disorders, Washington, DC, USA; 2 United Family Healthcare Hospitals, Beijing, China; 3 MobileODT Ltd., Tel Aviv, Israel*

## Running AVE at the point of Care

The EVA image portal – a secure location for saving patient data on the cloud – is an integral part of the EVA System.



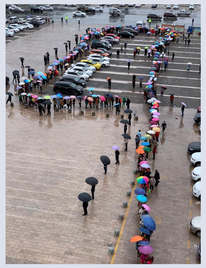
The EVA image portal provided the secure connection to the AVE classifier on the cloud. It allows for sending images to a remote server for analysis, and returns an answer.

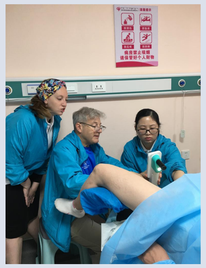



- Integrated button on portal
- Initiates EC2 instance on Amazon Web Services' servers
  - Input: single image
  - Output: prediction score
- Total time: ~5-10 min
- Feature piloted in Korea for 1 year
  - Almost 7000 patients

## Screening camp information

522-992 women/day self-swabbed for high-risk HPV (hrHPV+). Two Ampfire HPV PCR systems were run simultaneously to test the specimens. All hrHPV+ patients had digital colposcopy (DC) performed on the same day with the EVA system. Digital images were obtained, and all suspected lesions were biopsied. Suspected CIN1 and CIN2 lesions were treated with thermo-coagulation, and suspected CIN3 lesions were treated with LEEP.



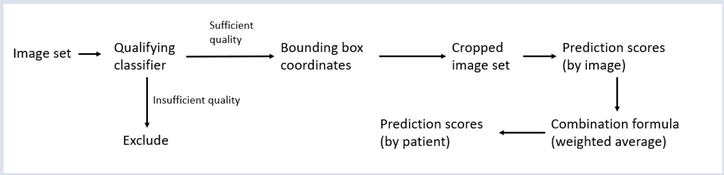


3345 women were screened/treated by 5 physicians in 5 days. 619 women (18.5%) were hrHPV+. 589 women underwent same-day colposcopy. Biopsy: 108 CIN1, 22 CIN2, 9 CIN3.

## AVE classifier

- Trained on manually annotated images from 1473 patients
- Data came from 17 countries, with heavy representation from Kenya and India.
- Based on Faster RCNN architecture running on Caffe

## AVE data analysis




During data analysis, a secondary ML classifier (the qualifying classifier) was used to filter out bad images prior to further analysis. The qualifying classifier also cropped the cervix region in the images, to ensure the prediction is not affected by artifacts.

There were multiple images captured per patient, each with varying quality. To combine the AVE prediction scores from multiple images in one patient, a weighted average of the scores was calculated, using the image quality scores as the weights.

## CONCLUSIONS

The tested implementation of AVE – uploading images to the patient file and running the classifier immediately – proved feasible for use in low resource settings.

The inadequacy rate of AVE – cases that did not sync – was very low (1.3%), and smaller than the rate at which colposcopic impression was not feasible.

In comparison to data from other studies on triage of HPV+ patients, the performance of AVE compared favorably to that of existing and emerging triage methods including cytology, dual p16/Ki67 stain, and DNA methylation.

## NEXT STEPS

External confirmation of biopsy + Adjudication on discordant cases (in progress)

Reassess accuracy and ROC given "new ground truth" labels

## References

Hu et al, J NCI 2019

Demarco et al, IPVC 2018

Wentzensen et al, J AMA Int Med, 2019

## Funding

This work was funded by Gynecologic Cancers Research Foundation.