

Automated Cervicography Using a Machine Learning Classifier

JY Kim¹, SR Kim², SW Song³, SY Kwon⁴, J Kim⁵, Ronen Nissim⁶, Jonah Mink⁶, David Levitz⁶

(1) Y-QUEEN WOMAN CLINIC
(2) Dr. Kim ob/gyn clinic, Seoul, Korea
(3) Roen ob/gyn clinic, Seoul, Korea

(4) Riz ob/gyn clinic, Seoul, Korea
(5) Jiin ob/gyn clinic, Seoul, Korea
(6) MobileODT, Tel Aviv, Israel

Abstract

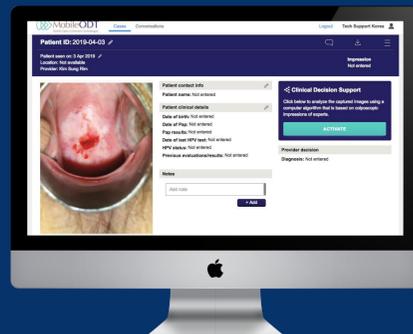
Objective: Demonstrate effectiveness of the first use of a prospective, real-time machine learning (ML) algorithm in a clinical setting. **Methods:** An ML classifier was developed from an existing image set from 1473 colposcopy patients (80% training, 20% validation). Annotations by two colposcopy experts were used as ground truth. The classifier was then integrated into a web service feature called from an image portal storing patient images and test results. The feature evaluates all images from the selected procedure, and provides both an automated impression and targeted feedback. This feature was piloted in a network of seven clinics in Korea, where combined cervicography and cytology are the screening standard of care. The results of the classifier were used to counsel patients on risk in order to improve loss to follow-up for high risk cases. **Results:** The ML classifier developed had an area under the (ROC) curve (AUC) of 0.93. The Korea pilot is the first ML algorithm on cervical images tested in a clinical setting. To date, 343 patients were enrolled, with provider utilization at 100%. Data from N=209 patients are included in this study, and laboratory results from N=134 patients are still pending. **Conclusion:** Preliminary results show widespread acceptance of AI at the point of care, and highlight potential to improve care and reduce costs related to cervical cancer screening.

Conclusion

- Utilization of CDS was very high, suggesting a need for such an automated QA tool
- CDS appeared to be more suitable for screening than digital cervicography
- Decision support through manual annotations does not always yield ground truth answers

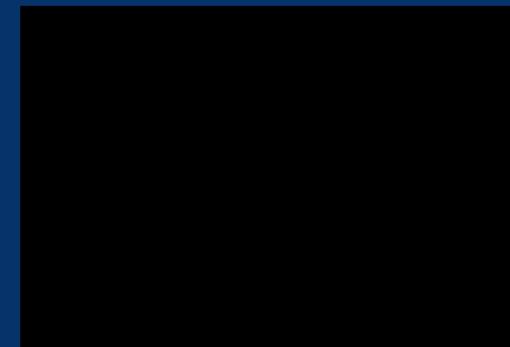
Material & Methods

- A clinical decision support (CDS) Classifier trained based on colposcopist annotations of images with an AUC of 93%
- Classifier deployed as an offline QA tool, following screening by cytology and digital cervicography
- Technology piloted in 7 clinics across Korea



Results

- CDS had similar positivity rates to cytology and cervicography (18-20%)
- CDS and cytology yielded inadequate / inconclusive results in 1% of patients
- Cervicography yielded inconclusive results in 47% of patients
- Biopsy results on 17 patients showed discrepancy between histopathology and training annotations



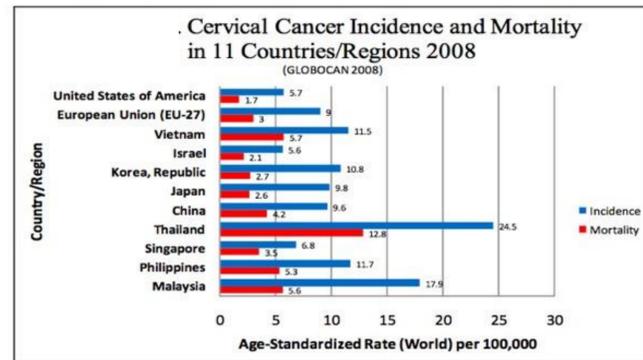
Click to play video



Introduction

Objective: assess the utility of a ML classifier to perform quality assurance on colposcopy images and annotations

Cervical cancer remains a leading cause of death for women worldwide.¹ While there are well established methods of screening for the disease for all resource settings, including HPV testing, cytology, VIA and cervicography, loss to follow-up remains a critical challenge for women to return for secondary screening and colposcopy following an abnormal screening result globally.²⁻⁴ As a single method to detect cervical cancer in the single visit has been established as high rate of effectiveness to limit loss to follow-up, challenges remain in track women across the screening cycle.



However, interpretation of medical images by “medical professionals” is highly subjective, with disagreements between experts occurring in approximately 1 in 3 patients. This includes cervical tissue imaged at the time of colposcopy. Mechanisms instituting quality assurance (QA) need to be put in place, to improve provider training and provide them with decision support in their clinical decision making.

Providers not only disagree with one another on colposcopy images, but also on digital cervicography images and visual inspection with acetic acid (VIA). QA is certainly necessary. While QA often works in smaller organizations, in larger organizations it is not feasible because one doctor cannot review so many images by junior providers.

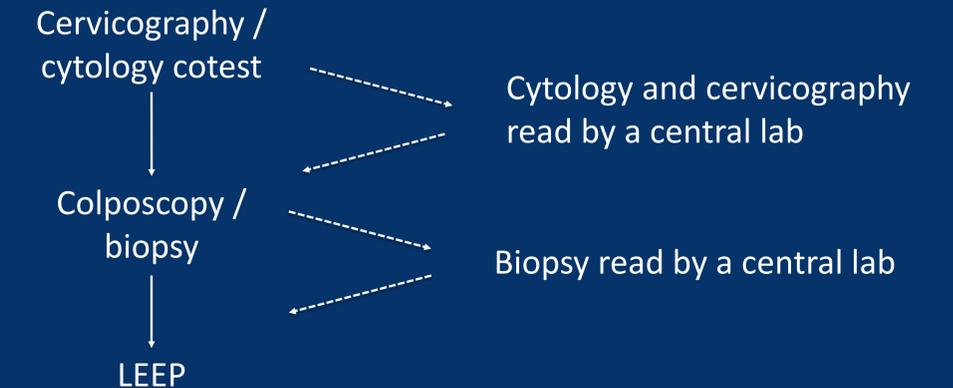
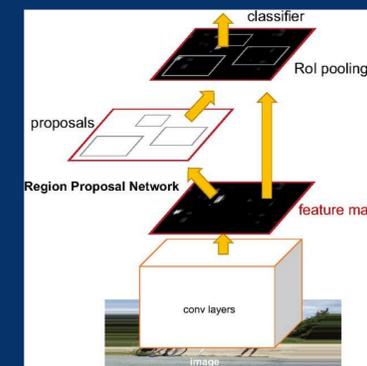
Cervical cancer management in Korea

Digital cervicography is part of standard of care!

- Images are read by cervicography experts in a remote central lab
- Reports take 2-3 days



Standard digital cervicography system in South Korea.



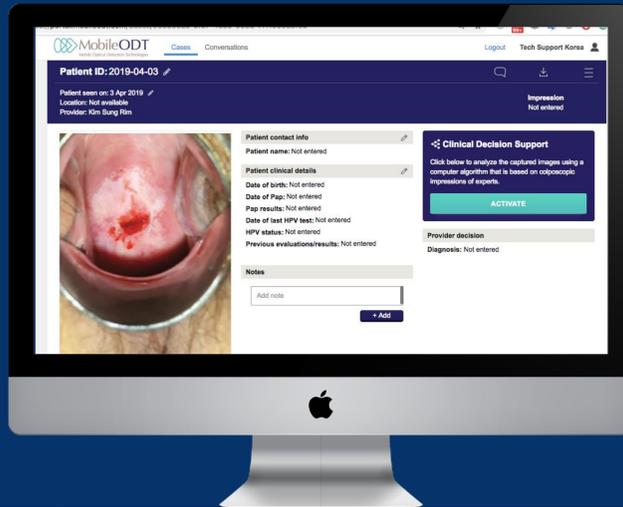
Further complicating matters is that there is a lack of proficient colposcopic experts in Korea,⁵⁻⁷ and providers trained in interpreting cervical images to understand risk of high-grade disease to counsel patients on the risk and need for follow-up. One solution to this challenge is to automate the QA process use an analysis tool developed using machine learning. A classifier can be trained to perform like expert colposcopist. Such a classifier can provide assistance to those providers who want to improve their practice, without the need to be working closely with an expert.

To address this gap, we developed a **clinical decision support (CDS)** classifier provide an automated second opinion on cervical images captured with the EVA System, a cloud-connected mobile colposcope. The classifier operates in offline mode in order to not affect the standard of care. Providers are able to capture images and assess themselves in near real time.

Neural network architecture used in CDS classifier.⁸

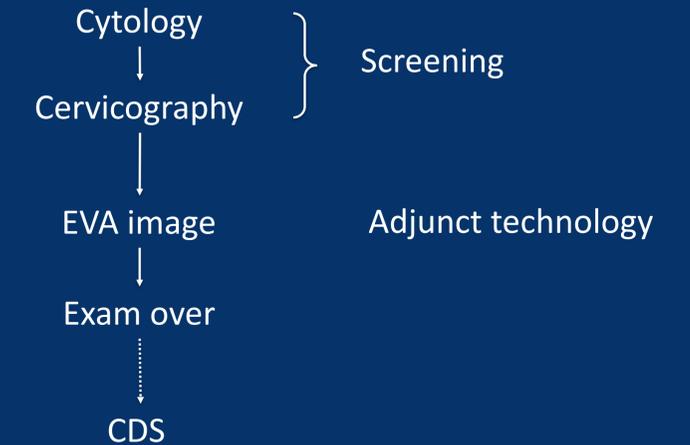
Materials and Methods

CDS on EVA Portal



Clinical Network

- **This is the first pilot to prospectively test machine learning algorithms in cervical cancer management**
- The pilot was conducted across 7 private sites in South Korea
- One or two providers captured images with the EVA System at each center
- Patient management was done based on standard of care
- After each exam, the provider opened the EVA online portal and activated CDS on captured images
- If a CIN 2+ result found, provider counseled patient on risk and importance of follow-up



Clinical decision support algorithm performance



All images in database were captured with the EVA System

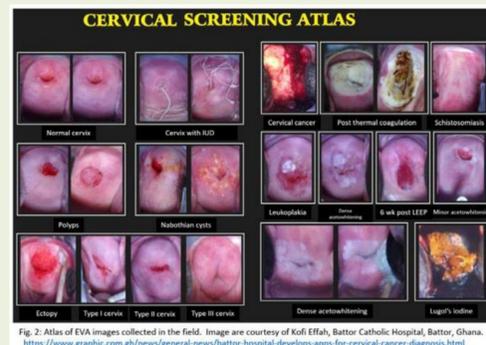


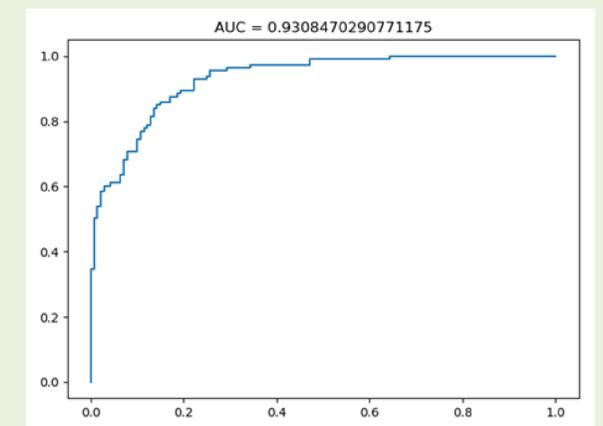
Fig. 2: Atlas of EVA images collected in the field. Image are courtesy of Kofi Effah, Bator Catholic Hospital, Bator, Ghana. <https://www.graphic.com.gh/news/general-news/bator-hospital-develops-apps-for-cervical-cancer-diagnosis.html>

- The EVA System has been used in >33 states in the US, and ~40 countries worldwide.
- >50,000 patients imaged with EVA
- Images are de-identified and stored on MobileODT portal

Data for classifier came from existing sources⁹

Expert impression	Total	Training set	Test set
Probable High-grade	665	532	133
Possible High-grade	848	0	848
Minor Abnormality	809	647	162
Normal			

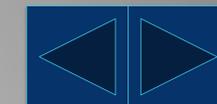
- High quality images from reputable clinics used
- Images reviewed for technical adequacy
- Adequate images reviewed by Gyn Oncologists at Rutgers (Mark Einstein, Akiva Novetzky, Jenna Marcus)



Classifier built around patient case, NOT images

Classifier score from images in the same session were combined by a weighted average., based on an image quality score.

Results



Cytology vs. CDS	CDS			TOTAL
	Positive (CIN 2+)	Negative (CIN 1-)	Insufficient for processing	
Cytology				
Normal	34	134	2	170
ASC-US	3	23	0	26
LSIL	5	6	0	11
HSIL	0	0	0	0
Unknown	0	2	0	2
TOTAL	42	165	2	209

Cervicography vs. CDS	CDS			TOTAL
	Positive (CIN 2+)	Negative (CIN 1-)	Insufficient for processing	
Cervicography				
Positive (CIN 1+)	1	21	0	22
Atypical	8	54	1	63
Negative	28	59	1	88
Unknown	5	31	0	36
TOTAL	42	165	2	209

Histopathology vs. CDS	CDS		TOTAL
	Positive (CIN 2+)	Negative (CIN 1-)	
Histopathology			
Normal	0	0	0
Cervicitis	0	0	0
CIN 1	3	13	16
CIN 2	0	0	0
CIN 3	0	1	1
TOTAL	3	14	17

CDS vs. cytology and cervicography

The CDS algorithm is built as a software tool to give providers rapid QA on their clinical decision making. Utilization of the algorithm was at 100%, meaning that clinicians are interested in receiving rapid feedback. As such, CDS could be a unique teaching tool for lesser skilled providers.

In comparison of CDS to cytology using an ASCUS threshold, CDS and cytology yielded similar screen positive rates: 37 cyt+ (17.9%) vs. 42 CDS+ (20.3%). Both technologies had an inadequacy rate <1%.

Cervicography, in comparison to CDS and cytology, had a similar screen-positive rate (20.0%). Interestingly, cervicography had a very high rate of inadequate reading / ambiguous results, with 99 of 209 images (47.4%) yielding an inconclusive result.

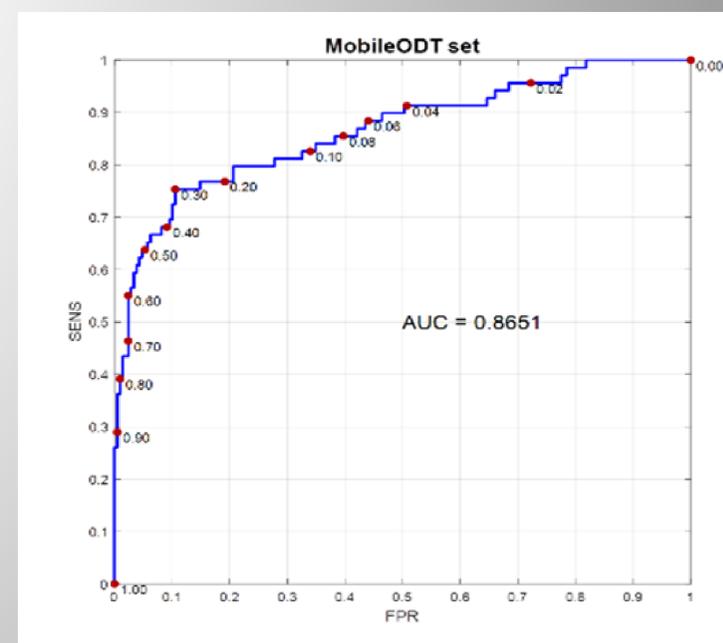
CDS vs. Histopathology

For the patients who went on to colposcopy with biopsy, a comparison was made between CDS and the biopsy result, using CIN 2+ as a threshold for positivity. There were N=17 cases in all, which is too small to draw meaningful conclusions.

In comparison to ground truth, CDS was accurate in 83% of the time. Any classifier is as good as the training data, which for CDS was manual annotations by colposcopy experts. When there are discrepancies between the training data and the ground truth (worst histopathology), the classifier performance will “degrade”, because it performs like the training data rather than ground truth. Indeed the CDS classifier was highly accurate in comparison to expert annotations (AUC = 93%).

To improve performance, the classifier needs to be trained on histopathology-correlated images, not expert annotations. Above we show preliminary results of a histopathology-correlated classifier, with an AUC of 86%.

Histopathology as ground truth



Case studies

TAP TO RETURN TO KIOSK MENU



Example Case (2)

Pap Test	EVA 3 Plus		Other Device		HPV Test	Cervical Punch Biopsy
Result	CDS result	Morphology	Result	Morphology	Result	Result
Normal	Negative	Acetowhite epithelium	Positive	Acetowhite epithelium	HPV 16, 70	CIN1
ASC-US	Negative		Positive	Erosion or Unclear	x	CIN1



On the other hand, these cases show when
The result of Cervical Punch Biopsy shows **CIN 1**,

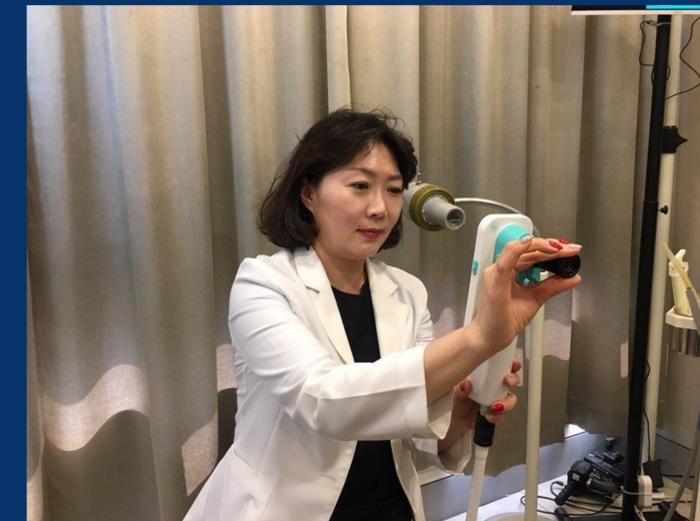
1. EVA System CDS detects **NO High Grade Lesion**
2. Other Cervicography Device says **positive**

These prove that CDS of EVA System makes
Less mistakes detecting high grade lesion.

LaonZ
Long-lasting Good Company

Copyright © 2019 LAONZ Inc Co., Ltd. All rights reserved. | Confidential

CDS Demo Video



References

1. de Martel C, Plummer M, Vignat J, Franceschi S. Worldwide burden of cancer attributable to HPV by site, country and HPV type. *Int J Cancer* . 2017;141(4):664–670.
2. Nene BM, Deshpande S, Jayant K, Budukh AM, Dale PS, et al. (1996) Early 168 detection of cervical cancer by visual inspection: A population-based study in rural India. *Int J cancer* 68: 770-773.
3. Bingham A, Bishop A, Coffey P, Winkler J, Bradley J, et al. (2003) Factors affecting utilization of cervical cancer prevention services in low resource settings. *Salud Publica de Mexico* 45.
4. Adadevoh S, Adu-Amankwah A, Ahmed S, Awuah B, Paul B, et al. (2004) A Qualitative Evaluation of the Acceptability and Feasibility of a Single Visit Approach to Cervical Cancer Prevention in Ghana. Baltimore: JHPIEGO
5. Min KJ, Lee YJ, Suh M, et al. The Korean guideline for cervical cancer screening. *J Gynecol Oncol*. 2015;26(3):232–239. doi:10.3802/jgo.2015.26.3.232.
6. Bruni L, Albero G, Serrano B, Mena M, Gómez D, Muñoz J, Bosch FX, de Sanjosé S. ICO/IARC Information Centre on HPV and Cancer (HPV Information Centre). *Human Papillomavirus and Related Diseases in Republic of Korea*. Summary Report 10 December 2018. [
7. Park, M.H. Cervical Cancer Screening in Korea. *The Korean Journal of Cytopathology* 2003;14(2): 43-52.
8. S. Ren, K. He, R. Girshick, and J. Sun, “Faster r-cnn: Towards real-time object detection with region proposal networks,” in *Advances in neural information processing systems*, 2015, pp. 91–99.
9. Demarco M, Hu L, Antani S, et al. Automated Visual Evaluation for Cervical Cancer Screening and Management: Promise and Limitations.; 2018.

Example Case (3)

Participant Number	Pap Test	EVA 3 Plus		Other Device ()		HPV Test	Cervical Punch Biopsy
	Result	CDS result	Morphology	Result	Morphology	Result	Result
5272	LSIL	Positive					CIN 1
19873	LSIL	Positive					CIN 1



2. Second two cases show our CDS detects **High Grads Lesion** while the result of biopsy is **CIN 1**

These are the **challenges** and what bring our goal of this training course.

LaonZ
Long-lasting Good Company

Copyright © 2019 LAONZ Inc Co., Ltd. All rights reserved. | Confidential