What are the Medical Benefits and Challenges of Integrating AI into the Process for Endometriosis?

Author Name:
Taneesha Gupta

Abstract:
Artificial intelligence (AI) is transforming the diagnosis and treatment of endometriosis by improving detection accuracy and enabling the development of more personalised treatments. This paper explores the use of AI across medical imaging, clinical and genomic data, comparing its effectiveness to traditional methods such as laparoscopy, a minimally invasive surgical procedure. While early studies show promising results, limitations including data bias, small datasets, and limited clinical validation remain significant barriers. Overall, this paper suggests that AI has the potential to revolutionise gynaecological care provided the issues of data privacy and algorithmic bias are addressed to prevent a wider gap in global healthcare.

Keywords - Artificial Intelligence (AI), Endometriosis, Machine Learning (ML), Laparoscopy, Algorithmic Bias

Introduction:

Endometriosis is a chronic condition that is often misunderstood, affecting women between the ages of 15 and 44. The condition affects an estimated 10–15% of all women of reproductive age worldwide, representing a global burden of over 22 million cases. Wang et al (2024). Characterised by the growth of endometrial-like tissue outside the uterus, it results in severe pain, bloating, and fatigue, which can profoundly impact a woman's daily functioning and quality of life. The urgency of this disease lies not just in its prevalence but also in the profound systemic failure to diagnose it efficiently.

Accurate medical diagnosis is essential for delivering effective treatment, but in reality, it’s not always guaranteed. Despite being essential to healthcare, human-led diagnosis is susceptible to error. Singh et al. (2014) report that approximately 1 in 20 U.S. adults (about 12 million people) experience a diagnostic error each year. These errors can result from systemic issues such as miscommunication, time constraints, and limited access to patient information. In the case of endometriosis, the lack of a non-invasive diagnostic method along with the variability of symptoms across patients adds to the complexity of the condition (average diagnostic delay is about 7-10 years in many countries).

In response to these challenges, artificial intelligence (AI) is rapidly emerging as a transformative tool in healthcare, particularly in diagnostics. AI systems can analyse vast amounts of data, from medical imaging and clinical records to genomic datasets, at speeds and accuracies beyond human capability. According to AIPRM (n.d.), disease diagnosis is currently the most common use of AI in healthcare, with 42% of healthcare organizations already utilizing it, and another 19% planning to adopt it within the next three years. This offers the potential to finally bypass the systemic constraints that have plagued endometriosis diagnosis for decades and provide earlier, more accurate detection. However, its impact hasn’t been evenly distributed, some conditions like endometriosis still suffer from delayed diagnosis and inconsistent care. Moreover, AI in medicine isn’t a flawless fix. While it holds immense promise, it also introduces ethical and clinical challenges.

This paper explores the benefits and challenges of integrating AI into the diagnostic process for endometriosis, shedding light on how this technology transform healthcare. Before diving into the transformative potential of AI, it is essential to first establish a foundation understanding of endometriosis, including its historical context and the different forms it takes.

Background: Historical Context and Types of Endometriosis

The first ever documented case of endometriosis is often credited to John A. Sampson in the early 20th century. Sampson described the condition and proposed the 'retrograde menstruation' theory as a possible cause; a theory that involves menstrual blood flowing backwards through the fallopian tube and into the pelvic cavity, instead of out of the body. However, clinical understanding of the disease came long before that: in the late 1800's by Thomas Cullen, who gave it the name 'Adenomyoma'. Cullen's observations laid the groundwork for understanding the pathological nature of the disease, although it took decades for the medical community to acknowledge its complexity.

Endometriosis is not a single, uniform condition; it exists in multiple forms, each associated with distinct symptoms, severity, and potential complications. There are mainly five types of endometriosis.

1. Superficial Peritoneal Endometriosis: This condition includes a tissue similar to the uterus lining growing on the peritoneum (the membrane lining the abdominal cavity). Affecting approximately 80% of patients (ISUOG, n.d.), this type is the most common form of endometriosis. Symptoms often include abdominal bloating, severe pelvic pain and mild to moderate discomfort during menstruation or ovulation.

2. Deep Infiltrating Endometriosis: A more severe form where the endometrial tissue grows 5mm beneath the peritoneal surface; into pelvic organs, leading to immense pelvic pain, pain during intercourse (Dyspareunia), bowel or urinary symptoms, and chronic fatigue.

3. Ovarian Endometriosis: Defined by the presence of on or within the ovaries which are filled with old menstrual blood. Symptoms include severe pelvic pain, bloating, and infertility due to disrupted ovulation.
These cysts might also rupture, leading to severe, sudden pain and risk of infection, often requiring surgical intervention.

4. Extra-pelvic Endometriosis: A rare presentation occurring outside the pelvic cavity, including sites such as the lungs or abdominal wall. Clinical features vary depending on the location.

5. Adenomyosis: Sometimes classified separately from endometriosis, adenomyosis occurs when tissue grows inside the uterus (myometrium, the muscular wall of the uterus) making it enlarged which can lead to heavier periods (menorrhagia) and very painful cramps (dysmenorrhea).

With the complexities of endometriosis and its long diagnostic journey established, the next section focuses on the emerging technologies poised to disrupt the current process: Artificial intelligence in medical diagnosis.

AI in Medical Diagnosis:
The term Artificial Intelligence (AI) was first coined by John McCarthy in 1956 during the Dartmouth Summer Research Project. Since then, it experienced a very slow advancement due to a lack of funds and adequate equipment in the mid 1980's. However, the early 2000s marked a period of rapid growth in AI research and development.

AI was first used in the healthcare industry in the early 1970's: Stanford researchers produced MYCIN, an AI program that helped identify blood treatment infections. While MYCIN never entered clinical use, it laid the groundwork for later medical AI systems. The last decade has witnessed a rapid increase of the application of AI in healthcare as the Food and Drug Administration (FDA) approved more than 500 AI-powered devices for various medical uses by 2022, 80% of which support diagnostic imaging.

Applications of AI in Healthcare:
Diagnostic and Medical Imaging: AI is helping doctors analyse medical scans like X-rays, CT scans and MRIs in order to detect patterns that might have been missed by the human eye. According to Time, A U.S./U.K. study reported AI models successfully reduced false-positive (a healthy person is mistakenly diagnosed with cancer) rates by up to 6% and false-negative (a sick person is mistakenly told they are healthy) rates by up to 9.4% in breast cancer detection. This not only increases chances of survival but also proves that AI integrated with a team of specialized doctors can significantly improve patient outcomes and diagnostic accuracy.

Drug discovery and development: AI has helped accelerate the tedious process of drug development by analysing biological and chemical data to identify potential drug candidates, predict toxicity and optimise clinical trials. During the COVID-19 pandemic, AI was used to screen thousands of compounds in record time, significantly shortening the discovery stages.

Personalised Medicine: Every human body is different and therefore some treatment plans that work great for one may not help another person suffering from the same condition. Personalised medicine includes using an individual's genetic data and health records in order to help diagnose and treat a disease. This is particularly valuable in oncology, where algorithms can predict tumour responses to therapies, improving treatment success rates.

Remote monitoring and wearable: AI-driven wearable devices and mobile health apps would allow continuous, non-invasive monitoring of patient health outside of clinic settings which would make it easier to track recovery and spot chances of relapsing. Moreover, smartwatches can track physiological data (data about the body's normal functions) such as heart rate, sleep, and menstrual cycles and by analysing it in real time, would be able to flag any abnormalities that might require medical attention. For endometriosis specifically, this technology holds significant promise for early detection and tracking of disease progression, potentially supplementing the diagnostic data currently obtained through invasive procedures like laparoscopy.

While AI has been applied successfully in multiple areas of healthcare, its use in gynaecology, specifically research on endometriosis remains limited. This presents an opportunity to explore how AI could transform the diagnosis and management of a condition that currently suffers from long delays, high misdiagnosis rates, and invasive procedures for diagnosis.

To understand the current state and the future potential of AI, the following literature review will examine existing studies on AI in medicine, focusing on its progress in gynaecology and the diagnosis of endometriosis.

Literature Review:

AI in Medicine

Since its initial launch into the industry in the 1970s, artificial intelligence (AI) has grown into one of the most promising tools in healthcare. Although earlier systems like MYCIN showcased potential for assisting in diagnosis, they lacked the data or processing capacity required to be used in clinical settings. Those developments in machine learning have now transformed AI into a practical diagnostic tool. McGenity et al. (2023) analysed 48 studies of digital pathology encompassing more than 152,000 images, using deep learning convolutional models (a type of machine learning that uses layers of artificial 'neurons' to analyse data, similar to how the human brain works) to identify malignant tissue structures. They reported a pooled sensitivity of 96.3 percent and specificity of 93.3 percent and reflecting the ability of AI to accurately recognize patterns. According to Hanis et al. (2022), machine learning methods show excellenct performance in mammography diagnosis of breast cancer as Convolutional Neural Network (CNN) reached 0.974 for sensitivity, specificity and AUC (Area Under the Curve, a way to measure the model's overall accuracy). While the integration of AI displays promising results, researchers note ongoing challenges in generalisability, as most AI models are trained on limited biological datasets, risking biased outputs that fail in real-world diverse populations.

AI in Gynaecology

AI is widely used in fields like oncology and cardiology, but its use in gynaecology is still limited. Khudhur et al. (2025) found that using reinforcement learning and vision-based AI in robotic surgeries improved precision and safety. In oncology and fertility, machine learning models have been tested for predicting complications and selecting embryos, showing better accuracy and outcomes. Patel et al. (2024) noted that while AI can predict risks such as preeclampsia, most systems are still in early stages and not part of daily clinical practice. Overall, even though AI performs well in research, doctors are cautious about using it due to concerns over reliability, workflow changes, and data privacy. As a result, gynaecology still lags behind other specialties in adopting AI. These barriers illustrate that gynaecology, while promising for AI, still lags behind other specialties in its implementation.

AI in Endometriosis

Endometriosis is one of the hardest gynaecological disorders to diagnose and research on the disease remains limited. Moro et al. (2025) reviewed AI applications in benign gynaecological diseases, reporting that ultrasound-based models could identify endometriotic lesions with profound precision.

Bendifallah et al. (2022) further demonstrated a machine learning model using 16 clinical variables, which achieved a sensitivity of 95–98% and specificity of around 80%, suggesting strçong potential for non-invasive diagnosis.

More advanced projects, such as the IMAGENDO study at the University of Adelaide, combine MRI and ultrasound through machine learning in hopes of providing a less invasive alternative to laparoscopy.

Likewise, deep learning applied to laparoscopic videos (Luz & Lima, 2025) has shown potential to assist surgeons in real-time lesion detection. Together, these innovations signal that AI could meaningfully transform how endometriosis is diagnosed and managed, paving the way for faster and less invasive clinical pathways.

Gaps and Challenges:

Despite promising developments, considerable gaps remain in current research on AI applications in endometriosis within the real world. The majority of studies base their findings on local data, often a single hospital, which restricts the external validity and evokes concerns regarding their generalisability (Luz & Lima, 2025). Furthermore, the reliability of these models in real life is questionable as few of them have been tested in clinical trials or in multiple clinical centres.

One of the challenges is the unavailability of standardized diagnostic imaging and labelling methods, which creates inconsistency and makes it difficult to build stable, generalisable algorithms (Moro et al., 2025). Ethical and methodological problems further complicate AI's integration. Deep learning models are often "black boxes," providing no insight into how they make decisions (meaning doctors don't know *why* the AI reached a certain conclusion), which challenges clinical accountability. The issues of privacy, permission, and data security are especially significant because reproductive health data are extremely sensitive (Luz and Lima, 2025).

Finally, persistent underfunding of endometriosis research continues to hinder advancements, causing AI innovation in this field to lag behind better-resourced fields like breast cancer. When combined, these challenges show that although AI has enormous potential to revolutionise the diagnosis and treatment of endometriosis, its clinical integration is still fragmented and lacking. To systematically address these critical gaps and evaluate the current body of fragmented research, the subsequent section outlines the specific methodology and selection criteria employed in this literature analysis.

Methodology:

The purpose of this research study is to conduct a comparative literature analysis, examining recent discoveries within the integration of AI in diagnostics and treatment for endometriosis. The goal of this paper is to evaluate the impact of endometriosis on women's health and quality of life, and to examine how emerging technologies, particularly artificial intelligence, can help reduce misdiagnosis and contribute to more effective management, ultimately improving patient outcomes and daily living.

Selection Criteria:

Publication Date - Only papers published within 2010 and 2025 were considered in order to capture modern AI development and evolution from its first application in the late 1900s.
Relevance - Papers focusing on AI applications in healthcare, particularly their impact on diagnostic accuracy, treatment outcomes, and associated challenges, were selected
Study design - Researched studies included clinical trials, observational studies, systematic reviews, and studies that assessed AI algorithms in diagnostic imaging, predictive modelling, or treatment planning.

Population Focus:

Clinical studies - Human studies across all age groups experiencing gynaecological conditions, particularly endometriosis.
AI model studies - Studies incorporating machine learning or AI-driven predictive or diagnostic tools relevant to gynaecology or healthcare as a whole.

Analytical Framework:
Diagnostic outcomes - Accuracy, specificity and early detection rates of AI-assisted diagnostics.
Treatment outcomes - Improvement in patient conditions, success rates of the provided treatments and accuracy of personalised therapy planning.
Challenges and benefits - Technical limitations, ethical considerations, cost effectiveness, data protection and security.
Emerging technologies - Emphasis on innovative AI approaches, imaging for AI lesion detection, and predictive modelling on disease progression.

Exclusion Criteria:
Papers lacking peer review or published before 2010
Studies not directly evaluating AI applications in healthcare or gynaecology
Articles not focused on clinical or diagnostic outcomes, rather, focused solely on technical AI development
Studies with insufficient sample size or unclear methodology

Having completed the methodology, the discussion section now begins, evaluating the synthesized research to determine whether AI's current performance meets its significant potential for transforming endometriosis diagnosis and treatment.

Discussion:

This review set out to evaluate how the integration of artificial intelligence influences diagnostic accuracy and treatment efficacy in endometriosis. The research presented demonstrates that AI integration produces measurable improvements in diagnostic accuracy and treatment planning. However, transitioning these models from theoretical frameworks to clinical environments requires addressing distinct methodological bottlenecks.

Diagnostic and Clinical Implications
Studies consistently show that AI-enhanced imaging (Moro et al., 2025; Liu et al., 2022) achieves higher sensitivity and specificity compared to traditional clinician-only approaches, making it easier to detect abnormalities in test scans and potentially accelerating the diagnostic process. By catching subtle abnormalities in ultrasound and MRI scans that might elude human perception, AI acts as a vital clinical safety net that could decrease the historical 7-to-10-year diagnostic delay that plagues endometriosis patients.
However, the lower pooled specificity (around 80%) observed in some endometriosis models introduces a clinical risk of false positives (Moro et al., 2025). In practice, a false positive means a healthy person is mistakenly diagnosed with a condition they do not have. For endometriosis, this could lead to overdiagnosis, causing unnecessary patient anxiety and potentially leading to unwarranted, invasive diagnostic surgeries like laparoscopy, ironically defeating the purpose of non-invasive AI tools.
Furthermore, while predictive machine learning models show promise in triaging patients by recurrence risk or predicting infertility (Bendifallah et al., 2022; Cao et al., 2025), their utility is hindered by the "black box" nature of deep learning. In gynaecology, where symptomatology is highly subjective and variable, physicians cannot blindly rely on an algorithmic output without understanding its underlying clinical logic. Therefore, the immediate future of this technology lies not in autonomous AI, but in human-AI collaborative models like HAICOMM (Wang et al., 2024), which preserve essential physician oversight and combine human knowledge alongside AI for interpretation.

Structural Barriers: Data Scarcity and Funding Realities

The primary obstacle to deploying reliable AI for endometriosis is the lack of diverse, large-scale datasets. As evidenced by the compiled literature, there is a stark contrast between sample sizes in general digital pathology - such as the 152,000 images studied by McGenity et al. (2023) - and gynaecology-specific datasets, which remain restricted to small-scale or single-centre cohorts (Moro et al., 2025; Luz & Lima, 2025). This disparity is deeply rooted in the historical underfunding of endometriosis research, causing AI innovation in this field to lag behind better-resourced fields like breast cancer (Hanis et al., 2022).

When an algorithm is trained on a homogeneous, local patient population or a single hospital's dataset, it severely restricts the AI model's external validity and generalizability (Luz & Lima, 2025). If deployed in the real world, such models may not perform reliably across different patient populations, ethnic groups, healthcare systems, or imaging machines. This raises severe concerns regarding algorithmic bias, where the AI's results are structurally unfair because of the limited data it was trained on. Without addressing this data scarcity, the clinical rollout of AI risks heightening global healthcare inequalities rather than reducing them, especially in lower-income or less-developed areas where access to AI-based care is already limited (Dungate et al., 2024; Luz & Lima, 2025).

A Framework for Implementation

To move past the current experimental standstill and achieve equitable, stable algorithms, the medical community must focus on two key structural strategies:

1. Establish Multicentre Collaborative Data Initiatives: Global collaboration between hospitals, research centres, and specialty groups is essential for pooling diverse clinical and imaging data. This diversity is the only mathematical defence against algorithmic bias, ensuring AI models work reliably across diverse patient populations and enhancing external validity.

2. Develop Large, Anonymized, and Standardized Datasets: Researchers must work with regulatory bodies and privacy organizations to create a large-scale repository of anonymized patient data, protecting highly sensitive reproductive health records. Crucially, establishing standardized diagnostic imaging and labelling methods within this repository is vital for building stable, generalisable algorithms (Moro et al., 2025). This requires active consultation with globally recognized health organizations, such as the World Health Organization (WHO) and the International Society of Ultrasound in Obstetrics and Gynaecology (ISUOG), to promote universal guidelines and ensure consistent data acquisition protocols.

Results:

Survey of the chosen literature reports show promising developments in the integration of AI in the diagnosis and treatment of endometriosis. While the current development seems to be halted in the testing phase, studies predict an increase in AI efficacy within the next decade. The key findings within the studied papers are listed in *table 1*.

Reference	AI method/model	Application Area	Dataset/ sample size	Key findings	Limitations
McGenity et al. (2023)	Digital pathology / CNN	Pattern recognition & diagnostic imaging	152,000 histology images (48 studies)	Sensitivity 96.3%, specificity 93.3%	Lack of dataset diversity
Hanis et al. (2022)	CNN for mammography	Diagnostic imaging (breast cancer)	Mammo gram dataset	Sensitivity, specificity, AUC = 0.974	Limited to breast imaging, not gynaecological
Khudhur et al. (2025)	Robotic-assisted AI systems	Gynaecologic oncology & fertility	Multiple surgical datasets	Improved surgical precision, faster outcomes	Early-stage integration, limited adoption
Patel et al. (2024)	Predictive ML model	Preeclampsia prediction	Clinical data (multi-centre)	High predictive accuracy	Lacks standardization for gynaecology
Moro et al. (2025)	Ultrasound & MRI-based ML	Endometriosis lesion detection	Small-scale ultrasoun d datasets	Sensitivity 95–98%, specificity ~80%	Retrospective study, no prospective validation
Luz & Lima (2025)	Deep learning on laparoscopic videos	Intraoperative lesion detection	2,000+ frames	Accurate lesion identificati on intra-op	Single-centre study, no external validation

Table 1

As evidenced by the compiled data, imaging-based models consistently demonstrate high diagnostic sensitivity, ranging from 95% to 98% in ultrasound-based machine learning. However, specificity rates fluctuate, with some models averaging around 80%. Furthermore, the data reveals a sharp contrast between sample sizes in general digital pathology (e.g., 152,000 images studied by McGenity et al.) and gynaecology-specific datasets, which remain restricted to small-scale or single-centre cohorts.

Final Conclusion:

This paper demonstrates that artificial intelligence holds transformative potential for endometriosis care, offering a data-driven solution to centuries of diagnostic delays and subjective clinical blind spots. By augmenting the accuracy of non-invasive imaging, AI represents a paradigm shift toward early, personalised intervention.

However, the technology has outpaced the medical infrastructure required to support it. The clinical integration of AI remains fragmented, held back by localized datasets, systemic underfunding, and a lack of standardized imaging protocols.

Ultimately, AI cannot fix a medical bias if it is trained on biased or limited data. Bridging the gap between experimental software and a trusted standard of care requires deliberate, global collaboration between clinicians, researchers, and policymakers to build ethical, equitable data repositories. Only when these systemic structural barriers are addressed can AI fulfil its promise: providing millions of women worldwide with a faster, safer, and deeply personalised path to healing.

References:

AIPRM. (n.d.). *AI in healthcare statistics*. https://www.aiprm.com/ai-in-healthcare-statistics/

Bendifallah, S., Ghoneim, T., Alsaqer, A., et al. (2022). Predicting infertility in women with endometriosis: A machine learning approach. *Journal of Clinical Medicine*, *11*(13), 3698. https://doi.org/10.3390/jcm11133698

Cao, S., Li, X., Zheng, X., & Liu, Y. (2025). Assessing the utility of artificial intelligence in endometriosis: Promises and pitfalls. *BMC Medical Imaging*, *24*, 153. https://pubmed.ncbi.nlm.nih.gov/38686828/

Dungate, B., Tucker, D., Goodwin, E., & Yong, P. (2024). The use of artificial intelligence in the diagnosis and management of endometriosis: A systematic review. *International Journal of Environmental Research and Public Health*, *21*(3), 268. https://doi.org/10.3390/ijerph21030268

Hanis, B., Noman, M., Jamil, A., & Javed, M. U. (2022). Diagnosis of breast cancer using CNN mammography. *International Journal of Biological and Biomedical Engineering*, *16*, 289–293. https://pubmed.ncbi.nlm.nih.gov/35885548/

International Society of Ultrasound in Obstetrics and Gynecology. (n.d.). *Superficial endometriosis*. https://www.isuog.org/clinical-resources/patient-information-series/patient-information-gynecological-conditions/superficial-endometriosis.html

Khudhur, A. J., Ismael, M. N., Al-Janabi, N. O. A., & Al-Janabi, F. A. (2025). Using AI in gynaecology. *International Journal of Research in Medical Sciences*, *11*(12), 4858–4863. https://pubmed.ncbi.nlm.nih.gov/39156405/

Liu, H., Li, X., Wang, Q., Li, W., & Chen, J. (2022). Deep learning-based model for accurate diagnosis of endometriosis using transvaginal ultrasound images. *Biomedical Engineering Online*, *21*(1), 1–13. https://doi.org/10.1186/s12938-022-01007-z

Luz, I., & Lima, M. J. (2025). A review of AI-driven digital innovations for endometriosis diagnosis. *Applied Sciences*, *15*(1), 108. https://link.springer.com/article/10.1007/s44326-025-00061-2

McGenity, A., Cheong, A., Braidwood, M., Maclean, A., & Cheong, J. L. (2023). A systematic review and meta-analysis of artificial intelligence diagnostic accuracy in prostate cancer histology identification and grading. *European Urology Oncology*, *6*(5), 473–481. https://pubmed.ncbi.nlm.nih.gov/37185992/

Moro, F., Tola, V., Monti, E., et al. (2025). Artificial intelligence in obstetrics and gynaecology: Current applications and future perspectives. *Obstetrics Journals*, *7*(1), 1–10. https://www.obstetricsjournals.com/archives/2025/vol7issue1/PartA/7-1-3-449.pdf

National Centre for Biotechnology Information. (2015). *Overview of diagnostic error*. https://www.ncbi.nlm.nih.gov/books/NBK338594/

Patel, R., Sharma, G., & Singh, R. (2024). Applications of deep learning models in laparoscopy for gynaecology. *Applied Sciences*, *14*(10), 1–18. https://www.mdpi.com/1648-9144/61/8/1460

Singh, H., Meyer, A. N., & Thomas, E. J. (2014). The incidence of diagnostic error in medicine. *BMJ Quality & Safety*, *23*(5), 415–421. https://pmc.ncbi.nlm.nih.gov/articles/PMC3786666/

Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, *25*(1), 44–56. https://pubmed.ncbi.nlm.nih.gov/30617339/

Wang, R., Li, Z., Sun, C., et al. (2024). Human-AI collaborative multi-modal multi-rater learning for endometriosis diagnosis. *arXiv*. https://arxiv.org/abs/2409.02046/

World Health Organization. (2023). *Endometriosis*. https://www.who.int/news-room/fact-sheets/detail/endometriosis