

The Role of Machine Learning Artificial Intelligent Models in Assisting Colposcopists: A Literature Review

Thomas J^{1*} Kovilveetil AN¹ and Ahmed AIH²

¹Canterbury Christ Church University, UK

²Medway Maritime Hospital, Canterbury Christ Church University, UK

*Corresponding author:

Jane Thomas,
Canterbury Christ Church University, UK

Received: 28 Oct 2024

Accepted: 28 Nov 2024

Published: 03 Dec 2024

J Short Name: AJSCCR

Copyright:

©2024 Thomas J, This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and build upon your work non-commercially.

Citation:

Thomas J. The Role of Machine Learning Artificial Intelligent Models in Assisting Colposcopists: A Literature Review. *Ame J Surg Clin Case Rep.* 2024; 8(4): 1-7

Keywords:

Artificial intelligence; Machine learning models; Colposcopy; Cervical cancer

1. Abstract

1.1. Background: Due to the subjective nature of colposcopic evaluation, automated methods for detecting cervical lesions have been introduced. However, colposcopic evaluation is time-consuming and is not exempt from error. This literature review aims to evaluate the diagnostic performance of Artificial Intelligence (AI) models concerning classification and biopsy processes and their role in assisting colposcopists.

1.2. Methods: Detailed searches were performed on four databases: Pubmed, Embase, Scopus, and Cochrane to find papers published from January 2020 to August 2023. Articles that applied Machine Learning (ML) models for the classification of cervical lesions, guidance on biopsy and those that evaluated the performance of AI assisted colposcopists were included in the review.

1.3. Results: AI-assisted colposcopists showed higher sensitivity, specificity, and diagnostic accuracy, suggesting that AI can serve as a potentially cost-effective tool, particularly in regions with limited access to experienced colposcopists. There are several limitations detected in the conducted studies and with AI technology as a whole, which have been highlighted in the review.

1.4. Conclusion: ML models that incorporate multidimensional data are proving to be beneficial in assisting human evaluation. The next target is to implement these models in prospective, randomised controlled trials, to eliminate all elements of bias and to provide more robust results.

Funding: No primary source of funding.

2. Introduction

Cervical cancer is the fourth most commonly detected cancer and

the fourth leading cause of cancer-related deaths among women on a global scale (Allahqoli et al., 2022). Human Papillomavirus (HPV) is responsible for about 70% of precancerous lesions and cervical carcinomas. HPV infection can lead to malignant changes, with oncoproteins disrupting tumor suppressor proteins (Bedell et al., 2020). Fortunately, about 90% of HPV infections resolve spontaneously within two years (Bedell et al., 2020). Due to its prolonged precancerous phase, cervical cancer is highly preventable and it is curable when detected early (Campos et al., 2022). Screening women at least once in their lives, especially at age 35, can reduce cervical cancer mortality risk by about 70% (Bedell et al., 2020).

Cervical cancer screening is categorized into three types: colposcopy (visual impression of the cervix), cytology (microscopic/cellular study of the cervix), and HPV status (Desai et al., 2022). Colposcopy involves evaluating the following features: response to acetic acid, lesion size, lesion margins and borders, vascular patterns, iodine uptake as well as signs of invasive cancer including necrosis (Bedell et al., 2020, Reich and Pickel, 2021). These serve as visual cues that guide biopsy sampling that will be sent for histopathological evaluation in order to confirm diagnosis (Allahqoli et al., 2022, Fu et al., 2022). Additionally, miscellaneous cervical features like condyloma, polyps, stenosis, and inflammation are noted. The accuracy of colposcopy is measured by the high-grade lesions that are missed in terms of false negative rates (Khan et al., 2017).

Unfortunately, in low- and middle-income countries (LMICs), colposcopy faces challenges including expert subjectivity, a lack of trained professionals/advanced infrastructure, and poor uniformity

due to variations in reporting and screening modalities (Xue et al., 2020). Colposcopy is a subjective examination, hence the diagnostic proficiency of numerous colposcopists falls short, and up to 40% of cases involving High-Grade Squamous Intraepithelial Lesions (HSIL) are misdiagnosed or missed in LMICs. Hence, the importance of Artificial Intelligence models cannot be emphasised enough when diagnosing cervical precancers and making decisions regarding management (Chen et al.,2022).

Artificial Intelligence (AI) has been increasingly used for disease diagnosis, including cervical cancer. Deep Learning methods can recognise images, learn classifications, extract regions of interest, and process data automatically (Allahqoli et al., 2022). They can handle big data information and generate objective outputs, facilitating decision-making processes in order to improve objectivity (Hashimoto et al., 2018, (Xue et al., 2020). Machine Learning techniques have the ability to process the large volume of data generated in healthcare, aiding in the formation of cost-effective guidelines and protocols while expanding the scope of clinical research and surgical training (Gao et al.,2020;Yangetal.,2021).

The pivotal role of Convolutional Neural Networks (CNNs) is to extract features from data input without human guidance, enabling automatic detection of attributes such as tumour size, location, and border extension (Hosny et al., 2018). Each image feature is filtered and brought into a training layer that the model must become proficient in, with multiple layers building upon this in forward propagation. In cases of errors or loss, the AI model undergoes backward propagation to revise and each time the model

goes through the training layer (epochs), the model becomes more efficient. Optimal image conditions for CNN models are determined after testing and validation cycles, revolving around factors like scanning models, image pixel size, and resolution (Cho et al., 2022). Training models on extensive datasets enhances their predictive accuracy, although the imbalance or lack of datasets leads to employing techniques like data augmentation to address bias. Data augmentation revolves around rotating or flipping one image to produce a new image that can serve as additional training data (Cui and Zhang, 2021).

AI should complement colposcopists rather than replace them, with a focus on guiding precise cervical biopsies to improve diagnostic performance (Xue et al., 2020). The main aim of the study is to conduct a literature review using a systematic approach regarding the efficacy of AI-related ML models in colposcopy procedures and their role in the management of cervical cancer.

3. Methods

3.1. Search Strategy and Information Sources

The relevant databases that were used to conduct the literature search included Pubmed, Embase, Web of Science, and Cochrane. Keywords related to “artificial intelligence”, “machine learning models”, “colposcopy” and “cervical cancer” were used to identify MeSh keywords, which is a list of standardised subject headings, that was employed to ensure a wide scope of available literature was searched. Additionally, Boolean operators (AND,OR) were employed to enhance the selection of entries.

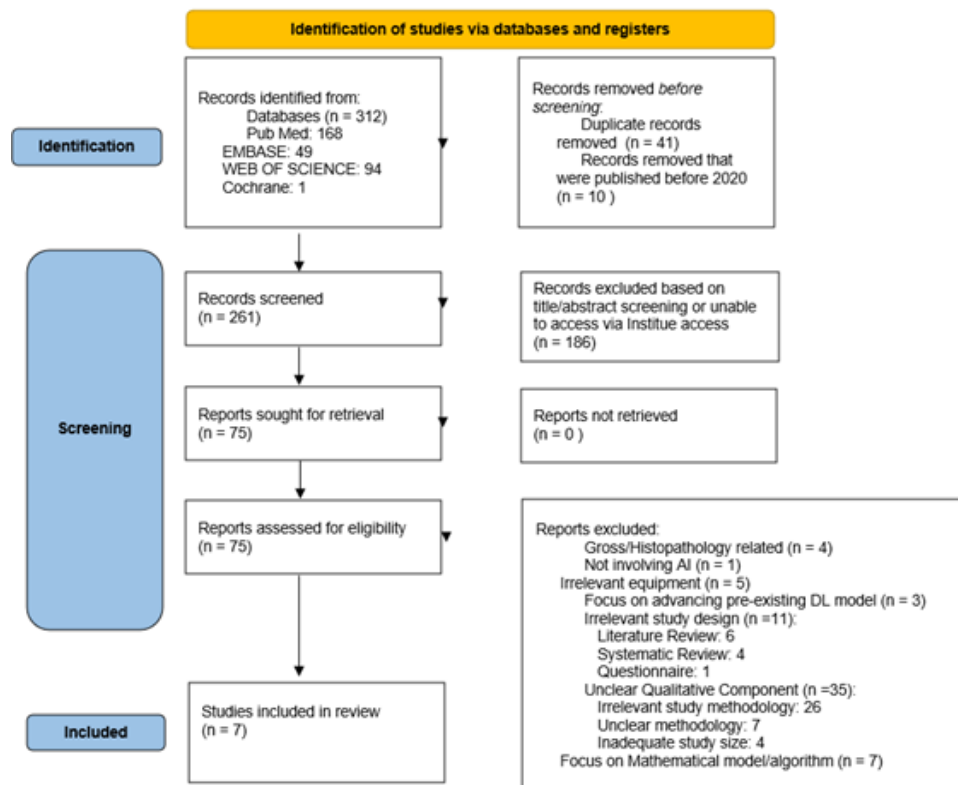


Figure 1: Showing study collection flowchart

3.2. Inclusion and Exclusion Criteria

The author included all types of observational, prospective, and retrospective studies. Additionally, all experimental studies including randomised controlled trials were searched for. Studies that were conducted anywhere in the world, published in the English Language, and only articles published from 2020 onwards were included. This is because there has been a sudden boom and change in practice in the field of Artificial Intelligence in the past few years only. Studies purely focused on cervical precancerous/cancerous lesions were included. Only studies focused on colposcopy as the main screening method involving AI/ML models and their integration with colposcopy were included. AI models incorporating the role of demographics and other screening methods including cytology and HPV status were also included. Additionally, only studies focused on making a comparison between AI models, independent colposcopists, and/or AI-guided colposcopists were included. Exclusion criteria include systematic reviews and meta-analyses, questionnaires, literature reviews, abstracts alone, case reports and conference presentations. Studies purely focused on comparing the efficacy of different types of AI models, studies on animals, the genomic field, molecular biomarkers, spectroscopy, mathematic models, and the prognosis of cervical cancers were excluded.

3.3. Data Extraction

Data from all articles were extracted and reported from the studies using a customised checklist to ensure that the paper met the selection process and to ensure the consistency of this step for all articles reviewed. The outcomes the author chose to specify in the review are the accuracy, sensitivity, and specificity of the AI models in classifying precancerous cervical lesions and the biopsy site sensitivity. Additionally, a main emphasis was placed on finding out if there was a significant difference when a colposcopist was guided by an AI model as opposed to independently making these decisions. The author did not have specific criteria for participant characteristics as different studies had slight variations in their participant inclusion criteria.

The effect measures that the author was particularly interested in include accuracy, sensitivity, and specificity rates. Additionally, p values, confidence intervals, Receiver Operating Characteristic (ROC) curves, and Area Under the Curve (AUC) values were focused on to make an analysis. The author was not able to find an adequate number of randomised control trials specific to the role of AI models in assisting junior doctors. This is specifically because AI models are still in the nascent stages and require further improvement and approval to be incorporated into randomised control trials.

4. Results

4.1. Study Selection

The MeSH and Boolean operator words as described in the Meth-

ods section were employed and the following flow diagram shows the number of hits from each database and how the author further refined the search strategy to finalise the number of papers that have been included in this review. A total of only one randomised control trial and 6 retrospective analytical studies were included. Wu et al. (2023) conducted a study measuring the effectiveness of an AI system called Colposcopic Artificial Intelligence Auxiliary Diagnostic System (CAIADS) in assisting junior colposcopists with cervical lesion grading and biopsy decisions. The study compared CAIADS, junior colposcopists, and CAIADS-assisted junior colposcopists using the same data while keeping them blind to the senior colposcopist's diagnosis. The study emphasized the importance of both human and AI collaboration rather than a direct comparison between them. Strengths of the paper included providing insights into AI's role in guiding colposcopists regarding biopsy decisions and sites. Additionally, the study has provided evidence showing that AI-assisted colposcopists have a narrower range of higher sensitivity (66.7–84.5%) compared to independent colposcopists (19.5 to 100%). An obvious limitation of this study is that the junior colposcopist was not given a washout period between when they were assessed as independent colposcopists and when they were guided by the ML model. Either a washout period should have been implemented or a new set of data should have been employed to avoid any form of bias that may result from memory or assumption. External validation and population-based studies are needed to further assess CAIADS' practicality (Wu et al., 2023).

Kim et al. (2022) conducted a study to assess the feasibility of using an AI system to assist in diagnosing high-grade CIN lesions compared to human interpretation of cervical images. The study included patients from two institutions in Korea who had undergone colposcopy-guided biopsy due to abnormal cervical cytology or positive HPV status in 2020. The Cerviray AI® system was developed with the idea of combining AI with expert involvement for screening high-grade cervical abnormalities. The average AUC is higher when AI models were combined with doctors compared to when the doctors made their interpretations independently (AUC 0.784; 95% CI 0.723-0.845 vs 0.734; 95% 0.669-0.799).

Conventional real-time colposcopic evaluation revolves around analysing the rate of acetowhite changes and slight differences in the response or light reflection over a period of time. However, in this study, only a single colposcopic image is evaluated and the subtle changes that occur over seconds cannot be accurately evaluated. This will affect the general reliability of the results, especially when it is to be considered for real-life clinical scenarios. The model had less training with samples of a normal cervix and this could have affected the final output of the model. (Kim et al., 2022). The same authors conducted a study in 2023 in the same institute but as a multicenter, crossover design, double-blind, randomized controlled trial (RCT) which is a strength in itself as it

reduces the possibility of bias. Based on the AUC of the ROC, the study found that AI-assisted interpretation resulted in significantly better performance than the control diagnosis, with improved sensitivity, specificity, and diagnostic accuracy (0.73 compared to 0.62, $p < 0.001$) when interpreters were assisted with the AI model. It showed that AI could assist non-specialist healthcare providers in making more accurate decisions regarding biopsies or referrals to higher centres. The study noted that individual experiences or preferences among healthcare providers might affect diagnostic accuracy when using AI (Kim et al., 2023).

Nonetheless, the study had limitations, including the need for additional endocervical evaluations in some cases, the lack of real-time colposcopic diagnoses, and the high amount of unknown information regarding HPV infections or other demographic criteria. Both the studies conducted in 2022 and 2023 only included acetic acid images and excluded images with Lugol's solution applied to the cervix, which is questionable because iodine images of the cervix are quite an integral part of a colposcopy. Nonetheless, both these studies concluded that AI interpretation of cervical images shows promise as an assistive tool in conjunction with human evaluation and has the potential to be a cost-effective diagnostic tool for high-grade cervical lesions, particularly in LMICs with limited access to experienced colposcopists. (Kim et al., 2023).

Chen et al. (2022) combined patient demographics, HPV results, cytology and colposcopic interpretations to evaluate the suitability of different ML models for predicting HSIL or worse in individuals referred for colposcopy. The main aim of the study was to convert the subjective nature of colposcopic evaluation into an objective predictive model appropriate for clinical application. Experienced colposcopists were asked to classify colposcopic impressions (normal, low grade lesions, high grade lesions, or cancer). Histopathological evaluation from an experienced pathologist was considered the gold standard. In this study, six common ML models with different algorithms to extract specific features were employed in the study. Additionally, the study emphasized the importance of considering patient demographics and other factors like age, gravidity, number of deliveries, and menopausal stage. Labour-induced cervical injury and elevated levels of pregnancy-related hormones and immunosuppression makes the Squamo-Columnar Junction (SCJ) prone to HPV infections. ML models integrating demographics and screening data significantly improved AUC (0.067 to 0.099) and sensitivity (11.55 to 14.88%). Reproductive factors and menopausal status were also linked to HSIL+, highlighting the complex relationship between these factors and cervical cancer risk. The absence of certain variables, such as HPV infection history, smoking status, and sexual behavior makes this study incomplete. Future efforts should focus on improving specialist diagnostic abilities and exploring the integration of non-imaging clinical data with AI-based colposcopy for a more holistic practice (Chen et al., 2022).

Liu et al. (2021) employed a robust approach by utilising a large dataset containing numerous patient cases, augmented images, and predefined pixel and resolution specifications. They conducted training, validation, and testing phases to ensure the AI system's competence. To assess the AI's performance, they randomly chose 300 samples from the test dataset and enlisted a junior colposcopist to provide diagnoses for comparative analysis. The AI system demonstrated diagnostic capabilities equally comparable with those of a senior colposcopist when distinguishing between Normal Cervix (NC) and Low Grade Squamous Intraepithelial Lesions (LSIL) lesions. Additionally, it enhanced the diagnostic accuracy of junior colposcopists. Certain patients with cervical polyps, cervical or uterine benign neoplasms, or any other cervical abnormalities/surgical interventions were excluded from this study. This exclusion criterion limits the practical replicability of the CAD model in real clinical scenarios. Junior colposcopists had a much lower accuracy and wider confidence interval: 0.753 (95% CI: 0.700 – 0.801) (Liu et al., 2021).

Xue et al. (2020), completed a study aimed to utilise CAIADS (Computer-Aided Intelligent Analysis and Diagnosis System) for classifying colposcopic impressions. CAIADS also sought to categorize colposcopic impressions into two biopsy grades: low and high. This classification was a key ingredient in determining appropriate biopsy thresholds during colposcopy by targeting the pathology-confirmed HSIL+. Patient records were divided into two groups: those with at least five adequate colposcopic images, taken at specific times (0 s, 60 s, 90 s, 120 s, and 150 s), and those with clinical data like cytology and HPV status. The colposcopic grading process involved cervix detection, and feature extraction, with the addition of other algorithms to identify a region of interest, segment the lesion area, and provide biopsy site guidance. An accurate lesion area segmentation helped reduce unnecessary biopsies outside of lesion regions.

CAIADS achieved an overall agreement of 82.2% when grading colposcopic impressions against the gold standard histopathological evaluation, surpassing original interpretations made by the expert. The inclusion of non-image data significantly improved CAIADS's diagnostic performance and this was reflected in higher AUC values when it came to biopsy site sensitivity. The overall agreement between CAIADS-graded colposcopic impressions and the gold standard histopathology was higher than the expert interpretations (82.2% versus 65.9%, $p < 0.001$). Additionally, incorporating non-image information significantly improved CAIADS's diagnostic performance, as evidenced by increased AUC values at different biopsy thresholds ($p < 0.001$). Furthermore, CAIADS outperformed the original interpretations by colposcopists when considering biopsy thresholds with a narrower and higher confidence interval and $p < 0.001$.

The authors recommend integrating CAIADS into local colposcopy clinics to support colposcopists during procedures. The main

drawback of this study is that it focused on grading colposcopic impressions related to purely cervical neoplastic lesions. They excluded any patients with other cervical pathologies including polyps, stenosis, and condyloma and this is not a representation of a universal population. To make the AI model more valuable and applicable on a global scale, the authors must consider refining their CAIAD's algorithms by training them to detect neoplastic changes amid other cervical pathologies (Xue et al., 2020).

Yuan et al. (2020) conducted a study involving three ML models that were developed and evaluated to enhance the accuracy of colposcopy-guided biopsies for cervical squamous intraepithelial lesions, especially HSILs. The ML models were compared with 5 colposcopist experts to identify whether they can serve as a tool to support colposcopists during their assessment and biopsy processes (Yuan et al., 2020). The study enrolled 22,230 cases for model training and evaluation; the included cases were divided into training, validation, and test sets between three separate models that had different algorithms. The authors of this study did not mention how the number of normal, LSIL, and HSIL cases were calculated to serve as training datasets for the three different AI models, each calibrated with a different algorithm (Yuan et al., 2020). The study excluded data that were diagnosed as invasive cervical cancer or images that included bleeding or vaginal discharge that may have disrupted the clarity of the image. This introduces limitations to the study as excluding these forms of data prevents the ML model from being adequately trained and prepared for real-time colposcopy scenarios and instead may lead to misclassification and prolong commencement of management (Yuan et al., 2020). The biopsy accuracy of colposcopists was higher (ranging from 22.22% to 30.57%) than the detection model in both ordinary and high-definition images. This meant that the detection model did require slightly more biopsies per case compared to colposcopists (2.79 vs. 2.39). Overall, the classification model achieved an accuracy of 84.10% in ordinary images, surpassing the accuracy of colposcopists and five experts. Surprisingly, when high-definition images were used the diagnostic accuracy of the AI model decreased to 63.8%. This disparity may be attributed to differences in image feature distributions such as brightness, contrast, color saturation, and other factors. High-definition images presented challenges because brighter and more saturated images lead to emphasis on aceto-white areas and iodine non-staining areas in normal cases. This leads to poor classification and unsatisfactory performance of the classification model (Yuan et al., 2020). Compared to many other studies, this study incorporated features from both acetic and iodine images during the training and validation phases. The inclusion of both sets of images enhances the ability to accurately detect cervical lesions and a clinically sound diagnosis (Yuan et al., 2020).

5. Discussion

There are several limitations to this literature review. As men-

tioned in the methods and results section, only research studies from Embase, Pubmed, Web of Science, and Cochrane were included. Grey literature, clinical trial registries, and other individual articles were not hand-picked via Google Scholar or any other means. In addition to time constraints, this means that the number of articles reviewed is limited and therefore does not provide a robust overview of the literature that is available. Another limitation of this review is that several of the selected articles are retrospective and therefore certain important demographic or clinical data that would have been beneficial to the outcome of the overall study may not have been collected. The reason why the majority of the studies that were conducted are retrospective is due to the ease of collecting a larger sample size this way. AI models are not commonly used in experimental studies or clinical studies on real-time data sets. The few that are done will only be performed on a small samples size and these studies are very unstable because any change in data input will completely change the AI algorithm and outcome results. Another limitation is that most of the studies is that they excluded cervical images that involved other cervical pathologies including stenosis or polyps. This means that the research conducted is not generalisable to other populations and is not representative of real-time scenarios that clinicians face on a day-to-day basis and therefore may not be robust enough to be implemented in clinical practice just yet.

Although the purpose of these AI models is to aid untrained professionals specifically in less developed countries, these models may not be valid enough to implement, especially because the quality of screening data or colposcopy images is not standardised. Different colposcopy practices employ different types of colposcopic equipment, including cervical annotation and classification and therefore the generalisability and the universal application come into question (Wu et al., 2023). Additionally, it may be demanding to require high-definition images that require large storage or memory space (Wu et al., 2023).

Colposcopic impressions are usually underclassified than overclassified and this is a result of subjective decisions made by colposcopists due to unclear diagnostic criteria with regards to colposcopy. Cytology, HPV screening, and demographic details should be considered when determining the need for biopsy, immediate management, or follow-up (Chen et al., 2022). AI models aid colposcopists to perform better, but perform even better when additional clinical information is fed through the system. Other useful information that could have been collected by these studies include smoking history, age of first sexual contact, and number of sexual partners.

Additionally, various AI models employ different algorithms and some require more time and memory to run and these issues lead to difficulty in universal application. More clinical trials using real-time colposcopy images should be employed to consider the effectiveness of AI models. Additionally, more studies need to be

conducted to review different AI models and different algorithms to compare and establish a superior method (Chen et al., 2022). Furthermore, in actual practice, the colposcopist notices the dynamic changes in the acetowhite epithelium over a period of time. Most of the studies available and included in this review purely focused on a single static colposcopy image. Perhaps, several other algorithms have to be generated and ML models must be trained to analyse time-sensitive colposcopic images from one subject/patient and to make a comprehensive analysis. This may take a few more years to perfect.

These artificial intelligence models have brought necessary benefits, but they certainly come with challenges. The influx of data compromises patient privacy and confidentiality and this has resulted in stricter control over patient data (Hosny et al., 2018). Cryptonets are advanced AI network models that prioritise data encryption and ensure that authorised individuals with decryption capabilities can analyse patient data. To allow for an effective analysis of extensive datasets, the following innovative developments are required: better internet speed, increased power consumption, evolved hardware, and larger memory space- all of which contribute to higher expenses (Yang et al., 2021). Furthermore, insurance companies have failed to reach a united decision regarding coverage for diagnostic AI services, which will then inevitably lead to patients bearing the costs of advanced services that were used without their specific request (Jiang et al., 2020).

Due to the potential deterioration of clinical skill sets and possible compromise of patient safety, physicians must exercise caution when it comes to developing reliance on AI technologies. There is rising concern that eventually, human intervention will not be able to override autonomous AI decisions and there is resistance to incorporating AI in medical practice, due to negative attitudes among health professionals (Wall & Krummel, 2020). AI systems include hidden layers that humans cannot interpret, called the “black box” design and this raises concerns about the accuracy and validity of results from a healthcare and patient perspective. Questions about the safety of DL models persist because AI, while powerful, will not be comparable to the unique interconnective and associative abilities of the human mind (Hosny et al., 2018). Additionally, a concern that must be addressed is whether clinicians should bear the sole responsibility for an inaccurate diagnosis and inappropriate mode of management, or if the blame should be shared with machines (Jiang et al., 2020).

AI in colposcopy faces challenges, including integration with existing equipment, interpretability, and the continued need for human guidance and management (Xue et al., 2020). Digital colposcopy offers advantages such as educational support, permanent record-keeping, quality control, and telemedicine consultations. This results in serving populations in need and improving clinical management and healthcare services (Bedell et al., 2020).

6. Conclusion

In summary, AI models offer time-saving benefits for clinicians and effectively pinpoint biopsies that demand immediate analysis. Furthermore, improvements in data storage are currently being made and this will lead to the emergence of hybrid networks that incorporate multi-dimensional data including colposcopy images, microbiology, proteomics, and even genomics. Although there are a few limitations, machine learning models hold a predominantly beneficial role in advancing precision medicine and providing customised healthcare delivery. AI-guided colposcopy should only be used as a means to assist junior colposcopists and not to replace them. Instead, AI models should continuously be monitored to improve their ability to assist colposcopists. AI-assisted colposcopists could lead to improved diagnostic decisions and enhance clinical management and aid in the scarcity of experienced colposcopists in LMICs (Xue et al., 2020).

It is essential to allow for adequate supervision of data and privacy that revolves around legal and ethical regulations. Establishing an ethical standard regarding personal data protection will remove the hindrances to innovation and further technological development for the improvement of human health (Xue et al., 2020). More prospective studies and RCTs should be conducted to ensure that robust AI models have adequately built algorithms to accurately classify real-time colposcopy images, guide professionals on the number and exact site for biopsies, and incorporate multi-modal clinical data to accurately provide a diagnosis. Additionally, AI models can be trained to provide an accurate prognosis and survival outcome. All research institutes and universities should take the initiative in translating their AI algorithms into practical clinical applications and public health services. This proactive step will not only facilitate the entry of AI-guided digital colposcopy into the healthcare market but also help reach the worldwide goal of eradicating cervical cancer (Xue et al., 2020).

References

1. Allahqoli L, Laganà AS, Mazidimoradi A, Salehiniya H, Günther V, Chiantera V, Karimi Goghari S, Ghiasvand MM, Rahmani A, Momenimovahed Z, Alkatout I. Diagnosis of cervical cancer and pre-cancerous lesions by artificial intelligence: A systematic review. *D*
2. Bedell SL, Goldstein LS, Goldstein AR, Goldstein AT. Cervical cancer screening: Past, present, and future. *Sexual Medicine Reviews*. 2020; 8(1):28–37.
3. Campos PL, Guimarães ICCV, Fialho SCAV, Martins CAO, Rodrigues FR, Velarde LGC, Monteiro DSA. Association of Swede Score and 2011 IFCPC nomenclature in women with abnormal cytology. *Revista Brasileira de Ginecologia e Obstetria / RBGO Gynecology and Obs*

4. Chen M, Wang J, Xue P, Li Y, Qiao YL. Evaluating the feasibility of machine-learning-based predictive models for precancerous cervical lesions in patients referred for colposcopy. *Diagnostics*. 2022; 12(12):3066.
5. Cho BJ, Kim JW, Park J, Kwon GY, Hong M, Jang SH, Bang H, Kim G, Park ST. Automated diagnosis of cervical intraepithelial neoplasia in histology images via deep learning. *Diagnostics*. 2022; 12(2):548.
6. Cui M, Zhang DY. Artificial intelligence and computational pathology. *Laboratory Investigation*. 2021; 101(4):412–422.
7. Desai KT, Befano B, Xue Z, Kelly H, Campos NG, Egemen D, Gage JC, Rodriguez A, Sahasrabudde V, Levitz D, Pearlman P, Jeronimo J, Antani S, Schiffman M, Sanjosé S. The development of ‘automated visual evaluation’ for cervical cancer screening: The promise
8. Fu L, Xia W, Shi W, Cao G, Ruan Y, Zhao X, Liu M, Niu S, Li F, Gao X. Deep learning-based cervical screening by the cross-modal integration of colposcopy, cytology, and HPV test. *International Journal of Medical Informatics*. 2022; 159: 104675.
9. Gao Y, Kruger U, Intes X, Schwaitzberg S, De S. A machine learning approach to predict surgical learning curves. *Surgery*. 2020; 167(2): 321–327.
10. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. *Nature Reviews Cancer*. 2018; 18(8): 500–510.
11. Jiang Y, Yang M, Wang S, Li X, Sun Y. Emerging role of deep learning-based artificial intelligence in tumor pathology. *Cancer Communications*. 2020; 40(4): 154–166.
12. Karasu Benyes Y, Welch EC, Singhal A, Ou J, Tripathi A. A comparative analysis of deep learning models for automated cross-preparation diagnosis of multi-cell liquid Pap smear images. *Diagnostics*. 2022; 12(8): 1838.
13. Khan MJ, Werner CL, Darragh TM, Guido RS, Mathews C, Moscicki AB, Mitchell MM, Schiffman M, Wentzensen N, Massad LS, Mayeaux EJ, Waxman AG, Conageski C, Einstein MH, Huh WK. ASCCP colposcopy standards: Role of colposcopy, benefits, potential harms, and te
14. Kim S, Lee H, Lee S, Song JY, Lee JK, Lee NW. Role of artificial intelligence interpretation of colposcopic images in cervical cancer screening. *Healthcare*. 2022; 10(3): 468.
15. Kim S, An H, Cho HW, Min KJ, Hong JH, Lee S, Song J, Lee JK, Lee NW. Pivotal clinical study to evaluate the efficacy and safety of assistive artificial intelligence-based software for cervical cancer diagnosis. *Journal of Clinical Medicine*. 2023; 12(12):
16. Liu L, Wang Y, Liu X, Han S, Jia L, Meng L, Yang Z, Chen W, Zhang Y, Qiao X. Computer-aided diagnostic system based on deep learning for classifying colposcopy images. *Diagnostics*. 2021; 9(13): 1045.
17. Munetoshi A, Hashimoto K. Artificial intelligence in gynecologic cancers: Current status and future challenges – A systematic review. *Artificial Intelligence in Medicine*. 2021: 102164.
18. Reich O, Pickel H. 100 years of iodine testing of the cervix: A critical review and implications for the future. *European Journal of Obstetrics & Gynecology and Reproductive Biology*. 2021; 261: 34–40.
19. Robertson S, Azizpour H, Smith K, Hartman J. Digital image analysis in breast pathology—from image processing techniques to artificial intelligence. *Translational Research*. 2018; 194: 19–35.
20. Stanley RJ, Sornapudi S, Addanki R, Stoecker W, Long R, Zuna R, Frazier S, Antani S. Automated cervical digitized histology whole-slide image analysis toolbox. *Journal of Pathology Informatics*. 2021; 12(1): 26.
21. Yuan C, Yao Y, Cheng B, Cheng Y, Li Y, Li Y, Liu X, Cheng X, Xie X, Wu J, Wang X, Lu W. The application of deep learning-based diagnostic system to cervical squamous intraepithelial lesions recognition in colposcopy images. *Scientific Reports*. 2020; 10(1)
22. Wall J, Krummel T. The digital surgeon: How big data, automation, and artificial intelligence will change surgical practice. *Journal of Pediatric Surgery*. 2020; 55: 47–50.
23. Wu A, Xue P, Abulizi G, Tuerxun D, Rezhake R, Qiao YL. Artificial intelligence in colposcopic examination: A promising tool to assist junior colposcopists. *Frontiers in Medicine*. 2023; 10.
24. Xue P, Tang C, Li L, Shen Y, Zhao Y, Chen J, Wu J, Li L, Wang W, Li Y, Cui X, Zhang S, Zhang W, Zhang X, Ma K, Zheng Y, Qian T, Tat M, Liu Z. Development and validation of an artificial intelligence system for grading colposcopic impressions and guiding b
25. Yang YC, Islam SU, Noor A, Khan S, Afsar W, Nazir S. Influential usage of big data and artificial intelligence in healthcare. *Computational & Mathematical Methods in Medicine*. 2021; 1–13.