

Original Article



Development of a prognostic prediction support system for cervical intraepithelial neoplasia using artificial intelligence-based diagnosis

Takayuki Takahashi ,^{1,2} Hikaru Matsuoka ,¹ Rieko Sakurai ,¹
Jun Akatsuka ,^{3,4} Yusuke Kobayashi ,² Masaru Nakamura ,²
Takashi Iwata ,² Kouji Banno ,² Motomichi Matsuzaki ,¹ Jun Takayama ,^{1,5}
Daisuke Aoki ,² Yoichiro Yamamoto ,³ Gen Tamiya ,^{1,5}

¹Statistical Genetics Team, RIKEN Center for Advanced Intelligence Project, Tokyo, Japan

²Department of Obstetrics and Gynecology, Keio University School of Medicine, Tokyo, Japan

³Pathology Informatics Team, RIKEN Center for Advanced Intelligence Project, Tokyo, Japan

⁴Department of Urology, Nippon Medical School Hospital, Tokyo, Japan

⁵Tohoku University Graduate School of Medicine, Miyagi, Japan



Received: Jan 18, 2022

Revised: Apr 7, 2022

Accepted: Apr 29, 2022

Published online: May 16, 2022

Correspondence to

Gen Tamiya

Statistical Genetics Team, RIKEN Center for Advanced Intelligence Project, Nihonbashi 1-Chome Mitsui Bldg., 15F, 1-4-1 Nihonbashi, Chuo-ku, Tokyo 103-0027, Japan.
Email: gen.tamiya@riken.jp

© 2022. Asian Society of Gynecologic Oncology, Korean Society of Gynecologic Oncology, and Japan Society of Gynecologic Oncology

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

ORCID iDs

Takayuki Takahashi

<https://orcid.org/0000-0002-3457-0031>

Hikaru Matsuoka

<https://orcid.org/0000-0003-1754-4568>

Rieko Sakurai

<https://orcid.org/0000-0001-6472-2801>

ABSTRACT

Objective: Human papillomavirus subtypes are predictive indicators of cervical intraepithelial neoplasia (CIN) progression. While colposcopy is also an essential part of cervical cancer prevention, its accuracy and reproducibility are limited because of subjective evaluation. This study aimed to develop an artificial intelligence (AI) algorithm that can accurately detect the optimal lesion associated with prognosis using colposcopic images of CIN2 patients by utilizing objective AI diagnosis.

Methods: We identified colposcopic findings associated with the prognosis of patients with CIN2. We developed a convolutional neural network that can automatically detect the rate of high-grade lesions in the uterovaginal area in 12 segments. We finally evaluated the detection accuracy of our AI algorithm compared with the scores by multiple gynecologic oncologists.

Results: High-grade lesion occupancy in the uterovaginal area detected by senior colposcopists was significantly correlated with the prognosis of patients with CIN2. The detection rate for high-grade lesions in 12 segments of the uterovaginal area by the AI system was 62.1% for recall, and the overall correct response rate was 89.7%. Moreover, the percentage of high-grade lesions detected by the AI system was significantly correlated with the rate detected by multiple gynecologic senior oncologists ($r=0.61$).

Conclusion: Our novel AI algorithm can accurately determine high-grade lesions associated with prognosis on colposcopic images, and these results provide an insight into the additional utility of colposcopy for the management of patients with CIN2.

Keywords: Colposcopy; Artificial Intelligence; Cervical Intraepithelial Neoplasia; Deep Learning

Jun Akatsuka 
<https://orcid.org/0000-0002-2671-7064>
 Yusuke Kobayashi 
<https://orcid.org/0000-0002-4503-2845>
 Masaru Nakamura 
<https://orcid.org/0000-0001-8683-0010>
 Takashi Iwata 
<https://orcid.org/0000-0001-7846-9261>
 Kouji Banno 
<https://orcid.org/0000-0003-2610-2833>
 Motomichi Matsuzaki 
<https://orcid.org/0000-0001-5811-798X>
 Jun Takayama 
<https://orcid.org/0000-0003-2803-6526>
 Daisuke Aoki 
<https://orcid.org/0000-0002-9596-8326>
 Yoichiro Yamamoto 
<https://orcid.org/0000-0003-0799-4338>
 Gen Tamiya 
<https://orcid.org/0000-0002-0597-415X>

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

Author Contributions

Conceptualization: T.T.; Data curation: T.T.; Formal analysis: M.H., S.R.; Funding acquisition: T.G.; Investigation: T.T., T.G.; Methodology: T.T., M.H., S.R., M.M., T.J., Y.Y.; Project administration: T.T., T.G.; Software: M.H., S.R.; Supervision: A.J., K.Y., N.M., I.T., B.K., M.M., T.J., A.D., Y.Y., T.G.; Validation: T.G.; Visualization: T.T., M.H., S.R.; Writing - original draft: T.T., M.H., S.R.; Writing - review & editing: K.Y., N.M., I.T., T.J., A.D., T.G.

Synopsis

High-grade lesion occupancy in the uterovaginal area was significantly correlated with CIN2 patients' prognosis. The number of high-grade lesions in 12 segments detected by an artificial intelligence (AI)-based system was comparable to that detected by senior colposcopists. The overall correct response rate of the AI algorithm for detecting high-grade lesions was 89.7%.

INTRODUCTION

Cervical cancer progresses from precursor lesions called cervical intraepithelial neoplasia (CIN), develops through persistent infection with high-risk human papillomavirus (HR-HPV), and is a leading cause of death among women worldwide [1-3]. The histological classification of cervical dysplasia is currently divided into the low-grade squamous intraepithelial lesion (LSIL) and high-grade squamous intraepithelial lesion (HSIL) [4,5]. In Japan, the histological classification is based on the American Society for Colposcopy and Cervical Pathology (ASCCP) [6]. However, disease classification, namely, grades CIN1, CIN2, and CIN3, is often used in clinical practice [4,5]. This disease grading serves as the basis for the follow-up treatment and management of patients, and LSIL and HSIL include CIN1 and CIN2 or CIN3 subsets, respectively [4,5]. CIN3 has a high risk of progression to cervical cancer [7], and CIN2 is known to sometimes regress, to generally have slow progression, and to occasionally progress to cancer [8,9]. Therefore, to avoid excessive treatment for the prevention of progression to cervical cancer, it is critical to carefully manage CIN2 patients with HSIL for appropriate therapeutic intervention based on the patient's progression because of heterogeneity [10]. When CIN2 is diagnosed, patients are followed up every 3–6 months with cytology and colposcopy and, in some cases, HR-HPV testing to consider further management strategy and possible therapeutic intervention [11]. However, the follow-up might be prolonged, increasing the burden on the patient, and the technique of the colposcopist varies based on the institution.

Colposcopy is an important imaging modality for the diagnosis and classification of cervical lesions [12,13]; therefore, it also plays a central role in cervical cancer prevention. During biopsy, the gynecologist would use a colposcope to manually examine the uterine cervix to determine the site for procuring the tissue samples for a complete microscopic evaluation [14,15]. However, determining the lesion of the cervical tissue that should be obtained relies on subjective assessment and requires substantial practical expertise of the gynecologist, thereby limiting the accuracy and reproductivity of colposcopy [16,17].

With the rapid advancement of artificial intelligence (AI) technology based on deep learning in recent years, it has been widely used in machine learning for medical imaging and applied to various automated processes, such as detection of lung nodules in computed tomography and classification of pathological images [18-20]. Significant progress has been made in the development of AI-based digital colposcopy to improve the efficiency and accuracy of the clinical diagnosis, and several pilot studies of computer algorithms applied to cervical images have been conducted [21-26]. Previous studies have shown excellent pathological differentiation accuracy and have focused on AI-based diagnostics for clustering between LSIL (CIN1) or HSIL (CIN2/3) [27-31]. To the best of our knowledge, no study has evaluated colposcopic imaging alone in patients with CIN2 to predict regression and progression. We hypothesized that the AI system would be able to analyze the colposcopic images objectively

and accurately, and as a result, it would be possible to extract specific findings associated with the prognosis of patients with CIN2 from the colposcopic images.

In this study, we aimed (1) to validate the utility of colposcopic findings as a prognostic factor in patients with CIN2, (2) to develop an AI algorithm that can evaluate the percentage of high-grade lesion occupancy in the uterovaginal area on colposcopy, (3) to confirm its diagnostic accuracy by comparing with the results obtained by multiple gynecologic senior oncologists, and (4) to evaluate the association with prognosis.

MATERIALS AND METHODS

1. Targets for analysis

We first investigated the prognostic indicators of CIN2 on 167 colposcopic images of 126 patients with CIN2 who were diagnosed and followed up by a gynecological oncologist at Keio University Hospital between 2016 and 2018. For the validation analysis of the AI system, we used 593 images from 179 patients who visited our clinic between 2012 and 2020 and were diagnosed with CIN2. Results of the cytology, colposcopy, targeted histology, and HPV examination were obtained from the medical record. We performed colposcopy based on the ASCCP Colposcopy Standards [6] and stratified 593 images as shown in **Table S1**.

The procedures of colposcopy and diagnosis are as follows. The application of 3% acetic acid was used to identify potential lesions; visual changes include response to acetic acid (acetowhitening), characteristics of lesion borders, surface contours, lesion size, and vascular patterns [6]. Changes were visually assessed colposcopically to direct biopsy placement. At least 2 or 3 lesions from the highest visible lesion were biopsied under colposcopic guidance in accordance with the ASCCP guidelines [6] and determined by senior gynecologists who are the supervising physicians of the Japanese Society of Gynecologic Oncology. It was decided to use the Bethesda System [32] and CIN [33] terminology in combination for the histological diagnosis, which was performed in accordance with standard World Health Organization procedures [34]. Several senior gynecologists annotated the study data to ensure objectivity. All patients with CIN2 in our study were diagnosed based on pathological findings. Only cytology and colposcopy data were included for patients who were referred from other hospitals and who were negative for intraepithelial lesion or malignancy (NILM), those who had atypical squamous cells of undetermined significance, and those with mild dysplasia or more. None of the patients with NILM were included. We performed cytology, colposcopy, and targeted histology in patients with CIN1, 2, and 3 via standard procedures as previously reported [35]. The notation of the transformation zone (TZ) includes TZ1 to TZ3 [36]. All cases were well observed in TZ1, and histology was performed for the TZ sample appropriately. Using 593 colposcopic images of CIN2 (90% training data and 10% test data; cross-validation method) from 179 cases, we constructed an AI algorithm and evaluated its prognostic performance. The data were collected mainly from patients who had undergone conization between 2012 and 2020. Of these, 80 colposcopic images that were inappropriate as training data (excessive bleeding or the entire vagina not included in the colposcopic images) were excluded and 593 were finally enrolled.

2. Convolutional neural network-based learning of colposcopic image findings

At our hospital, colposcopy was performed using a colposcope OCS500® (Olympus Corp., Tokyo, Japan). For colposcopic observation, a cotton ball fully soaked in 3% acetic acid

was lightly pressed against the vaginal area of the uterus, and the observation focused on detecting the abnormal findings while processing the acetic acid. The colposcopic images and their areas were color-coded into low- and high-grade lesions in the uterovaginal area based on the diagnoses of several gynecologic oncologists in the electronic medical record. The images were prepared as training data (**Fig. S1**).

3. Analysis for the evaluation of colposcopy results as prognostic indicators

To validate colposcopy as a prognostic indicator for patients with CIN2, survival analysis was performed using the Cox proportional hazards model, one of the most common survival models [37]. The baseline time point was defined as the time at diagnosis of CIN2. The event was defined as the pathological diagnosis of CIN3. The diagnosis of CIN2 and CIN3 is not an irreversible event; therefore, only the first diagnosis of CIN2 and CIN3 was considered. Consequently, we considered 2 prognostic indicators: (1) the proportion of low- and high-grade lesions occupying the surface area of the portio vaginalis including the vaginal lesion and (2) proportion of high-grade lesions occupying the surface area of the portio vaginalis including the vaginal lesion. Both prognostic factors were measured at baseline and treated as numeric variables in each univariate survival analysis. The proportion of lesions was normalized to be an integer from 0 to 12, which originates from a conventional analogy of a clock face in clinical practice to identify which directions the lesions are located on the face of the cervix during colposcopy. This analogy was utilized for the proposed AI algorithm, in which colposcopic images were segmented into 12 sections to calculate the proportion of lesions as the prognostic indicators. The details are presented in the following sections.

4. Construction of an AI algorithm that can calculate the lesion occupancy rate of high-level findings on colposcopy

We constructed the AI algorithm in 3 steps, as shown in **Fig. 1**. The database was trained using Intel® X® CPU E5-1680 v3 @ 3.20 CPU and GeForce GTX 1080, and the total memory was trained on Ubuntu 18.04.3 LTS. The codes were programmed in the TensorFlow 2.4.2 framework with Python 3.7.3. In step 1, we mixed 2 samples (randomly adjusted to the strongest lesion sites) to create a new sample, Mix Match [38], to expand the data. In step 2, we performed predictive classification using U-Net [39], which is often applied to a variety of medical image segmentation problems. U-Net is an algorithm that can classify vaginal and high-grade lesions in a pixel-by-pixel manner. A total of 593 acetate-processed images were provided as input to the model, and because U-Net learns in a pixel-by-pixel manner after annotation, the segmentation model outputs a final prediction region consisting of pixels with the possibility of lesions with low- or high-level findings and a specific vaginal area. As a cross-validation method, pixel-wise annotation classification was performed on the images extracted from 593 colposcopic images. Moreover, we added a dropout layer before the last layer with a dropout rate of 0.6 to avoid overfitting due to the small amount of data.

In step 3, we calculated the high-grade lesion detection rate and overall area under the curve of the average correct response rate using the AI algorithm. The parameters had a learning rate of 0.001, 50 learning cycles, and a batch size of 16.

5. Verification of the correlation between the number of recognized lesions occupying the sites of advanced findings and percentage of advanced lesions recognized by gynecologic oncologists

We constructed an AI algorithm that can be a prognostic predictor of CIN2 and verified this system.

AI-based system to predict progression of CIN2

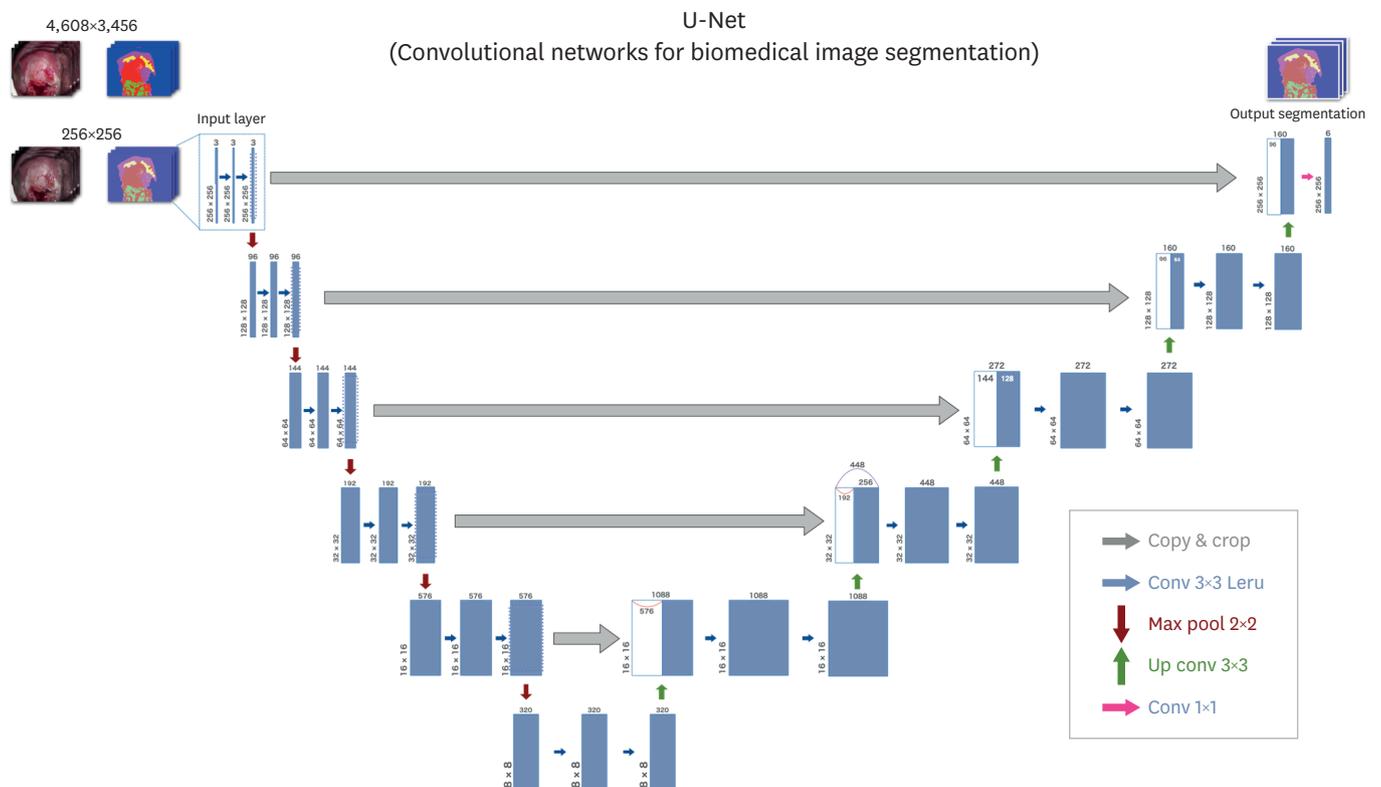
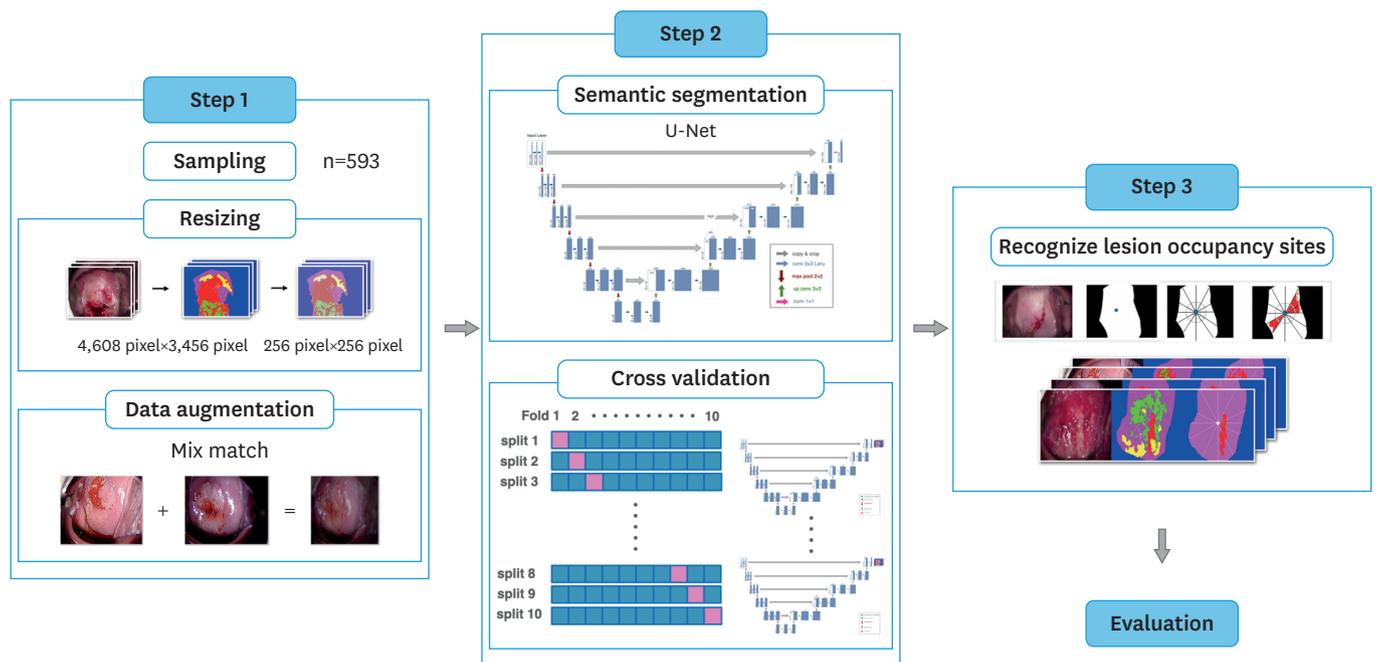


Fig. 1. Steps in building an artificial intelligence algorithm. Step 1: We collected 593 images for the pre-processing step and used the mix-match method to expand the data. Step 2: We used U-net. Step 3: The lesion detection rate and average correct rate of high-level findings using the artificial intelligence algorithm were calculated.

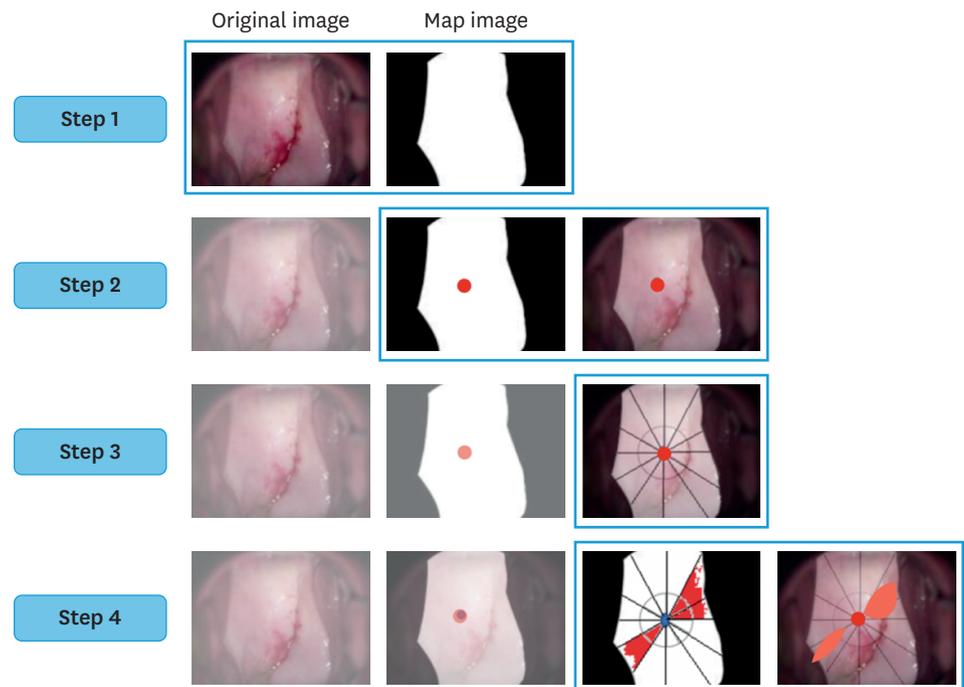


Fig. 2. Step 1: The vaginal and non-vaginal areas were distinguished, and a map image was produced. Step 2: The center of gravity of the uterine vaginal cervix was calculated. Step 3: Twelve sections of the uterine vaginal cervix are shown. Step 4: The artificial intelligence algorithm calculates the number of divisions occupied by high-grade lesions.

The workflow of the AI algorithm's recognition of high-grade lesions was performed in the following steps (**Fig. 2**):

- Step 1: Determine whether the lesion is vaginal or non-vaginal.
- Step 2: Calculate the center of gravity of the cervix.
- Step 3: Divide the cervix into 12 equal parts with the center of gravity as the center and display the 12 parts with the constructed AI algorithm.
- Step 4: Detect and count the altitude findings.

Final evaluation: Calculation of the lesion detection rate using an AI algorithm and verification of its capability as a prognostic predictor of CIN2.

First, we created a tool that automatically displays the percentage of high-grade lesions in the 12 segments, and then this was applied to the constructed AI algorithm to analyze the correlation between the results obtained and percentage of high-grade lesions recognized by multiple gynecologic oncologists. For the final evaluation, we tested prognostic relevance to determine whether the number of high-grade lesions using the AI algorithm was equivalent to the prognostic predictors of CIN2 identified by senior colposcopists. The number of high-grade lesions in the AI algorithm was evaluated as a prognostic factor for CIN2 using Cox regression analysis.

6. Ethics approval

This study was approved by the Clinical Research Ethics Review Committee of Keio University Hospital guidelines for clinical research (approval number: 20190165, approval date: September 30, 2019) and was conducted according to the principles of the Declaration of Helsinki. In this retrospective observational study, we adopted an opt-out approach.

Table 1. Hazardousness for Cox proportional hazards models

Characteristic	Rho	χ^2 value	p-value
Regions occupied by high-grade lesions on colposcopy	0.0329	0.0333	0.855

7. Statistical analyses

The statistical software used was R 3.6.0 software (R Foundation, Vienna, Austria) and Pearson’s correlation function was used to evaluate the correlation coefficient, with a p-value <0.05 being considered a statistically significant difference.

RESULTS

1. Utility of colposcopic findings as prognostic factors for patients with CIN2

We first identified the prognostic factors of patients with CIN2 by analyzing colposcopic images by using the Cox proportional hazards model. We first tested the proportionality of the hazards using Schoenfeld residuals, which resulted in no impropriety for the assumption that the hazard ratio does not depend on time (p=0.855) (**Table 1**). Further, there was no significant pattern of change with time based on a plot of the Schoenfeld residuals shown in **Fig. S2**; therefore, the validity of our use of the Cox proportional hazards model was verified.

Based on colposcopic images, there was no association with prognosis when low- and high-grade lesions were considered as lesion sites (p=0.30, data not shown), although there was a significant association with prognosis (p<0.01) when only high-grade lesions were considered as lesion sites (**Table S2**). We also divided the entire uterine vagina into 12 sections and analyzed the relationship between the proportion of high-grade lesions and prognosis. The results showed that the hazard increased 1.299-fold (95% confidence interval [CI]=1.064–1.586) when high-grade lesions increased in one of the 12 segments of the uterine vagina (p<0.05) (**Fig. 3**).

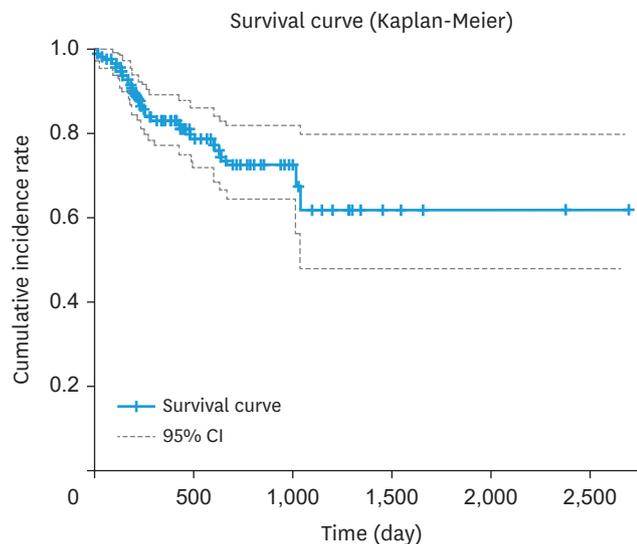


Fig. 3. Progression time from CIN2 to CIN3 based on findings by senior colposcopists. The Kaplan-Meier curve shows the survival time from baseline (time of CIN2 diagnosis) to the time of CIN3 diagnosis. Baseline: time point of diagnosis of CIN2; Event: first time point to CIN3 diagnosis; Survival time: time from baseline to event (n=126). CI, confidence interval; CIN, cervical intraepithelial neoplasia.

Table 2. Detection rate of advanced findings by artificial intelligence algorithm*

Fold	Detection rates for high-grade lesions (%)	Detection rates for the vaginal cervix (%)	Detection rates for other parts (%)	Total accuracy (%)
1	60.4	90.8	83.1	92.6
2	60.6	96.7	76.2	91.8
3	60.9	91.7	76.4	85.5
4	54.8	95.6	68.9	91.2
5	66.6	90.9	83.6	89.7
6	65.2	92.7	87.4	92.7
7	75.0	84.9	82.4	85.2
8	62.2	93.3	83.9	93.3
9	54.4	97.2	75.8	86.7
10	60.6	91.7	81.3	88.2
Average	62.1	92.5	79.9	89.7

*Mean detection rate for dense acetowhite epithelium by artificial intelligence using a 10-fold segment cross-validation method.

The results suggest that the diagnosis of high-grade lesions contributes to the prediction of the risk of CIN2 developing into CIN3 after colposcopy. This supports the importance of the proposed AI algorithm for detecting high-grade lesions for predicting CIN2 prognosis, and thus, we then evaluated the performance of our novel AI algorithm.

2. Performance of our AI algorithm in detecting high-grade lesion

The detection rate of high-grade lesions by AI was 62.1% for recall, and the overall correct response rate was 89.7% (Table 2). In addition, the percentage of high-grade lesions occupying the 12 segments that were recognized by the AI algorithm showed a significant correlation with that recognized by multiple gynecologic oncologists ($r=0.606$, 95% CI=0.554–0.658, $p<0.001$) (Fig. S3). This suggests that the diagnosis based on our AI algorithm is consistent with that detected by skilled senior colposcopists.

3. Our AI algorithm could reflect the disease status of CIN2 and predict the prognosis

As a final evaluation of the AI algorithm, we examined whether the time from CIN2 to CIN3, predicted by the percentage of high-grade lesions and recognized by the AI algorithm, correlates with the prognosis predicted by the actual data of diagnoses from senior colposcopists. Results show that when the number of high-grade lesions recognized by the AI algorithm increased by one segment in the uterovaginal area, the hazard ratio of going from CIN2 to CIN3 increased 1.125 times (95% CI=1.015–1.248, $p=0.025$), which is considered a good result for a prognostic prediction system for CIN2 (Fig. 4).

DISCUSSION

In this study, we first verified whether colposcopy can be used to predict the prognosis of patients with CIN2 in addition to its typical function of diagnosis. In addition, we have created a system that can objectively and accurately evaluate colposcopy imaging findings related to CIN2 prognosis by utilizing deep learning.

Since appropriate management of CIN can help prevent the progression to invasive cancer, it is important to carefully assess the disease status of patients with CIN2. Therefore, the establishment of further indicators apart from the HPV subtype of predictive disease progression for CIN2 will highly contribute to early therapeutic intervention.

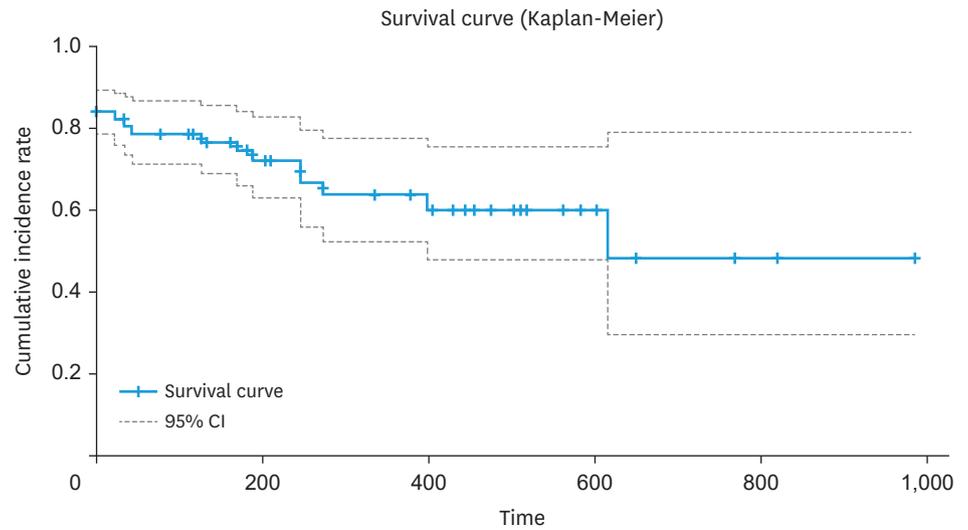


Fig. 4. Progression time from CIN2 to CIN3 based on findings by senior colposcopists. The Kaplan-Meier curve shows the survival time from baseline (time of CIN2 diagnosis) to the time of CIN3 diagnosis. Baseline: time of CIN2 diagnosis; Event: time of the first CIN3 diagnosis; Survival time: time from baseline to event (n=179). CI, confidence interval; CIN, cervical intraepithelial neoplasia.

Colposcopy is widely used to detect cervical precancer and invasive cancer, although some clinical settings lack experienced colposcopists who are capable of making an accurate diagnosis. The experience of the colposcopist significantly affects the accuracy of colposcopy in evaluating cervical lesions and identifying the lesion sites [17]. Junior colposcopists have difficulties in identifying the site with the most lesions and fail to classify the lesion in approximately half of the cases [17]. Therefore, the development of a robust evaluation system is needed. Our proposed AI algorithm can address this problem and estimate with high sensitivity the proportion of the cervix with high-grade lesions. It also showed a relationship between the proportion and risk of CIN2 progression. We believe that our novel algorithm can be applied during colposcopy imaging, especially by inexperienced colposcopists and practitioners. This could help physicians initiate early treatment of patients with CIN2, despite not being experts in colposcopy. In addition, the output of the proportion of high-grade lesions on a 1/12-unit image is based on human observation; therefore, the interpretation is direct and easy for colposcopists who are accustomed to the original procedure. To the best of our knowledge, this study is the first successful case of applying AI technology in clinical practice as a system for predicting the prognosis of CIN2. Our system outperformed the conventional colposcopy method in terms of accuracy and robustness and may be used by junior and senior colposcopists. Furthermore, it may help suspected patients with a progressive disease receive appropriate medical care.

In this study, we found that the percentage of high-grade lesions was significantly correlated with the progression of CIN2. The detection rate of high-grade lesions in our AI algorithm was 61.5%, with an overall mean correct rate of 89.7%. In addition, the percentage of high-grade lesions recognized by the AI algorithm was significantly correlated with that recognized by the physician. The percentage of high-grade lesions on colposcopy recognized by the AI significantly increased the hazard by 1.125 times when the lesion occupied one segment of the vaginal area. These findings demonstrate that the use of AI, which can objectively and accurately determine the occupancy rate, can be an effective tool in predicting the prognosis of CIN2 on the spot based on the change in color tone, which is difficult with conventional

colposcopy. In the evaluation of colposcopy, it is necessary to evaluate not only white lesions, but also borders, surface contours, and vascular patterns; however, in this study, the training numbers are small and we mainly evaluated the white lesions. We can obtain good results by training data for the white lesions and by adding borders, surface contours, and vascular patterns in further studies.

There are 4 limitations to this study that should be addressed in future research. First, the colposcopic images included in this study were not originally taken as training data for machine learning. For example, the orientation of the cervix in the colposcopic images varied depending on the colposcopist who recorded the images. Therefore, the detection rate was reduced during the development of the AI algorithm. To obtain a better quality of training data, it would be effective to align the orientation of the vaginal cervix when taking colposcopic images and to eliminate blood adhesions and vaginal secretions. Second, this study only collected data from a single center, which might have introduced unexpected biases in our dataset. Therefore, to improve the accuracy of prognosis, a more extensive dataset is warranted, and we hope that clinical trials, such as a prospective interventional study, will be conducted in multiple centers to apply the proposed AI prognostic system in clinical practice. Colposcope images can vary in diagnostic accuracy due to differences in magnification and resolution. Therefore, differences in colposcopy conditions among facilities may affect the performance of AI-based diagnosis. As a result, we believe it is necessary to increase the number of low-resolution studies and to improve compatibility and feasibility in the future. If this type of deep learning is achieved, the algorithm could be applied in colposcopy using smartphones [40], which has been reported in recent years. Third, we considered only one prognostic factor in the pre-analysis. As mentioned, testing HPV subtypes is useful for predicting the prognosis of CIN2. Therefore, future studies should consider evaluating the combination of HPV subtypes and other prognostic factors. We would also like to consider exploring other prognostic factors. Fourth, if the lesion persists, evisceration was performed if the patient desires, and follow-up for prognosis has been censored; therefore, this study may be associated with competing risks. Nevertheless, these limitations should not hinder present and future efforts to further improve and to develop AI systems.

In conclusion, our novel AI algorithm can accurately determine high-grade lesions associated with prognosis on colposcopic images, and these results provide an insight into the additional utility of colposcopy for the management of patients with CIN2.

SUPPLEMENTARY MATERIALS

Table S1

The classification of 593 images for target analysis of artificial intelligence algorithm

[Click here to view](#)

Table S2

Estimation results when only the strongest findings of lesion occupation on colposcopy are included in the explanatory variables

[Click here to view](#)

Fig. S1

Colposcopic findings. (A) Normal colposcopic findings. (B) Abnormal colposcopic findings. The red and blue outlines indicate the low-grade lesion and high-grade lesion, respectively. The image areas were recorded from the electronic medical records of patients, as assessed by senior colposcopists.

[Click here to view](#)

Fig. S2

Confirmation of proportional hazards (plot of Schoenfeld residuals). The vertical axis shows the value of the residuals substituted by the degree of high-grade lesion occupancy on colposcopy as a parameter, whereas the horizontal axis shows the number of days. The plot of Schoenfeld residuals shows that there is no significant pattern of change over time.

[Click here to view](#)

Fig. S3

Relationship of the percentage of high-grade lesion occupancy identified by AI and senior colposcopists. The proportion of high-grade lesion occupying the 12 segments that was recognized by the AI algorithm showed a significant correlation with that recognized by multiple gynecologic oncologists (correlation coefficient=0.606, 95% confidence interval=0.554–0.658, $p < 0.001$).

[Click here to view](#)

REFERENCES

1. Castellsagué X. Natural history and epidemiology of HPV infection and cervical cancer. *Gynecol Oncol* 2008;110:S4-7.
[PUBMED](#) | [CROSSREF](#)
2. Wang CW, Liou YA, Lin YJ, Chang CC, Chu PH, Lee YC, et al. Artificial intelligence-assisted fast screening cervical high grade squamous intraepithelial lesion and squamous cell carcinoma diagnosis and treatment planning. *Sci Rep* 2021;11:16244.
[PUBMED](#) | [CROSSREF](#)
3. Schiffman M, Solomon D. Clinical practice. Cervical-cancer screening with human papillomavirus and cytologic cotesting. *N Engl J Med* 2013;369:2324-31.
[PUBMED](#) | [CROSSREF](#)
4. Waxman AG, Chelmow D, Darragh TM, Lawson H, Moscicki AB. Revised terminology for cervical histopathology and its implications for management of high-grade squamous intraepithelial lesions of the cervix. *Obstet Gynecol* 2012;120:1465-71.
[PUBMED](#) | [CROSSREF](#)
5. Darragh TM, Colgan TJ, Cox JT, Heller DS, Henry MR, Luff RD, et al. The lower anogenital squamous terminology standardization project for HPV-associated lesions: background and consensus recommendations from the College of American Pathologists and the American Society for Colposcopy and Cervical Pathology. *J Low Genit Tract Dis* 2012;16:205-42.
[PUBMED](#) | [CROSSREF](#)
6. Khan MJ, Werner CL, Darragh TM, Guido RS, Mathews C, Moscicki AB, et al. ASCCP colposcopy standards: role of colposcopy, benefits, potential harms, and terminology for colposcopic practice. *J Low Genit Tract Dis* 2017;21:223-9.
[PUBMED](#) | [CROSSREF](#)
7. Loopik DL, Int'Hout J, Ebisch RM, Melchers WJ, Massuger LF, Siebers AG, et al. The risk of cervical cancer after cervical intraepithelial neoplasia grade 3: a population-based cohort study with 80,442 women. *Gynecol Oncol* 2020;157:195-201.
[PUBMED](#) | [CROSSREF](#)

8. Tainio K, Athanasiou A, Tikkinen KA, Aaltonen R, Cárdenas J, Hernández, et al. Clinical course of untreated cervical intraepithelial neoplasia grade 2 under active surveillance: systematic review and meta-analysis. *BMJ* 2018;360:k499.
[PUBMED](#) | [CROSSREF](#)
9. Holowaty P, Miller AB, Rohan T, To T. Natural history of dysplasia of the uterine cervix. *J Natl Cancer Inst* 1999;91:252-8.
[PUBMED](#) | [CROSSREF](#)
10. Castle PE, Stoler MH, Solomon D, Schiffman M. The relationship of community biopsy-diagnosed cervical intraepithelial neoplasia grade 2 to the quality control pathology-reviewed diagnoses: an ALTS report. *Am J Clin Pathol* 2007;127:805-15.
[PUBMED](#) | [CROSSREF](#)
11. World Health Organization. WHO guidelines for screening and treatment of precancerous lesions for cervical cancer prevention [Internet]. Geneva: World Health Organization; 2021 [cited 2021 July 6]. Available from: <https://www.who.int/publications/i/item/9789240030824>.
12. Kyrgiou M, Tsoumpou I, Vrekoussis T, Martin-Hirsch P, Arbyn M, Prendiville W, et al. The up-to-date evidence on colposcopy practice and treatment of cervical intraepithelial neoplasia: the Cochrane colposcopy & cervical cytopathology collaborative group (C5 group) approach. *Cancer Treat Rev* 2006;32:516-23.
[PUBMED](#) | [CROSSREF](#)
13. O'Neill E, Reeves MF, Creinin MD. Baseline colposcopic findings in women entering studies on female vaginal products. *Contraception* 2008;78:162-6.
[PUBMED](#) | [CROSSREF](#)
14. Manopunya M, Suprasert P, Srisomboon J, Kietpeerakool C. Colposcopy audit for improving quality of service in areas with a high incidence of cervical cancer. *Int J Gynaecol Obstet* 2010;108:4-6.
[PUBMED](#) | [CROSSREF](#)
15. Bekkers RL, van de Nieuwenhof HP, Neesham DE, Hendriks JH, Tan J, Quinn MA. Does experience in colposcopy improve identification of high grade abnormalities? *Eur J Obstet Gynecol Reprod Biol* 2008;141:75-8.
[PUBMED](#) | [CROSSREF](#)
16. Darragh TM, Colgan TJ, Cox JT, Heller DS, Henry MR, Luff RD, et al. The lower anogenital squamous terminology standardization project for HPV-associated lesions: background and consensus recommendations from the College of American Pathologists and the American Society for Colposcopy and Cervical Pathology. *Arch Pathol Lab Med* 2012;136:1266-97.
[PUBMED](#) | [CROSSREF](#)
17. Bifulco G, De Rosa N, Lavitola G, Piccoli R, Bertrando A, Natella V, et al. A prospective randomized study on limits of colposcopy and histology: the skill of colposcopist and colposcopy-guided biopsy in diagnosis of cervical intraepithelial lesions. *Infect Agent Cancer* 2015;10:47.
[PUBMED](#) | [CROSSREF](#)
18. Ardila D, Kiraly AP, Bharadwaj S, Choi B, Reicher JJ, Peng L, et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat Med* 2019;25:954-61.
[PUBMED](#) | [CROSSREF](#)
19. Nam S, Chong Y, Jung CK, Kwak TY, Lee JY, Park J, et al. Introduction to digital pathology and computer-aided pathology. *J Pathol Transl Med* 2020;54:125-34.
[PUBMED](#) | [CROSSREF](#)
20. Yamamoto Y, Tsuzuki T, Akatsuka J, Ueki M, Morikawa H, Numata Y, et al. Automated acquisition of explainable knowledge from unannotated histopathology images. *Nat Commun* 2019;10:5642.
[PUBMED](#) | [CROSSREF](#)
21. Simões PW, Izumi NB, Casagrande RS, Venson R, Veronezi CD, Moretti GP, et al. Classification of images acquired with colposcopy using artificial neural networks. *Cancer Inform* 2014;13:119-24.
[PUBMED](#) | [CROSSREF](#)
22. Asiedu MN, Simhal A, Chaudhary U, Mueller JL, Lam CT, Schmitt JW, et al. Development of algorithms for automated detection of cervical pre-cancers with a low-cost, point-of-care, pocket colposcope. *IEEE Trans Biomed Eng* 2019;66:2306-18.
[PUBMED](#) | [CROSSREF](#)
23. Miyagi Y, Takehara K, Nagayasu Y, Miyake T. Application of deep learning to the classification of uterine cervical squamous epithelial lesion from colposcopy images combined with HPV types. *Oncol Lett* 2020;19:1602-10.
[PUBMED](#) | [CROSSREF](#)
24. Song D, Kim E, Huang X, Patrino J, Munoz-Avila H, Heflin J, et al. Multimodal entity coreference for cervical dysplasia diagnosis. *IEEE Trans Med Imaging* 2015;34:229-45.
[PUBMED](#) | [CROSSREF](#)

25. 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. *J Natl Cancer Inst* 2019;111:923-32.
[PUBMED](#) | [CROSSREF](#)
26. Bao H, Bi H, Zhang X, Zhao Y, Dong Y, Luo X, et al. Artificial intelligence-assisted cytology for detection of cervical intraepithelial neoplasia or invasive cancer: a multicenter, clinical-based, observational study. *Gynecol Oncol* 2020;159:171-8.
[PUBMED](#) | [CROSSREF](#)
27. Massad LS, Jeronimo J, Katki HA, Schiffman M; National Institutes of Health/American Society for Colposcopy and Cervical Pathology Research Group. The accuracy of colposcopic grading for detection of high-grade cervical intraepithelial neoplasia. *J Low Genit Tract Dis* 2009;13:137-44.
[PUBMED](#) | [CROSSREF](#)
28. Yang B, Pretorius RG, Belinson JL, Zhang X, Burchette R, Qiao YL. False negative colposcopy is associated with thinner cervical intraepithelial neoplasia 2 and 3. *Gynecol Oncol* 2008;110:32-6.
[PUBMED](#) | [CROSSREF](#)
29. Ito Y, Miyoshi A, Ueda Y, Tanaka Y, Nakae R, Morimoto A, et al. An artificial intelligence-assisted diagnostic system improves the accuracy of image diagnosis of uterine cervical lesions. *Mol Clin Oncol* 2022;16:27.
[PUBMED](#) | [CROSSREF](#)
30. Yuan C, Yao Y, Cheng B, Cheng Y, Li Y, Li Y, et al. The application of deep learning based diagnostic system to cervical squamous intraepithelial lesions recognition in colposcopy images. *Sci Rep* 2020;10:11639.
[PUBMED](#) | [CROSSREF](#)
31. Tan X, Li K, Zhang J, Wang W, Wu B, Wu J, et al. Automatic model for cervical cancer screening based on convolutional neural network: a retrospective, multicohort, multicenter study. *Cancer Cell Int* 2021;21:35.
[PUBMED](#) | [CROSSREF](#)
32. Nayar R, Wilbur DC. *The Bethesda system for reporting cervical cytology. Definitions, criteria, and explanatory notes.* 3rd ed. New York, NY: Springer; 2015.
33. Richart RM. Cervical intraepithelial neoplasia. *Pathol Annu* 1973;8:301-28.
[PUBMED](#)
34. Carcangiu M, Kurman RJ, Carcangiu ML, Herrington CS. *WHO classification of tumours of female reproductive organs.* Lyon: International Agency for Research on Cancer; 2014.
35. Nishio H, Iwata T, Nomura H, Morisada T, Takeshima N, Takano H, et al. Liquid-based cytology versus conventional cytology for detection of uterine cervical lesions: a prospective observational study. *Jpn J Clin Oncol* 2018;48:522-8.
[PUBMED](#) | [CROSSREF](#)
36. Quaas J, Reich O, Küppers V. Explanation and use of the Rio 2011 colposcopy nomenclature of the IFCCP (International Federation for Cervical Pathology and Colposcopy): comments on the general colposcopic assessment of the uterine cervix: adequate/inadequate; squamocolumnar junction; transformation zone. *Geburtshilfe Frauenheilkd* 2014;74:1090-2.
[PUBMED](#) | [CROSSREF](#)
37. Cox DR. Regression models and life-tables. *J R Stat Soc Series B Stat Methodol* 1972;34:187-202.
[CROSSREF](#)
38. Berthelot D, Carlini N, Goodfellow I, Oliver A, Papernot N, Raffel C. MixMatch: a holistic approach to semi-supervised learning. In: *Proceedings of the 33rd International Conference on Neural Information Processing Systems (NIPS'19)*; 2019 Dec 8-14; Vancouver, Canada. La Jolla, CA: Neural Information Processing Systems Foundation; 2019. p. 5049-59.
39. Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. In: *Proceedings of the 18th International Conference on Medical Image Computing and Computer Assisted Intervention*; 2015 Oct 5-9; Munich, Germany. Cham: Springer; 2015. p. 234-41.
40. Aydın S, Karasu AF, Maraşlı M, Bademler N, Kıran G, Dural HR. Reliability and diagnostic performance of smartphone colposcopy. *Int J Gynaecol Obstet* 2021;155:404-10.
[PUBMED](#) | [CROSSREF](#)