REVIEW ARTICLE

Artificial Intelligence in Chronic Pelvic Inflammatory Disease Management: A Comprehensive Review of Integrated Diagnostic Frameworks and Adaptive Therapeutic Systems

Jiayi Yang[1] **and Qian Hong**[2,*]

[1] School of Engineering, Hong Kong University of Science and Technology, Hong Kong 999077, China
[2] Department of Obstetrics and Gynecology, Jiangxi Fifth People's Hospital, Nanchang 330046, China

Abstract

Chronic Pelvic Inflammatory Disease (CPID) poses significant challenges to women's health, necessitating advanced management strategies. This paper provides a comprehensive review of artificial intelligence (AI)-based health management techniques for PID, focusing on their potential to enhance diagnosis, treatment personalization, and long-term monitoring. By synthesizing Bayesian probabilistic frameworks with ensemble Machine Learning architectures, we systematically evaluate AI-driven solutions for PID pathophysiology analysis, therapeutic efficacy prediction, and patient-specific intervention planning. These approaches collectively enhance diagnostic precision while addressing key challenges in therapeutic personalization and longitudinal care coordination. This review significantly advances intelligent PID care by resolving fundamental challenges in heterogeneous data integration, algorithmic transparency, and cross-institutional collaboration, ultimately offering a scalable blueprint for AI-powered gynecological health systems.

Keywords: PID, bayesian network, machine learning.

1 Introduction

Chronic Pelvic Inflammatory Disease (CPID, also abbreviated as PID), encompasses a variety of infectious processes that can harm the endometrium, fallopian tubes, ovaries, and pelvic peritoneum. While most cases of PID are caused by sexually transmitted infections (STIs), bacteria linked to bacterial vaginosis (BV) are also known contributors. Around 15% of untreated chlamydial infections advance to PID, with the rate potentially being higher in cases of gonococcal infections [1, 2]. Delayed diagnosis increases the risk of inflammatory complications such as infertility, ectopic pregnancy, and chronic pelvic pain [3].

The diagnosis of PID involves several clinical criteria, including an oral temperature above 38.3°C, abnormal purulent cervical discharge or cervical tenderness, and

Academic Editor:
Thinagaran Perumal

Submitted: 27 May 2025
Accepted: 27 June 2025
Published: 17 August 2025

Vol. 2, **No.** 3, 2025.
10.62762/TIS.2025.511235

***Corresponding author:**
Qian Hong
qian.h@163.com

Citation
Yang, J., & Hong, Q. (2025). Artificial Intelligence in Chronic Pelvic Inflammatory Disease Management: A Comprehensive Review of Integrated Diagnostic Frameworks and Adaptive Therapeutic Systems. *ICCK Transactions on Intelligent Systematics*, 2(3), 169–189.

the presence of abundant white blood cells in vaginal fluid microscopy. Elevated erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP) levels indicate inflammation. Laboratory tests may confirm cervical infections with Neisseria gonorrhoeae or Chlamydia trachomatis. Nucleic acid amplification tests (NAAT) are useful for detecting these bacteria, while saline microscopy can identify Trichomonas vaginalis or BV. Imaging, such as transvaginal sonography or computed tomography (CT), may be required to identify abscesses, and laparoscopy could be used for confirmation in unclear cases.

PID is often treated with empiric antibiotic therapy, using antibiotics that cover likely pathogens without waiting for culture results. A common regimen includes doxycycline combined with either ceftriaxone or azithromycin. For severe cases or when abscesses are present, inpatient treatment with broader-spectrum antibiotics like clindamycin and gentamicin may be necessary. Clinical improvement should be observed within 72 hours. After initial recovery, follow-up care of PID is extremely crucial, which can help reduce the probability of subsequent recurrence of the disease. Follow-up care involves treating sexual partners and advising abstention from intercourse until therapy is completed. Routine screening for Chlamydia trachomatis is recommended for women under 35 or those at high risk. Prevention focuses on reducing the risk of infection spread, emphasizing patient education on the importance of screening, treatment, and prophylaxis for STIs [4–6].

PID significantly affects women's health worldwide. According to the survey from CDC [4], PID can lead to severe reproductive health issues, including chronic pelvic pain, ectopic pregnancy, and infertility. The long-term impact is particularly concerning, with 1 in 8 women diagnosed with PID experiencing difficulties in getting pregnant. This underscores the critical health and societal implications of this condition.

Moreover, PID often stems from untreated STIs like Neisseria gonorrhoeae or Chlamydia trachomatis. The rise of antibiotic-resistant strains further complicates treatment, increasing the urgency for innovative approaches to diagnosis and management. Given its asymptomatic nature in many cases, PID is frequently underdiagnosed, delaying treatment and worsening outcomes [4]. These challenges highlight the need for advanced, AI-based solutions in PID management, with the potential to significantly improve PID diagnosis and treatment by addressing key clinical challenges, including symptom variability, delayed diagnoses, and the integration of multimodal data [7, 8].

Based on the recent applications of AI-based PID management and treatment algorithms mentioned above, it can be observed that the integration of AI in the management of chronic PID has shown significant promise.

Nowadays, PID remains a formidable clinical challenge, with its insidious progression and heterogeneous presentation complicating timely diagnosis and effective intervention. Current management strategies, largely reliant on static diagnostic criteria and empirical antibiotic protocols, struggle to address three persistent barriers: the subtlety of early-stage symptoms, rising antimicrobial resistance patterns, and unpredictable transitions to chronic sequelae. This review catalyzes a paradigm shift by systematically evaluating how artificial intelligence (AI) redefines PID care through three groundbreaking dimensions.

First, we demonstrate the unprecedented capability of hybrid AI architectures in synthesizing fragmented clinical data: quantifying diagnostic uncertainties through Bayesian probabilistic reasoning while leveraging Machine Learning to decode complex inflammatory signatures from imaging and biomarker trajectories. Second, the paper introduces adaptive therapeutic systems that dynamically optimize treatment regimens using real-time feedback loops, integrating pharmacological responses with individual immune profiles. Third, our analysis reveals how cascading resilience models, originally designed for critical infrastructure protection, provide novel insights into predicting and mitigating PID progression risks through immunological network failure simulations.

Beyond technological innovation, this work establishes critical benchmarks for ethical AI implementation in women's health, addressing data sovereignty concerns through decentralized learning frameworks while enhancing clinician trust via explainable decision pathways. By bridging computational models with clinical workflows, the proposed solutions offer a transformative roadmap for transitioning from reactive PID management to proactive, precision-driven care - a vital step toward reducing the global burden of preventable reproductive morbidity.

2 Bayesian Network

The diagnostic criteria for PID present substantial complexity due to the multifaceted interactions between clinical symptoms, laboratory findings, and heterogeneous patient factors. Bayesian Networks, a cornerstone of artificial intelligence methodologies, have emerged as a critical tool for navigating this intricacy. A Bayesian Network is a probabilistic graphical model that encodes causal and statistical relationships among variables through a **Directed Acyclic Graph (DAG)**. In this framework, nodes represent random variables (e.g., clinical features, disease states, or biomarkers), while directed edges signify conditional dependencies between these variables. Crucially, the absence of cycles in the graph ensures that no variable can indirectly influence itself, preserving logical consistency in probabilistic reasoning.

At its core, a Bayesian Network operates on the principles of **conditional probability** and **modular factorization**. The joint probability distribution of all variables in the network is decomposed into a product of conditional probabilities, each dependent only on its direct parent nodes. This is formalized by the equation:

$$P(X_1, \ldots, X_n) = \prod_{i=1}^{n} P(X_i|\text{parents}(X_i)) \quad (1)$$

where parents(X_i) denotes the set of nodes directly influencing X_i. For instance, if symptom S_1 depends on disease D, and biomarker B_1 depends on both D and S_1, the joint probability $P(D, S_1, B_1)$ would factorize as $P(D) \cdot P(S_1|D) \cdot P(B_1|D, S_1)$. This factorization reduces computational complexity by leveraging **conditional independence** assumptions: two variables are conditionally independent if their probabilistic relationship is entirely mediated by a third variable (or set of variables). For example, if two symptoms S_1 and S_2 are caused solely by disease D, they become independent once D is known, simplifying the model to $P(S_1, S_2|D) = P(S_1|D) \cdot P(S_2|D)$.

Equation (1) captures the principle of factorizing joint probabilities in Bayesian Networks, where each clinical variable (e.g., pelvic pain, CRP level, cervical tenderness) depends only on its direct causal or correlational parents. In a PID diagnosis context, this means the network does not need to consider all variables simultaneously, but rather in a modular and interpretable way. This helps simplify inference even when dealing with high-dimensional, incomplete clinical datasets by leveraging conditional independence. For instance, if fever and vaginal discharge are both influenced by PID, the model can treat them independently once PID is accounted for, significantly reducing complexity in decision-making [9].

The strength of Bayesian Networks lies in their ability to perform **bidirectional inference**. Using **Bayes' theorem**, they update probabilities dynamically as new evidence is incorporated in equation:

$$P(H|E) = \frac{P(E|H)P(H)}{P(E)} \quad (2)$$

where H represents a hypothesis (e.g., a disease) and E represents observed evidence (e.g., symptoms or test results). This allows clinicians to answer both **diagnostic queries** (e.g., "What is the probability of PID given pelvic pain and elevated CRP?") and **prognostic queries** (e.g., "If PID is confirmed, what is the likelihood of developing infertility?"). The network's edges, while often interpreted as causal relationships, need not strictly represent causality; they may instead encode associative or diagnostic pathways derived from clinical data. This flexibility enables the integration of diverse knowledge sources, including epidemiological studies, clinical guidelines, and real-world patient datasets.

This is Bayes' theorem in action: given observed evidence E (e.g., adnexal tenderness, fever), Equation (2) computes the posterior probability for hypothesis H (PID). For instance, a patient presenting both tenderness and elevated CRP might shift from a low prior suspicion to a high posterior probability, guiding clinicians toward timely antibiotic therapy even before confirmatory tests [9].

In medical applications, Bayesian Networks address three critical challenges: **uncertainty quantification**, **confounder adjustment**, and **missing data handling**. For PID diagnosis, where symptoms such as pelvic pain, fever, and cervical motion tenderness lack specificity, the network quantifies uncertainty by computing posterior probabilities for competing diagnoses (e.g., PID vs. endometriosis vs. urinary tract infection). Each node is associated with a **conditional probability table (CPT)** that defines the likelihood of its states given combinations of parental states. For instance, the CPT for "Fever" might specify that fever occurs in 80% of PID cases, 30% of endometriosis cases,

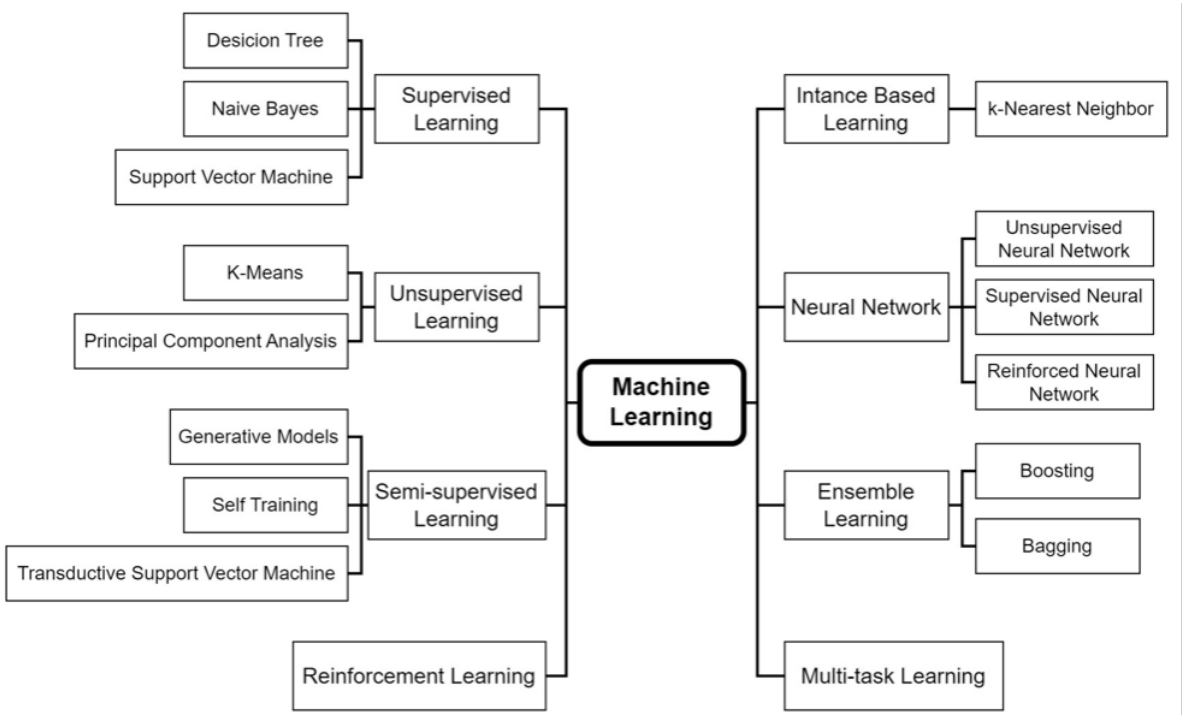


Figure 1. Classification of ML [12].

and 10% of urinary tract infections. These probabilities are refined iteratively as patient-specific data (e.g., lab results, imaging findings) is entered into the model.

Recent advancements in Bayesian Network methodologies have further enhanced their clinical utility. **Structure learning algorithms**, such as constraint-based or score-based approaches, automatically infer optimal network configurations from data, reducing reliance on expert-defined architectures. **Parameter learning techniques**, including maximum likelihood estimation and Bayesian updating, improve the accuracy of CPTs by incorporating population-level and patient-level data. Additionally, hybrid models combining Bayesian Networks with Machine Learning (e.g., deep learning for feature extraction) enable the integration of high-dimensional data, such as ultrasound images or cytokine profiles, into diagnostic frameworks.

The application of Bayesian Networks in PID diagnostics exemplifies their capacity to synthesize fragmented clinical information into a unified probabilistic model. By translating subjective clinical judgments into quantifiable probabilities, these networks mitigate diagnostic variability and enhance decision-making transparency. Future developments are poised to leverage real-time adaptive learning, multi-modal data fusion, and explainable AI interfaces, bridging the gap between theoretical models and bedside practice [10, 11].

3 Machine Learning

Machine Learning (ML) is a subfield of artificial intelligence that develops computational models capable of extracting patterns from data without explicit programming. Its applications span diverse domains including predictive analytics, medical imaging, and clinical decision support systems [12]. Within healthcare, ML algorithms are broadly categorized into supervised and unsupervised learning paradigms, with supervised learning demonstrating exceptional utility in inflammatory disease management due to its predictive accuracy and interpretability [7].

As illustrated in Figure 1, contemporary ML algorithms encompass multiple branches, each tailored to specific data characteristics and clinical requirements. Among these, three methodologies have gained prominence in PID management:

3.1 Supervised Learning

This paradigm trains models using labeled datasets containing input-output pairs, enabling precise mapping of clinical features to diagnostic outcomes. The training process involves minimizing prediction errors through iterative optimization, yielding models capable of generalizing to unseen patient data [12]. Two key supervised algorithms in PID management include:

3.1.1 Naive Bayes Classifier

Based on Bayes' theorem with conditional independence assumptions, this probabilistic model calculates posterior probabilities for disease classes using equation:

$$P(y \mid x_1, ...x_n) \propto P(y) \prod_{i=1}^{n} P(x_i \mid y) \tag{3}$$

where y represents disease classes and x_i clinical features. Despite its simplified assumptions, it achieves high computational efficiency in text-based symptom classification [7].

This computes the likelihood of a diagnosis y (e.g., PID) by assuming conditional independence of observed features. In fast-paced triage situations, this allows rapid computation of posterior probabilities based on available data (e.g., fever, leukocytosis). Despite the simplifying assumption, Naive Bayes has proven surprisingly effective in medical diagnostic tasks, providing clinicians with probabilistic confidence even when tests are unavailable [13].

3.1.2 Decision Tree

This interpretable algorithm recursively partitions feature space using entropy-based rules as shown in equation:

$$Gini(S) = 1 - \sum_{i=1}^{k} {p_i}^2 \tag{4}$$

where p_i denotes class proportions in subset S. Clinicians favor its transparent decision pathways for treatment planning, particularly when handling mixed categorical/continuous EHR data [12].

Gini impurity quantifies how well a potential symptom or test result splits patients into homogeneous PID vs. non-PID groups. A low Gini after splitting by "tenderness present" signals that most cases in that node share the same diagnosis, effectively mimicking the clinical reasoning: "if tenderness is present, then PID probability is high." This intuitive, stepwise partitioning aligns well with diagnostic checklists used in practice [13].

3.2 Multi-Task Learning

Multi-Task Learning (MTL) enhances model robustness by jointly optimizing related clinical prediction tasks through shared latent representations. This framework is mathematically expressed as equation:

$$\min_{W} \sum_{t=1}^{T} \ell_t(\omega_t) + \lambda \parallel W \parallel_* \tag{5}$$

where $W = [w_1, \ldots, w_T]$ parameterizes T tasks, and $\|W\|_*$ denotes the nuclear norm for knowledge transfer [12]. In PID management, MTL effectively addresses data scarcity in rare disease subtypes by leveraging shared pathophysiological patterns. $\ell_t(\omega_t)$ represents task-specific losses—for example, predicting PID severity and treatment response concurrently. The nuclear-norm regularization $\|W\|_*$ encourages sharing patterns across tasks. Clinically, this enables learning from features like inflammatory markers to improve multiple prediction goals simultaneously, enhancing model performance in the face of limited PID data [14].

3.3 Artificial Neural Networks

Artificial Neural Networks (ANNs) employ hierarchical nonlinear transformations to decode complex medical patterns:

$$h^{(l)} = \sigma(W^{(l)} h^{(l-1)} + b^{(l)}) \tag{6}$$

where σ denotes activation functions (e.g., ReLU), $W^{(l)}$ weight matrices, and $b(l)$ biases at layer l. Deep learning variants demonstrate superior performance in medical image analysis for PID severity staging, achieving AUC scores exceeding 0.92 in recent trials [7].

This formula describes how an artificial neural network processes input data (e.g., ultrasound pixel values, CRP levels) through successive hidden layers to generate higher-level diagnostic features. Although hidden-layer representations are abstract, visualization techniques (e.g., saliency mapping) can highlight which clinical or imaging features most influenced a PID diagnosis—providing transparency and building clinician trust in oncology-like diagnostic support systems [15].

The ML approaches employed in PID management each demonstrate unique capabilities and limitations shaped by their algorithmic foundations, and Table 1 illustrates the differences between them. Supervised learning methods, particularly naive Bayes classifiers and decision trees, provide clinically interpretable frameworks for symptom classification and treatment

logic. While naive Bayes achieves computational efficiency through probabilistic assumptions, its simplification of biomarker interactions may overlook critical inflammatory pathways. Decision trees offer transparent decision boundaries that align with clinical workflows, yet risk overfitting to institution-specific protocols. These methods collectively address diagnostic prioritization but require meticulous data labeling and may struggle with complex feature interdependencies inherent in chronic PID pathophysiology.

MTL introduces a paradigm shift by leveraging shared pathophysiological patterns across PID subtypes, enhancing model stability for rare disease variants. By optimizing multiple clinical tasks through nuclear norm regularization, MTL reduces data dependency compared to single-task models. However, its effectiveness hinges on predefined task relationships, which may inadvertently obscure subtype-specific inflammatory signatures. ANNs excel in decoding nonlinear patterns from medical imaging data, demonstrating particular utility in pelvic lesion quantification and severity staging. While their hierarchical architectures achieve state-of-the-art performance in image-based assessments, the requirement for large annotated datasets and inherent opacity of feature transformations pose challenges for clinical validation and ethical auditing.

Within the proposed PID diagnostic framework, Bayesian Networks (BNs) serve as foundational probabilistic models that encode domain knowledge and manage uncertainty inherent in clinical data. These graphical models, typically structured as directed acyclic graphs, associate symptoms (e.g., pelvic tenderness, fever) with latent disease states using conditional probability tables [11]. Bayesian inference enables clinicians to update disease probabilities in light of new evidence—such as ultrasound findings or laboratory results—making them well-suited for structured decision support in PID, even when data are incomplete or noisy. Hybrid implementations often combine BN structure learning with expert input and parameter estimation via maximum-likelihood or expectation-maximization algorithms—approaches shown effective in multi-disease CDSS prototypes.

Complementing BNs, the framework integrates ML classifiers: ranging from Random Forests to gradient-boosted ensembles and support vector machines to detect complex patterns from multimodal inputs, including imaging features and patient metadata. Such models are trained on curated PID datasets, leveraging feature importance metrics (e.g., SHAP/FI) to enhance interpretability and precisely focus on critical predictive variables. These ML classifiers excel in stratifying disease severity and predicting treatment response, with reported AUCs often exceeding 75% in similar chronic inflammatory contexts [16].

At the apex of the model stack, ANNs, especially Bayesian neural networks (BNNs) and convolutional neural networks (CNNs), handle high-dimensional inputs such as pelvic ultrasound frames and raw EMR text. BNNs are employed to quantify uncertainty in their outputs, offering probabilistic predictions with confidence intervals critical in clinical decision-making [17]. Meanwhile, CNNs process imaging data and are often paired with ensemble BNs to improve both performance and interpretability. Such dual-stack architectures maintain diagnostic accuracy while ensuring that outputs remain explainable and clinically actionable.

Overall system efficiency is achieved through layer-wise modular integration: fast, lightweight ML models handle initial screening, while more compute-intensive BNNs are invoked only when necessary. Approximate inference algorithms and truncated BN structures help to bound computational costs [18, 19]. Real-world performance metrics demonstrate that frameworks like BayCANN, which replace complex simulators with ANNs metamodels, achieve significant speedups (e.g. 5 times faster training) without sacrificing accuracy [20]. Similarly, edge-optimized neural architectures (e.g., LogNNet) deliver near-clinical accuracy ($\sim$90%) using only a few kilobytes of memory [21].

As a result, the combined framework is not only computationally viable but also compatible with real-time clinical settings. BNs inference runs in milliseconds on modest hardware; shallow ML classifiers and lightweight CNNs can process imaging frames at tens of FPS. With attention-based pruning or early-exit strategies, latency can be reduced further—supporting continuous PID monitoring via ultrasound or telemedicine platforms. Overall, this hybrid model balances accuracy, interpretability, and computational efficiency, making it well-suited for deployment in real-world clinical workflows.

Table 1. Comparative analysis of ML approaches in PID.

Algorithm	Clinical Utility	Technical Constraints
Naive Bayes	Rapid symptom triage	Oversimplifies biomarker interactions
Decision Tree	Protocol-transparent decisions	Sensitive to institutional bias
MTL	Subtype-adaptive predictions	Requires manual task engineering
ANNs	Image-based severity staging	Demands large annotated datasets

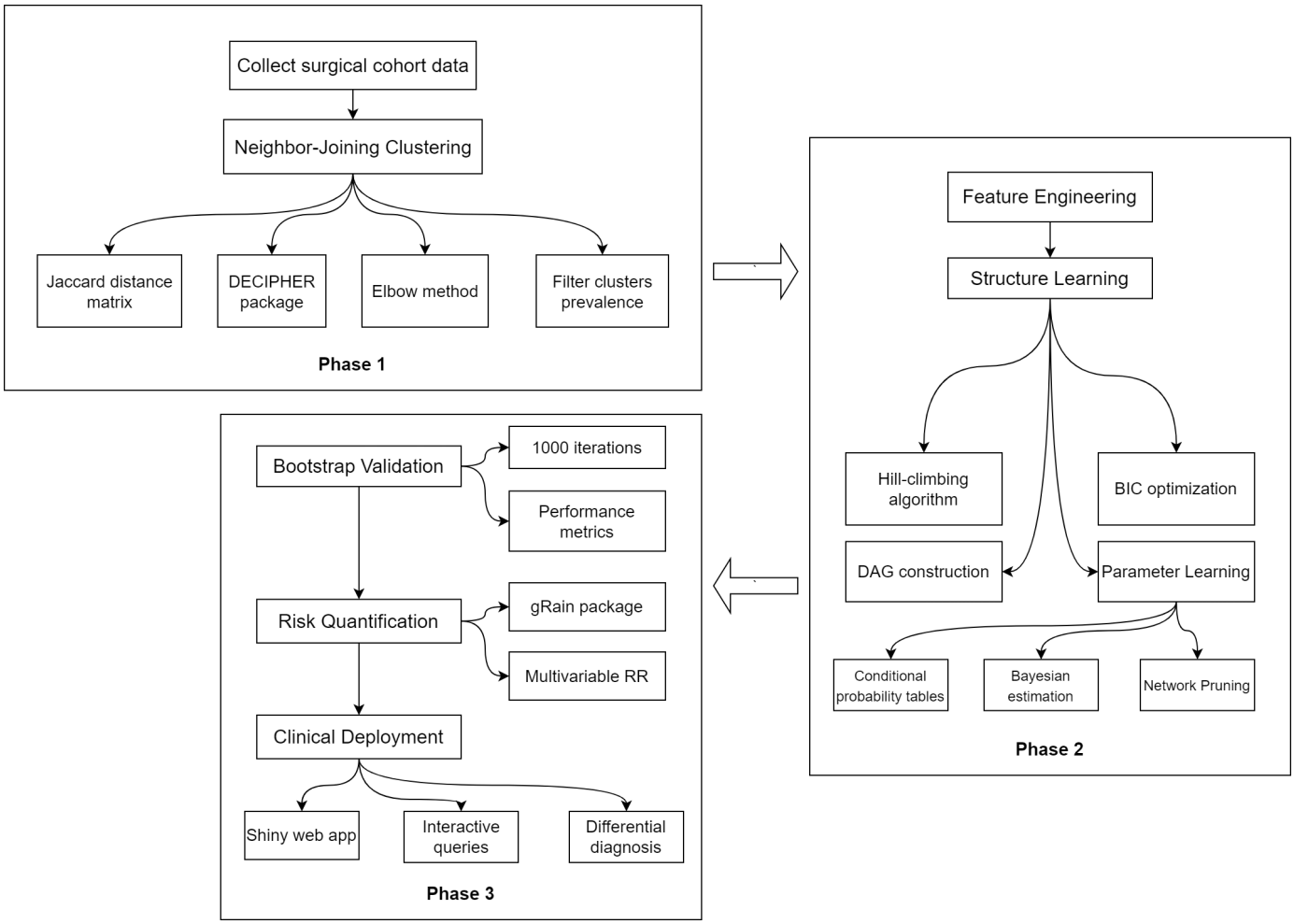


Figure 2. Research flow of [22].

4 Applications

4.1 Bayesian Network Based Approach

In recent years, Bayesian estimation methods have gained prominence in medical research due to their capacity to handle uncertainty, integrate prior knowledge, and model complex variable relationships. This methodological framework, grounded in probabilistic reasoning, effectively synthesizes clinical observational data with domain expertise, demonstrating unique advantages in disease diagnosis, treatment evaluation, and pathological mechanism analysis. Its core value lies in visualizing conditional dependencies among variables through probabilistic graphical models, providing interpretable quantitative evidence for clinical decision-making. Notably, Bayesian approaches exhibit superior adaptability over traditional statistical methods when processing multidimensional, heterogeneous medical data.

In the field of gynecological disorders, the groundbreaking study by Kiser et al. [22] exemplifies the application of BNs. By constructing a hierarchical clustering model encompassing 155 anatomical pain sites, the researchers deconstructed the complex pain mechanisms of endometriosis into 15 clinically interpretable regions. Their BN not only quantified the synergistic diagnostic enhancement effects of concurrent symptoms such as deep dyspareunia and cyclic sciatica but also

outperformed conventional regression models in sensitivity and specificity. This spatiotemporal symptom analysis framework offers critical insights for managing PID. For instance, unilateral adnexal tenderness and diffuse pelvic tenderness may reflect distinct etiological pathways, necessitating differentiated therapeutic strategies. The model's innovation further resides in its ability to integrate multimodal data, including antibiotic response patterns and transvaginal ultrasound features, revealing latent associations between CRP trajectories and endometrial thickness through conditional probability modeling. These advancements provide novel tools for distinguishing PID-related chronic pelvic pain from other gynecological conditions.

Figure 2 demonstrates the analytical framework progresses of [22] through three methodical stages. Initially, spatial pain patterns undergo neighbor-joining clustering with Jaccard distance metric optimization, transforming granular anatomical coordinates into clinically meaningful regions. Subsequently, a hill-climbing algorithm constructs the Bayesian Network's architecture, iteratively optimizing conditional dependencies while maintaining clinical interpretability through node pruning. The final validation phase employs bootstrap resampling to establish diagnostic performance metrics, ultimately operationalized through an interactive web interface enabling probabilistic symptom-diagnosis mapping. Particularly notable is the algorithm's capacity to handle non-linear symptom interactions, for instance, modeling how concurrent dyspareunia and right hypochondrium pain synergistically elevate endometriosis risk beyond simple additive effects. This phased approach effectively bridges spatial pain mapping with multivariable probabilistic reasoning, overcoming traditional regression limitations in handling high-dimensional symptom combinations.

Bayesian methods also excel in optimizing diagnostic systems. Soe et al. [2] developed a sexually transmitted infection diagnostic tool that integrates ML with Bayesian Networks, combining patient-reported symptoms, demographic characteristics, and laboratory indicators to significantly improve diagnostic accuracy in resource-limited regions. Similarly, Lamb et al. [23] employed BNs to classify clinical features of interstitial cystitis/bladder pain syndrome, enabling personalized treatment plans based on urinary frequency and pain intensity. These studies collectively demonstrate the unique advantages of Bayesian frameworks in handling heterogeneous data and balancing diagnostic sensitivity with specificity. The visual representation of probabilistic relationships—such as causal chains between intrauterine device use history and parametrial tenderness—aligns closely with clinical reasoning, enhancing clinician trust in AI-driven decision systems.

In treatment assessment, Bayesian Network meta-analysis (NMA) has emerged as a pivotal tool in evidence-based medicine. Li et al. [24] applied this approach to compare non-pharmacological interventions for primary dysmenorrhea, identifying acupuncture combined with thermotherapy as the most effective pain-relief strategy. Baroncini et al. [25] extended Bayesian NMA to chronic low back pain research, confirming that personalized physical therapy combined with cognitive behavioral therapy yields optimal long-term outcomes. In orthopedics, Migliorini et al. [26] evaluated polyethylene liner wear rates in hip arthroplasty using Bayesian NMA, providing quantitative evidence for material selection. These cases highlight Bayesian methods' potential in addressing clinical heterogeneity and establishing hierarchical evidence for therapeutic interventions, even in the absence of direct comparative trials.

Surgical studies emphasize uncertainty quantification. Sun et al. [27] employed Bayesian NMA to compare surgical approaches for Hirschsprung's disease, demonstrating the superiority of transanal endorectal pull-through in complication control. Yang et al. [28] applied this framework to assess novel therapies for peripheral nerve injuries, offering methodological support for evidence-based selection of neurotrophic factors and stem cell therapies. At the molecular level, Kang et al. [29] integrated meta-analysis with Bayesian modeling to systematically identify endometriosis-associated gene networks, focusing on inflammatory cytokines and angiogenesis regulators. These investigations collectively illustrate the bridging role of Bayesian methods in connecting clinical observations with mechanistic insights.

Notably, Bayesian approaches show unique value in merging traditional medicine with modern analytics. Sumbul et al. [30] innovatively combined ML with Bayesian Networks to validate the efficacy of sesame seeds and rose essential oil in treating uncomplicated PID, optimizing phytomedicine classification through feature importance analysis. This interdisciplinary synergy not only strengthens the evidence base for traditional therapies but also paves new pathways for

their integration into modern healthcare systems.

Despite advantages in prior knowledge integration and missing data handling, Bayesian methods face challenges. The exponential growth of computational complexity in high-dimensional models, as seen in Kiser's pain model [22], and the subjectivity of prior distributions remain critical limitations. Current solutions involve advanced sampling techniques (e.g., Markov chain Monte Carlo, MCMC) and sensitivity analyses to ensure model robustness. Future research must prioritize multimodal data integration (e.g., genomics with electronic health records), develop dynamic Bayesian Networks for real-time disease monitoring, and enhance computational efficiency and interpretability in clinical settings.

Bayesian estimation methods have become integral to modern medical research paradigms. Their strengths in uncertainty quantification and personalized healthcare support exemplified by clinical breakthroughs such as Kiser's pelvic pain model: signal a new era of probabilistic reasoning in medical data analysis. With advancing computational technologies and interdisciplinary collaboration, these methods hold promise for transcending traditional statistical limitations, delivering predictive and adaptive decision-support systems for complex disease management.

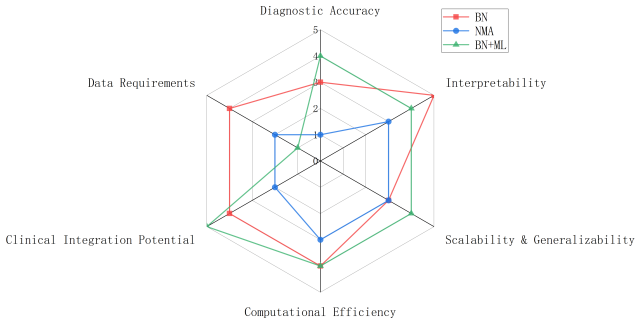


Figure 3. Comparison of some Bayesian network methods mentioned in Section 4.1.

Figure 3 presents a comparative analysis of representative Bayesian Network methodologies, including pure Bayesian Network (BN), Bayesian Network integrated with ML (BN+ML) and NMA. The above three algorithms will be evaluated and compared in the following six dimensions.

The first dimension is Diagnostic Accuracy. The BN approach achieved moderate diagnostic performance. In the endometriosis pain study, the BN reached a sensitivity of ∼81% but a specificity of ∼42%, indicating it caught most true cases but with many false positives [22]. We assign BN an average score in accuracy: it improved association detection over traditional methods but remains imperfect as a standalone diagnostic tool. NMA, in contrast, is not a diagnostic method but an evidence-synthesis tool; its output informs treatment efficacy rather than diagnosing patients [24]. Consequently, NMA scores the least for "diagnostic accuracy" in this context, as it does not directly identify conditions. The BN+ML approach demonstrated high diagnostic accuracy across multiple sexually transmitted infection (STI) conditions (AUC 0.75–0.95 for most targets) [2]. Both its ML classifiers and Bayesian models showed robust performance (accuracy ∼0.68–0.99) with no significant differences for most conditions. We assign BN+ML a highest score in accuracy, reflecting its strong overall diagnostic capability (noting a few lower-performing cases like cervicitis).

The second dimension is Interpretability. BN offers excellent interpretability, which can be fully marked due to their graphical structure and probabilistic relationships. Clinicians can visualize how symptoms like chronic pelvic pain, dyspareunia, and subfertility interconnect and lead to a diagnosis in the BN model. This transparency aligns with clinical reasoning, as the BN revealed specific pain patterns indicative of endometriosis that make intuitive sense. NMA results are moderately interpretable since they present aggregate effect sizes and rank interventions, which experts can interpret (e.g. exercise and acupuncture ranked top for pain reduction) but which are not as intuitively clear as a diagnostic model [24]. The BN+ML approach is fairly interpretable. While ML models alone can be "black boxes," combining them with Bayesian reasoning and using feature importance techniques yielded understandable insights [2]. In the symptom checker study, the models identified key symptoms for each condition such as urethral discharge for gonorrhea and pelvic pain for PID, mirroring clinical decision patterns. Such interpretability approaches (including SHAP analyses) helped bridge the gap between complex ML models and clinician understanding.

The third dimension is Scalability and Generalizability. We rate BN models as average on scalability. They handled the moderate-sized endometriosis dataset well and could incorporate multiple pain features, but extending a BN to many more variables or entirely new patient populations would require careful re-training or expert reconfiguration. In [22], the network was pruned to 18 key nodes for efficiency,

suggesting some limitation in handling overly complex feature sets. NMA is similarly as BN: it can scale to include more studies or treatments as evidence grows, but it is constrained by available clinical trials. The dysmenorrhea NMA, for instance, could only analyze eight non-pharmacological interventions due to limited trials, and noted generalizability concerns since most studies were in Asian populations [24]. The BN+ML approaches the highest ranking for scalability and generalizability. Its ML component excels with large datasets and was validated on unseen patients with good performance. It can be retrained on new data and extended to additional conditions (the study suggests expanding to more STIs and integrating image analysis). Generalizability is largely good, though the authors caution that a single-center dataset may not capture all populations, which means a reminder that large, diverse data are needed to fully generalize the model [2].

The fourth dimension is Computational Efficiency. The BN approach is relatively efficient. Learning the network structure and probabilities from a few hundred cases is feasible with bnlearn in R, and once built, inference (querying the risk of diagnoses given symptoms) is quick [22]. NMA receives an average score here. NMA involve iterative simulations or complex modeling; the referenced study used specialized software (ADDIS) for consistency models: manageable for 33 studies, but computational burden grows with network complexity [24]. Additionally, NMA is an offline analysis, not intended for real-time use. The BN+ML approach is the best. Training multiple ML models on >10,000 records is computationally heavier than the BN's training, but it remained practical (the study successfully evaluated numerous algorithms and performed bootstrap validations). At runtime, the symptom checker's predictive algorithms run efficiently on a web server, providing instant results to users [2]. In sum, both BN and ML+BN approaches are computationally tractable for clinical deployment, whereas NMA is resource-intensive but acceptable within research settings.

The fifth dimension is Clinical Integration Potential. We assign BN the top for integration potential. A BN model, once validated, can be incorporated into clinical decision support systems to estimate diagnosis probabilities. The endometriosis BN, for example, could be used by gynecologists to input a patient's pain profile and get an individualized risk assessment for endometriosis or other pelvic pathologies [22]. Its interpretable nature would likely aid clinician buy-in for use in practice. NMA has lower direct integration potential. Its value is in informing guidelines and broad management strategies, rather than providing patient-specific recommendations on the fly [24]. Clinicians benefit from NMA through updated clinical practice guidelines on interventions, but an NMA isn't a tool one uses during a patient visit. The BN+ML approach stands out with a full score in integration potential. In the STI symptom checker study, the authors actually deployed a web-based tool (iSpySTI) using a Bayesian model to guide patients [2]. They then enhanced it with ML, demonstrating a ready pathway to real-world use. Such a symptom checker can be used directly by patients or frontline providers for preliminary assessment, effectively integrating AI into the care pathway for earlier diagnosis and referral. The ease of web and mobile deployment for the ML models, plus their ability to incorporate images and other data streams, makes this approach highly amenable to clinical and public health integration (e.g. telehealth services).

The sixth dimension is Data Requirements. The BN approach is comparatively data-efficient. It derived meaningful diagnostic insights from a few hundred cases: a relatively small dataset by AI standards [22]. It can also leverage expert knowledge to structure the model, potentially reducing the data needed for training. However, they do require quality data on all relevant variables to perform well; missing or sparse inputs (e.g. certain rare symptoms) could limit accuracy. NMA is data-hungry, as it relies on a robust corpus of clinical trials. The example NMA included 33 RCTs with 2,826 participants, and still many interventions had only a handful of studies each, leading to calls for larger high-quality research to firm up conclusions [24]. Conducting an NMA in a niche area of PID management requires that sufficient studies exist – a high bar for data availability. The BN+ML approach also has heavy data requirements. Building a reliable symptom checker for 12 conditions needed over 10,000 patient cases, and even then certain diagnoses (female cervicitis, male UTI) lacked enough positives to achieve acceptable accuracy [2]. ML thrives on big data; models improve with more examples of each condition. Thus, while the ML+BN system performed well, it demanded extensive prospectively collected data and would need even more data to expand to additional conditions or to different clinical settings.

In this visual summary, the BN (red) displays

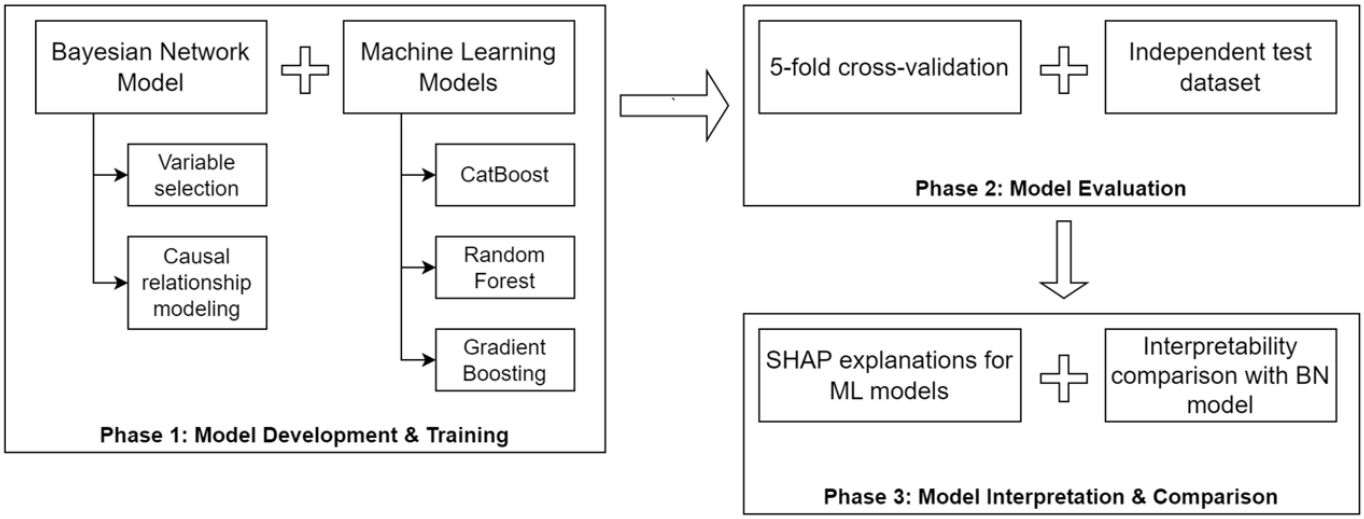


Figure 4. Research flow of [2].

a balanced but moderate profile – it excels in interpretability and offers decent integration potential, reflecting how its transparent probabilistic model can be readily understood and used in clinical decision support. Its diagnostic accuracy and scalability are middling, suggesting the BN is useful for parsing complex symptom relationships but may require refinement or more data to improve predictive power and to generalize widely. The BN+ML approach (green) shows high scores in accuracy and clinical integration, indicating that this hybrid method achieves reliable diagnostic predictions and has already been implemented in practice (e.g. a patient-facing symptom checker). Notably, BN+ML also scores well in scalability, owing to ML's capacity to handle large-scale data and new input types, like images, for broader PID applications. One of its weak point is data requirement: he radar plot confirms that a ML-driven solution demands substantial data for training, a consideration for clinicians and researchers aiming to adopt similar tools.

Meanwhile, the NMA approach (blue) has a unique shape reflecting its role: it ranks high in interpretability (results can be translated into clear clinical guidance) and contributes strongly to evidence-based management rather than diagnosis. The NMA's minimal presence on the diagnostic accuracy axis reinforces that it's an analytic technique not used for identifying diseases. Instead, its value is seen in practice by informing treatment choices – for example, confirming that exercises and acupuncture likely benefit dysmenorrhea patients – which complements diagnostic-focused AI by optimizing therapy. From a clinical standpoint, this comparison highlights that BN offers an interpretable framework suited for differential diagnosis in chronic PID (e.g. distinguishing endometriosis or infection causes of pelvic pain), BN+ML systems provide high-accuracy tools that can be integrated into patient care pathways (though they require significant data investment), and NMA contributes an overarching evidence synthesis to guide interventions. Each approach thus plays a distinct but potentially complementary role in an integrated AI-driven PID management framework, combining diagnostic acumen with evidence-based therapy to improve patient outcomes.

4.2 Machine Learning Based Approach

In 2024 [2], N.N. Soe, Janet M Towns, etc. explored a novel symptom checker tool based on ML and Bayesian Network algorithms for diagnosing common STIs and related genital diseases. The core of the study is the development and evaluation of an online symptom checker called "iSpySTI," which analyzes user-inputted data on gender, behaviors, and symptoms to predict the likelihood of 12 STIs and genital conditions.

The main steps of the research are shown in Figure 4: Model Development and Training, Model Evaluation and Model Interpretation and Comparison. Firstly, a BN was used to construct an initial symptom checker model, predicting disease probabilities through variable selection and causal relationship modeling. Various ML algorithms, including CatBoost, Random Forest, and Gradient Boosting, were applied to the same dataset to build binary classification models to distinguish whether diseases occurred; Secondly, the diagnostic performance of ML models and the Bayesian Network model was compared using

5-fold cross-validation and an independent test dataset, evaluating metrics such as Receiver Operating Characteristic (ROC-AUC) score, sensitivity, and specificity; Thirdly, Shapley Additive Explanations (SHAP) were used to interpret the key predictors of the ML models, and these were compared with the interpretability of the Bayesian Network model [2].

The study found that the ML and Bayesian Network models performed better than the pure Bayesian Network in diagnosing certain diseases (such as male balanitis, molluscum contagiosum, and genital warts), while the Bayesian model demonstrated superior predictive performance in cases of urethral gonorrhea, female PID, and urinary tract infections.

Based on [12], the ML algorithms utilized includes decision trees, support vector machines (SVM), and neural networks. These algorithms were selected for their ability to handle complex data and improve diagnostic accuracy.The implications of this study extend to the management and treatment of PID, a condition included in the study's scope. An accurate symptom checker can facilitate early detection and timely intervention for PID, reducing complications and public health impacts. By applying the ML approaches used in this study, AI-based PID management systems can be enhanced, leading to more effective treatment outcomes.

Wang et al. [31] compared Cox regression and ML models, demonstrating that the XGBoost algorithm outperformed traditional methods in predicting sepsis progression among PID patients, particularly due to its ability to capture nonlinear relationships among multiple parameters. Similarly, Yu et al. [32] developed a random forest model integrating clinical signs, MRI radiomics, and psychological assessment data to predict the risk of myofascial pelvic pain syndrome (MPPS), emphasizing the importance of integrating biopsychosocial features. Tokuc et al. [33] further validated that combining neutrophil-to-lymphocyte ratio with ML significantly improved the prediction accuracy of urethroplasty success.

For superficial peritoneal endometriosis, Santos et al. [34] constructed an ML model analyzing dysmenorrhea patterns, pelvic examination findings, and serum CA125 levels, achieving higher diagnostic accuracy than conventional clinical evaluations. Zhao et al. [35] developed a multimodal diagnostic system combining ultrasound imaging, metabolomics, and pain scales to enhance specificity in diagnosing deep infiltrating endometriosis. Notably, Okui [36] identified a vulvodynia-predominant subtype of bladder pain syndrome using unsupervised learning, providing insights for personalized treatment. Liu et al. [37] integrated single-cell sequencing and clinical data to reveal the critical role of the NFKBIZ gene in promoting PID progression via regulation of the IL-6/STAT3 pathway. Zhou et al. [38] identified immune-related biomarkers, including CXCL10 and TLR4, and developed an ensemble model to differentiate interstitial cystitis from other lower urinary tract disorders, offering new avenues for targeted therapies.

Kumar et al. [39] designed a wrist pulse analysis system with optimized feature selection for early detection of pelvic infections. Xiang et al. [40] proposed a multimodal model integrating perianal fistula imaging, gut microbiome profiles, and serum inflammatory markers to predict Crohn's disease risk. These innovations underscore the transformative potential of ML and bioinformatics in bridging clinical gaps across pelvic disorders.

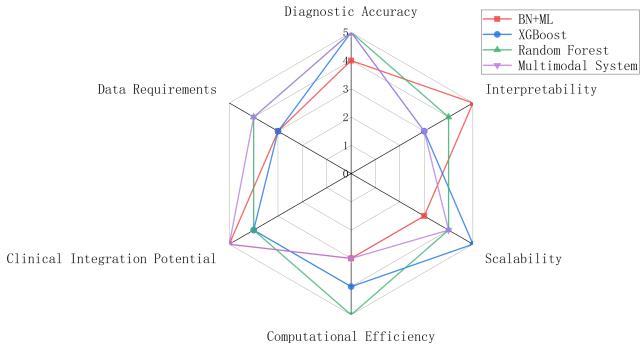


Figure 5. Comparison of some machine learning based approach methods mentioned in Section 4.2.

Figure 5 presents a comparative analysis of representative ML methodologies, including BN+ML, XGBoost algorithm, random forest model and multi-model method. Among the compared models, the multimodal ML system demonstrated the highest diagnostic accuracy, achieving an AUC of 0.85 and sensitivity of 86.2% by integrating random forest algorithms with non-invasive biomarkers such as CA125 and NLR for endometriosis prediction [35]. XGBoost also displayed excellent performance in predicting the risk of PID progression, with multiple models achieving AUCs exceeding 0.90 [31]. Similarly, the random forest model used in myofascial pelvic pain syndrome detection yielded an outstanding AUC of 0.942 [32]. As for the BN+ML system, while robust, exhibited slightly lower accuracy in certain

STI conditions, such as cervicitis, with AUCs ranging between 0.75 and 0.95 across diseases in the iSpySTI platform [2].

Interpretability remains a critical factor for clinical adoption. The BN+ML framework offers a distinct advantage in this dimension, as its Bayesian component provides transparent probabilistic reasoning between symptoms and outcomes, supporting clinical trust and validation [2]. The random forest model maintains relatively high interpretability through accessible feature importance measures, enabling clinicians to assess key predictors [32]. In contrast, XGBoost, though powerful, suffers from the opacity typical of gradient boosting models, requiring advanced tools such as SHAP values for explanation, which may not be practical in everyday clinical environments [31]. The multimodal approach integrating multiple serological indicators, while clinically useful, also results in a more complex decision structure, making it less intuitive for medical practitioners [35].

The XGBoost model demonstrated superior scalability due to its capacity to handle high-dimensional data and maintain performance across large patient cohorts [31]. The multimodal system, which relies on standardized clinical laboratory data, also shows promise in this regard, especially given the ubiquity of CA125 and inflammatory markers in gynecologic diagnostics [35]. Random forest models, while generally scalable, may require more tuning when applied to larger, noisier datasets, although they perform reliably in medium-sized clinical samples [32]. BN+ML systems are somewhat limited in this domain due to the inherent complexity of Bayesian Network structure learning, which does not scale efficiently when the number of input variables increases significantly [2].

In terms of computational efficiency, the random forest algorithm consistently delivered high-speed performance with minimal parameter tuning, making it well-suited for integration into clinical decision support systems [32]. XGBoost, although powerful, requires intensive training and parameter optimization phases, which increase computational demands [31]. The BN+ML model, while clinically valuable, involves both probabilistic inference and ML components, which can be more resource-intensive than purely tree-based models [2]. The multimodal system, integrating serologic data with ML, exhibits moderate efficiency; while its individual components are computationally light, the combined modeling of multiple biomarkers requires careful feature selection and validation [35].

BN+ML models exhibit the highest clinical integration potential, as evidenced by the iSpySTI platform's deployment in real-world settings, offering patients and medical staff diagnostic support with real-time feedback [2]. Multimodal systems, relying on routine and minimally invasive tests like CA125 and NLR, are highly adaptable to existing diagnostic protocols in gynecology, further reinforcing their clinical usability [35]. Random forest models are suitable for integration into electronic medical records or diagnostic interfaces, though they require additional validation for varied PID subtypes [32]. XGBoost models, despite their high accuracy, remain less represented in live clinical tools due to their complexity and lower explainability [31].

In terms of data dependence, the multimodal system ranks favorably, utilizing standard clinical markers that are routinely collected, allowing broad accessibility in healthcare settings [35]. The random forest model also performs well with moderately sized structured datasets, enabling deployment in medium-resource environments [32]. XGBoost and BN+ML systems, however, require large, well-labeled datasets to realize their full potential. The BN+ML method is particularly sensitive to diverse input features and comprehensive training data across disease categories [2], while XGBoost's performance may degrade when applied to smaller or noisier datasets without rigorous preprocessing [31].

4.3 Artificial Neural Network Approach

In 2023, Vibha Verma and Yadvendra Singh [41] underscored the profound burden of inflammatory diseases on women's health, with a specialized focus on genitourinary tract infections—a framework directly applicable to advancing the diagnosis and management of PID. Their proposed Artificial Intelligence-Based Approach (AIBA) exemplifies a paradigm shift in addressing the multifactorial nature of such conditions through the strategic convergence of ML methodologies. By employing supervised learning algorithms trained on clinical and multimodal datasets, AIBA establishes predictive models capable of distinguishing PID from clinically similar conditions such as endometriosis or urinary tract infections. The framework further incorporates MTL architectures to optimize both diagnostic accuracy and personalized treatment recommendations, utilizing hierarchical

neural networks to process microbiological data, inflammatory markers, and imaging features, thereby capturing interactions between etiological agents and host responses.

This study introduces AIBA to enhance diagnostic precision and therapeutic management of female genitourinary inflammatory pathologies, with a focus on recurrent urinary tract infections (UTIs) and chronic pelvic inflammatory disorders. The research addresses limitations in conventional diagnostic approaches, which often rely on subjective symptom assessment and delayed microbial culture results, contributing to antibiotic misuse and recurrent infections. Leveraging deep learning architectures, the proposed model integrates clinical metadata, biochemical markers, and histopathological imaging data to establish multi-parametric diagnostic criteria [41].

The methodological framework, as shown in Figure 6, employs convolutional neural networks (CNN) for feature extraction from cellular morphology patterns in urine sediment analysis, coupled with regression-based optimization of therapeutic response predictions. Experimental validation compared AIBA against traditional clinical decision protocols (MUTI) and pregnancy-specific diagnostic guidelines (UTIA), demonstrating significant improvements in accuracy, precision, and computational efficiency [41].

When evaluated in the context of chronic PID management, BN, ML, and ANNs (e.g., AIBA) demonstrate distinct strengths and limitations, which are presented in Figure 7. From a clinical application perspective, BN excels in interpretability and explainability, enabling clinicians to visualize probabilistic relationships between symptoms and diagnoses—a feature that significantly enhances decision-making transparency. BN modeling effectively distinguishes symptom patterns in endometriosis, albeit with moderate diagnostic accuracy and limited scalability due to structural rigidity. In contrast, ML models such as XGBoost and Random Forest demonstrated superior diagnostic accuracy and scalability, as shown in studies on PID progression and MPPS detection, with AUC values often exceeding 0.90 [31, 32]. These models, though less interpretable, have shown strong integration into patient-centered tools like the iSpySTI platform [2].

On the other hand, ANN-based models such as AIBA excel in computational efficiency and real-time adaptability due to their deep learning architecture [41]. However, they present the greatest challenge in explainability and clinical transparency. Despite achieving high classification accuracy across inflammatory disease types, the AIBA model lacks clear interpretive outputs, which may limit trust and adoption in clinical practice. Technically, ANNs represent a significant advancement in pattern recognition and automation, yet their "black-box" nature hinders clinical utility without supplementary explainability frameworks.

Looking forward, the most promising direction lies in hybrid frameworks that combine the clinical clarity of BNs, the scalability and robustness of traditional ML models, and the adaptability of ANN architectures. By integrating interpretability with computational power, such multi-paradigm systems could offer both precision and transparency in future AI-enhanced PID management tools.

5 Currently Available Technologies

5.1 Cross-Domain Relevance of MFRNet in PID Management

PID presents a range of diagnostic and therapeutic challenges—unclear lesion boundaries in ultrasound or CT images, lesions that span from micro-inflammation to full abscess, and confounding anatomical structures such as bowel loops—that closely mirror the problems addressed by the MFRNet framework [42]. In particular, the Spatial-Optimized Feature Attention (SOFA) module sharpens object boundaries, while the Context-Aware Channel Refinement (CACR) module effectively suppresses irrelevant anatomical noise. Moreover, MFRNet's use of brightness and contrast augmentation during training enhances its robustness in low-quality imaging environments, a feature directly translatable to real-time PID monitoring scenarios.

In image-driven diagnostic frameworks, MFRNet's dual attention mechanism provides a powerful template for detecting subtle pelvic lesions. The SOFA module refines the borders of inflammatory exudates—such as tubal wall thickening—while CACR optimizes channel weights to suppress vascular clutter and other confounders. Quantitatively, MFRNet achieves a 12.7% reduction in mean absolute error (MAE) on the ECSSD benchmark, underscoring its potential to improve lesion delineation in complex pelvic MRI or CT scans [42].

For adaptive therapeutic systems, MFRNet demonstrates an architecture capable of real-time

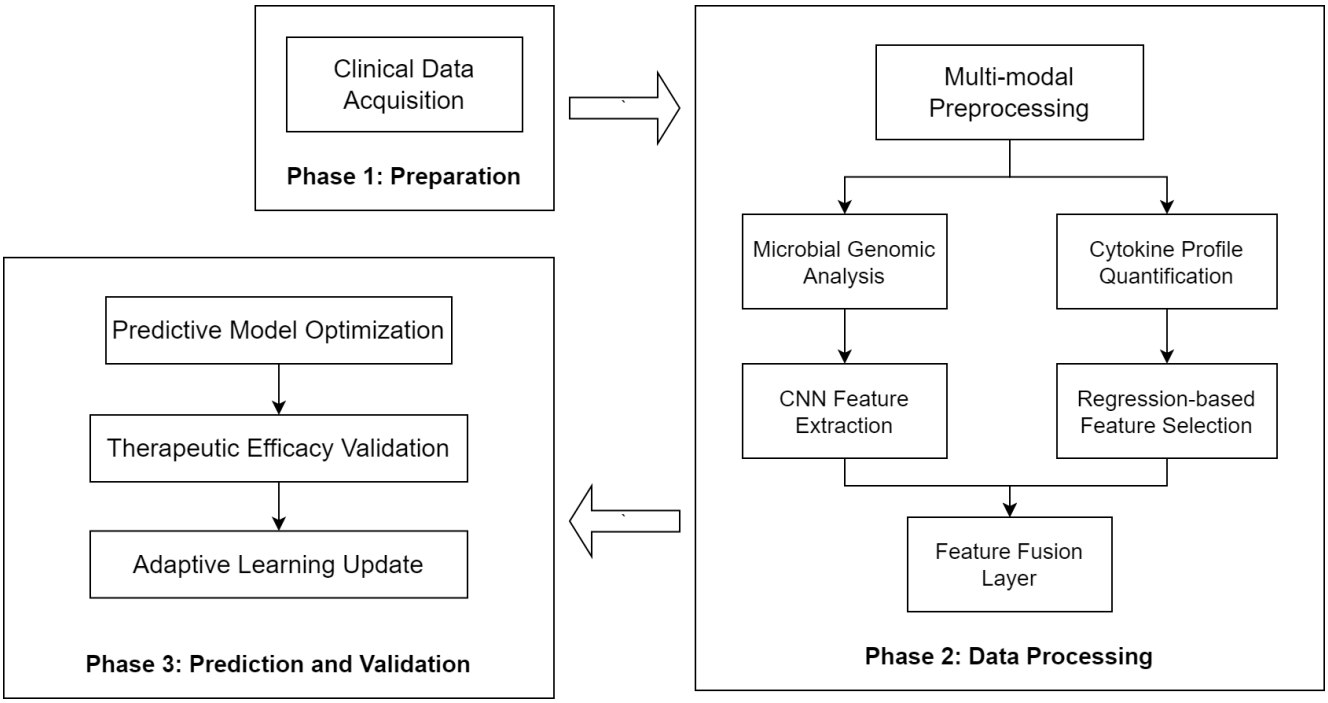


Figure 6. Research flow of [41].

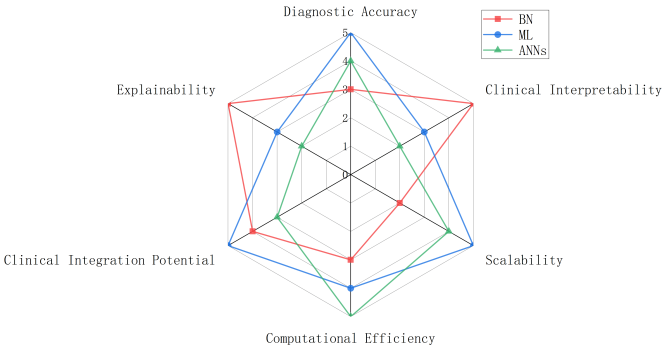


Figure 7. Comparison of BN, ML and ANNs.

video analysis at 38.28 frames per second, making it well-suited for dynamic endoscopic monitoring of PID lesion progression [42]. Its brightness-contrast augmentation strategy directly counteracts variability in transvaginal ultrasound quality, ensuring consistent performance across patient-specific factors such as adiposity or mucosal secretions. These capabilities are summarized in Table 2, which outlines how specific MFRNet components address key PID clinical needs, including boundary definition, noise reduction, real-time monitoring, and data augmentation.

From a data-centric perspective, MFRNet's multi-scale feature fusion informs the design of PID datasets by advocating hierarchical lesion annotation, from pixel-level inflammation to macroscopic adhesions, that mirrors its six-stage EfficientNet-B7 backbone extraction [42]. This approach aligns with radiologists' multi-level analysis of pelvic inflammation, promoting richer training data and more nuanced diagnostic models.

Despite these strengths, a domain shift between natural imagery and medical imaging necessitates fine-tuning of MFRNet on PID-specific datasets. In particular, adjusting the dilation rates within the CACR module may enhance sensitivity to micro-abscesses commonly encountered in early-stage PID. Looking ahead, extending MFRNet to integrate multi-modal inputs—such as combining RGB imaging with thermal pelvic thermography—could further bolster diagnostic accuracy. Finally, clinical validation against expert radiologist annotations using the boundary-aware MAE metric will be critical to establish its real-world efficacy in PID management.

5.2 Cross-Domain Relevance of YOLOv8 Framework in PID Management

PID also poses unique challenges for image-driven diagnostics, where the ability to localize diffuse inflammatory regions amid anatomical clutter is critical. The Convolutional Block Attention Module (CBAM) can be repurposed to highlight PID-specific markers, such as subtle tubal wall thickening or fluid accumulations by sequentially applying channel and spatial attention [43]. In this adaptation, CBAM's channel attention would prioritize feature maps containing inflammatory signals, while its spatial attention would refine the localization of these signals

Table 2. MFRNet technology modules which can be applied in PID management.

PID Clinical Need	MFRNet Component	Implementation Example
Boundary Definition	SOFA Module	Delineating pyosalpinx margins
Noise Reduction	CACR Channel Attention	Suppressing bowel gas artifacts
Real-time Monitoring	EfficientNet-B7 Backbone	Portable ultrasound deployment
Data Augmentation	Brightness Adjustment	Standardizing obese patient scans

in ultrasound or CT slices. This dual-attention pipeline offers a pathway to more precise delineation of lesion boundaries that are otherwise obscured by bowel loops or vascular noise [43].

Beyond static imaging, real-time therapeutic monitoring of PID requires models that can process video streams at high frame rates on edge devices. The optimized YOLOv8 framework achieves up to 126 FPS through a lightweight backbone and Focus module enhancements [43]. This throughput enables continuous analysis of endoscopic or ultrasound video feeds, making it feasible to track lesion dynamics and patient responses during treatment sessions. For instance, integrating a fall-detection–inspired pipeline, originally designed to flag abrupt posture changes—into a PID telemedicine system could allow simultaneous detection of pain-expression cues or sudden anatomical shifts indicative of acute exacerbations [43].

Building robust AI systems for PID also mandates careful dataset design to prevent bias and ensure generalization across patient subgroups. The DiverseFALL10500 benchmark incorporates extensive variability in lighting, subject demographics (including BMI ranges), and camera angles [43]. Translating these principles, a PID dataset should include transvaginal scans from diverse populations, with variations in adiposity, mucosal secretions, and probe orientations to train models that remain reliable under real-world clinical conditions [43]. This heterogeneity promotes resilience against domain shifts and helps guard against performance drop-offs when encountering underrepresented patient cohorts.

However, transferring techniques from general vision to medical imaging is not without hurdles. Whereas fall detection hinges on macroscopic posture cues, PID diagnosis demands sensitivity to micro-scale adhesions and early abscesses. The CBAM's Focus module, which reduces computational overhead in natural images, must be reconfigured, potentially through increased dilation rates or higher-resolution inputs to capture microscopic pathologies without sacrificing inference speed. Such domain-specific adjustments will be essential to fully leverage attention-based models for the nuanced task of PID management.

6 Current Shortcomings and Controversies

Despite significant advances, the interpretability of AI models remains a critical concern. Most deep learning systems operate as opaque "black boxes," with decision-making processes hidden even from expert users [44, 45]. This opacity challenges clinician trust and limits adoption: healthcare professionals are less willing to rely on tools they cannot explain, and patients may object to decisions they cannot comprehend [46]. To mitigate this, the field is actively exploring explainable-AI (XAI) methods such as Grad-CAM, feature visualization, and structured segmentation pipelines; however, a known trade-off persists: the most transparent models often sacrifice predictive performance [47].

A second major challenge is dataset bias and privacy. AI models are only as effective as their training data, and imbalanced datasets which lack demographic diversity risk embedding systemic disparities [48]. Additionally, sensitive patient data invoke stringent regulatory protections under frameworks like GDPR, PIPL, and HIPAA [49]. Addressing these issues requires more representative, multi-center data collection and privacy-preserving training techniques (e.g., federated learning, synthetic data), although these introduce additional complexity and validation hurdles.

Thirdly, regulatory and legal accountability pose barriers to clinical deployment. Clinicians remain ultimately responsible for patient outcomes, yet opaque AI tools complicate attribution of errors and due diligence [50]. While some argue that AI-generated diagnoses can be accepted without full transparency, this approach conflicts with existing medical ethics and legal standards that demand explainability for informed consent and error accountability [50].

Table 3. YOLOv8 Framework Modules which can be Applied in PID Management.

PID Clinical Need	YOLOv8 Framework Contribution	Impact
AI-Driven Diagnostic Imaging	CBAM for Lesion Emphasis	Improves inflammation localization
Real-Time Patient Monitoring	High-FPS Model Design	Enables low-latency symptom detection
Medical Data Curation	Diversity Strategies	Mitigates dataset bias in PID

Finally, the generalizability of AI systems to diverse real-world environments is not ensured. Many models exhibit strong performance under controlled conditions but fail when exposed to varied clinical settings, equipment differences, or patient populations, especially outside of high-resource regions [51]. Bridging this gap calls for comprehensive multi-site trials, continuous post-deployment monitoring, and alignment with real-world clinical workflows.

In summary, while AI offers tremendous promise in PID management, deploying it responsibly demands transparent, equitable, and context-aware systems. Focusing on interpretability, representative data practices, regulatory compliance, and clinical validation will be essential to realize its full potential as a trusted, real-world clinical assistant.

In addition to the shortcomings mentioned above that need improvement, this technology also raises many controversial issues. High-dimensional EHR data frequently exhibit systematic biases: notably selection bias, measurement bias, temporal bias, and implicit bias, which can seriously impair the fairness and generalizability of AI-based PID diagnostics. A recent systematic review identified six major bias types within EHR-AI models and found a concerning lack of real-world deployment, despite widespread fairness assessment using metrics like statistical parity and equal opportunity [52, 53]. Furthermore, Al-Sahab et al. [54] emphasize that selection and information biases arise both at the data capture stage and during research pipelines, warranting extreme caution in inferring causality from secondary EHR data. To mitigate these distortions, implementing strategies such as re-sampling/re-weighting, data completeness checks, harmonization protocols, and active bias monitoring during model training is crucial. In the context of chronic PID, these measures can reduce demographic skew, such as underrepresentation of rural or low-income women, and thereby support more equitable diagnostic support tools.

Deploying AI-driven PID diagnostic tools in low-resource or rural healthcare settings presents unique challenges encompassing infrastructure, human capacity, governance, and culture. A targeted review of AI in LMICs outlined 40 challenges, ranging from data quality, context awareness, regulatory frameworks, training resistance, financial constraints, to infrastructure and scalability [55]. Additionally, large-scale implementation studies in general healthcare identify structural barriers: poor internet connectivity, limited computing infrastructure, absence of robust data governance, and low clinician trust due to lack of explainability [56]. Practical strategies observed in telehealth and mHealth deployments (e.g. smartphone-based ultrasound or chatbots) include edge computing, offline-first interfaces, multi-language support, and federated learning constructs to overcome data privacy and connectivity constraints. For chronic PID specifically, pairing low-cost diagnostic hardware with AI models trained on context-representative data, and deploying them via cloud-sync echo systems, could create a viable path for scalable adoption in under-served clinics.

To ensure longevity and equity in PID–focused AI systems, it is essential to combine rigorous data quality management with resource-sensitive deployment planning. Robust pipelines incorporating bias auditing, EHR data quality assessment, and fairness-aware model training can ensure internal validity. Simultaneously, building sustainable adoption models through co-design with front-line providers, training frameworks, regulatory clarity, and flexible technological architectures can enhance external validity across diverse health systems. Embedding these pillars, such as quality, explainability, adaptability and inclusivity, is critical to realizing AI's promise in chronic PID management without exacerbating health disparities.

7 Conclusion and outlook

There are several avenues for future research that could further enhance its effectiveness and applicability:

1. Integration of Diverse AI Techniques: Combining Bayesian Networks, ML, and deep learning is a optimal method to create more robust diagnostic and predictive models. For instance, using deep learning for image analysis and Bayesian Networks for probabilistic reasoning could provide a more comprehensive understanding of PID progression. Also, the use of reinforcement learning will develop adaptive treatment strategies that adjust in real-time based on patient responses and disease progression.

2. Big Data Handling and Scalability: Scalable Algorithms used in developing AI models can efficiently process and analyze large-scale medical datasets, including electronic health records, imaging data, and genomic information. High-Dimensional Data Analysis can create algorithms which can handle the complexity and variability of high-dimensional data, providing more accurate predictions and insights.

3. Real-Time Monitoring and Prediction: Leveraging wearable devices and IoT technologies to continuously monitor patients and predict potential flare-ups or complications in real-time is crucial in AI-based PID management and treatment technology. The creation of AI-driven early warning systems that alert healthcare providers to potential complications will also enable timely interventions.

4. Ethical Considerations: Ethical challenges can be addressed by developing frameworks that ensure patient data privacy and security in AI-driven healthcare systems, which can be achieved through investigating and mitigating biases in AI algorithms to ensure equitable treatment recommendations and avoid disparities in care.

In the evolving landscape of AI-based medical diagnostics and treatment, the integration of Bayesian Networks and ML algorithms emerges as a promising trend, poised to revolutionize healthcare outcomes. Bayesian Networks, with their probabilistic framework, excel in capturing the uncertainty inherent in medical diagnosis, particularly in the early stages of disease when symptoms are often vague or overlapping. Conversely, ML algorithms, such as decision trees, support vector machines, and neural networks, offer robust pattern recognition capabilities, enabling the detection of complex relationships within large datasets. The synergy between these approaches can enhance the accuracy and reliability of diagnostic tools, as demonstrated in the study of symptom checking for STIs.

This combined approach allows for a flexible application of the most suitable algorithm depending on the disease stage. In the initial phases, where symptoms are subtle, Bayesian Networks can provide probabilistic insights, suggesting potential diagnoses. As the disease progresses and more data becomes available, ML models can refine these predictions, leveraging a broader range of features for enhanced accuracy. Furthermore, in treatment phases, these integrated models can predict the efficacy of different treatment options based on patient characteristics and disease progression, thereby optimizing treatment plans.

Data Availability Statement

Not applicable.

Funding

This work was supported without any funding.

Conflicts of Interest

The authors declare no conflicts of interest.

Ethical Approval and Consent to Participate

Not applicable.

References

[1] Sultana, A., Mehdi, S., Rahman, K., Fazmiya, M. J. A., Heyat, M. B. B., Akhtar, F., & Baig, A. A. (2022). Recent advancements of pelvic inflammatory disease: A review on evidence-based medicine. *Computational Intelligence in Healthcare Applications*, 101-120. [Crossref]

[2] Soe, N. N., Towns, J. M., Latt, P. M., Woodberry, O., Chung, M., Lee, D., ... & Fairley, C. K. (2024). Accuracy of symptom checker for the diagnosis of sexually transmitted infections using machine learning and Bayesian network algorithms. *BMC Infectious Diseases, 24*(1), 1408. [Crossref]

[3] Curry, A., Williams, T., & Penny, M. L. (2019). Pelvic inflammatory disease: diagnosis, management, and prevention. *American family physician, 100*(6), 357-364.

[4] Centers for Disease Control and Prevention. *Pelvic Inflammatory Disease (PID) Treatment and Care*. Retrieved from https://www.mayoclinic.org/diseases-conditions/pelvic-inflammatory-disease/diagnosis-treatment/drc-20352600

[5] Gradison, M. (2012). Pelvic inflammatory disease. *American family physician, 85*(8), 791-796.

[6] Sutton, M. Y., Sternberg, M., Zaidi, A., Louis, M. E. S., & Markowitz, L. E. (2005). Trends in pelvic inflammatory disease hospital discharges and ambulatory visits, United States, 1985–2001. *Sexually transmitted diseases, 32*(12), 778-784. [Crossref]

[7] Revzin, M. V., Mathur, M., Dave, H. B., Macer, M. L., & Spektor, M. (2016). Pelvic inflammatory disease: multimodality imaging approach with clinical-pathologic correlation. *Radiographics, 36*(5), 1579-1596. [Crossref]

[8] Haggerty, C. L., Gottlieb, S. L., Taylor, B. D., Low, N., Xu, F., & Ness, R. B. (2010). Risk of sequelae after Chlamydia trachomatis genital infection in women. *The Journal of infectious diseases, 201*(Supplement_2), S134-S155. [Crossref]

[9] Kyrimi, E., McLachlan, S., Dube, K., Neves, M. R., Fahmi, A., & Fenton, N. (2021). A comprehensive scoping review of Bayesian networks in healthcare: Past, present and future. *Artificial Intelligence in Medicine, 117*, 102108. [Crossref]

[10] Stephenson, T. A. (2000). An introduction to Bayesian network theory and usage. *IDIAP-RR 00-03*, 78-83.

[11] Polotskaya, K., Muñoz-Valencia, C. S., Rabasa, A., Quesada-Rico, J. A., Orozco-Beltrán, D., & Barber, X. (2024). Bayesian networks for the diagnosis and prognosis of diseases: A scoping review. *Machine Learning and Knowledge Extraction, 6*(2), 1243-1262. [Crossref]

[12] Pandey, D., Niwaria, K., & Chourasia, B. (2019). Machine learning algorithms: a review. *Mach. Learn, 6*(2).

[13] Kononenko, I. (2001). Machine learning for medical diagnosis: history, state of the art and perspective. *Artificial Intelligence in medicine, 23*(1), 89-109. [Crossref]

[14] Zhang, Y., & Yang, Q. (2021). A survey on multi-task learning. *IEEE transactions on knowledge and data engineering, 34*(12), 5586-5609. [Crossref]

[15] Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., ... & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical image analysis, 42*, 60-88. [Crossref]

[16] Muñoz-Valencia, C. S., Quesada, J. A., Orozco, D., & Barber, X. (2023). Employing Bayesian Networks for the Diagnosis and Prognosis of Diseases: A Comprehensive Review. *arXiv preprint arXiv:2304.06400*.

[17] Ngartera, L., Issaka, M. A., & Nadarajah, S. (2024). Application of bayesian neural networks in healthcare: three case studies. *Machine Learning and Knowledge Extraction, 6*(4), 2639-2658. [Crossref]

[18] McLachlan, S., Dube, K., Hitman, G. A., Fenton, N. E., & Kyrimi, E. (2020). Bayesian networks in healthcare: Distribution by medical condition. *Artificial intelligence in medicine, 107*, 101912. [Crossref]

[19] Yet, B., Perkins, Z. B., Rasmussen, T. E., Tai, N. R., & Marsh, D. W. R. (2014). Combining data and meta-analysis to build Bayesian networks for clinical decision support. *Journal of biomedical informatics, 52*, 373-385. [Crossref]

[20] Jalal, H., Trikalinos, T. A., & Alarid-Escudero, F. (2021). BayCANN: streamlining Bayesian calibration with artificial neural network metamodeling. *Frontiers in Physiology, 12*, 662314. [Crossref]

[21] Velichko, A. (2021). A method for medical data analysis using the LogNNet for clinical decision support systems and edge computing in healthcare. *Sensors, 21*(18), 6209. [Crossref]

[22] Kiser, A. C., Schliep, K. C., Hernandez, E. J., Peterson, C. M., Yandell, M., & Eilbeck, K. (2024). An artificial intelligence approach for investigating multifactorial pain-related features of endometriosis. *Plos one, 19*(2), e0297998. [Crossref]

[23] Lamb, L. E., Janicki, J. J., Bartolone, S. N., Ward, E. P., Abraham, N., Laudano, M., ... & Chancellor, M. B. (2024). Risk classification for interstitial cystitis/bladder pain syndrome using machine learning based predictions. *Urology, 189*, 19-26. [Crossref]

[24] Li, X., Hao, X., Liu, J. H., & Huang, J. P. (2024). Efficacy of non-pharmacological interventions for primary dysmenorrhoea: a systematic review and Bayesian network meta-analysis. *BMJ Evidence-Based Medicine, 29*(3), 162-170. [Crossref]

[25] Baroncini, A., Maffulli, N., Schäfer, L., Manocchio, N., Bossa, M., Foti, C., ... & Migliorini, F. (2024). Physiotherapeutic and non-conventional approaches in patients with chronic low-back pain: a level I Bayesian network meta-analysis. *Scientific reports, 14*(1), 11546. [Crossref]

[26] Migliorini, F., Betsch, M., Maffulli, N., Schäfer, L., Hildebrand, F., Kubach, J., & Pasurka, M. (2024). Rate of revision and wear penetration in different polyethylene liner compositions in total hip arthroplasty: a Bayesian network meta-analysis. *Scientific reports, 14*(1), 21162. [Crossref]

[27] Sun, D., Zhang, X., Xu, Q., Li, Y., Zhang, Q., Wang, D., ... & Li, A. (2024). Duhamel and transanal endorectal pull-throughs for Hirschsprung disease: a Bayesian network meta-analysis. *BMC surgery, 24*(1), 132. [Crossref]

[28] Yang, Y., Gu, W., Xu, S., Wang, S., Shi, H., Zhang, L., ... & Du, Y. (2024). Treatment for peripheral nerve injury: a protocol for a systematic review and Bayesian

network meta-analysis. *BMJ open, 14*(12), e090497. [Crossref]

[29] Kang, J., Ahn, K., Oh, J., Lee, T., Hwang, S., Uh, Y., & Choi, S. J. (2025). Identification of Endometriosis Pathophysiologic-Related Genes Based on Meta-Analysis and Bayesian Approach. *International Journal of Molecular Sciences, 26*(1), 424. [Crossref]

[30] Sumbul, Sultana, A., Heyat, M. B. B., Rahman, K., Akhtar, F., Parveen, S., ... & Malik, A. (2024). Efficacy and classification of Sesamum indicum linn seeds with Rosa damascena mill oil in uncomplicated pelvic inflammatory disease using machine learning. *Frontiers in Chemistry, 12*, 1361980. [Crossref]

[31] Wang, Q., Sun, J., Liu, X., Ping, Y., Feng, C., Liu, F., & Feng, X. (2024). Comparison of risk prediction models for the progression of pelvic inflammatory disease patients to sepsis: Cox regression model and machine learning model. *Heliyon, 10*(1). [Crossref]

[32] Yu, H., Zhao, H., Liu, D., Dong, Y., Nai, M., Song, Y., ... & Li, X. (2024). Prediction of myofascial pelvic pain syndrome based on random forest model. *Heliyon, 10*(11). [Crossref]

[33] Tokuc, E., Eksi, M., Kayar, R., Demir, S., Topaktas, R., Bastug, Y., ... & Ozturk, M. (2024). Inflammation indexes and machine-learning algorithm in predicting urethroplasty success. *Investigative and Clinical Urology, 65*(3), 240. [Crossref]

[34] Santos, L. L. A., de Azevedo, M. C., Shimamura, L. K., Nogueira, A. A., Candido-dos-Reis, F. J., Schor, E., ... & Poli-Neto, O. B. (2024). Machine learning as a clinical decision support tool for diagnosing superficial peritoneal endometriosis in women with dysmenorrhea and acyclic pelvic pain. *Medical Research Archives, 12*(12). [Crossref]

[35] Zhao, N., Hao, T., Zhang, F., Ni, Q., Zhu, D., Wang, Y., ... & Mi, X. (2024). Application of machine learning techniques in the diagnosis of endometriosis. *BMC Women's Health, 24*(1), 491. [Crossref]

[36] Okui, N. (2024). Unsupervised machine learning reveals a vulvodynia-predominant subtype in bladder pain syndrome/interstitial cystitis. *Cureus, 16*(6). [Crossref]

[37] Liu, B., Wang, Q., Dong, H., Zhang, J., Pu, X., & Yu, J. (2025). Integrative bioinformatics and experimental approaches reveal NFKBIZ as a key regulator of inflammatory factors and chemokines in pelvic inflammatory disease progression. *Gene Reports, 39*, 102194. [Crossref]

[38] Zhou, T., Zhu, C., Zhang, W., Wu, Q., Deng, M., Jiang, Z., ... & Zou, C. (2025). Identification and validation of immune and diagnostic biomarkers for interstitial cystitis/painful bladder syndrome by integrating bioinformatics and machine-learning. *Frontiers in Immunology, 16*, 1511529. [Crossref]

[39] Kumar, S., Veer, K., & Kumar, S. (2024). Development of a novel wrist pulse system for early diagnosis of pathogenic bacterial infections using optimized feature selection with machine learning approaches. *Biomedical Signal Processing and Control, 87*, 105503. [Crossref]

[40] Xiang, Y., Yang, F., Yuan, F., Gong, Y., Li, J., Wang, X., ... & Yang, B. (2025). Development and Validation of a Multimodal Machine Learning Model for Diagnosing and Assessing Risk of Crohn's Disease in Patients With Perianal Fistula. *Alimentary Pharmacology & Therapeutics, 61*(5), 824-834. [Crossref]

[41] Verma, V., & Singh, Y. (2023, February). Improved Analysis of Inflammatory Diseases in Women using Artificial Intelligence based Approach. In *2023 IEEE International Conference on Integrated Circuits and Communication Systems (ICICACS)* (pp. 1-7). IEEE. [Crossref]

[42] Chandramohan, A., Bhat, T. A., John, R., & Simon, B. (2020). Multimodality imaging review of complex pelvic lesions in female pelvis. *The British Journal of Radiology, 93*(1116), 20200489. [Crossref]

[43] Saraei, M., Lalinia, M., & Lee, E. J. (2025). Deep Learning-Based Medical Object Detection: A Survey. *IEEE Access, 13*, 53019-53038. [Crossref]

[44] Amann, J., Blasimme, A., Vayena, E., Frey, D., Madai, V. I., & Precise4Q Consortium. (2020). Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC medical informatics and decision making, 20*(1), 310. [Crossref]

[45] Javed, H., El-Sappagh, S., & Abuhmed, T. (2024). Robustness in deep learning models for medical diagnostics: security and adversarial challenges towards robust AI applications. *Artificial Intelligence Review, 58*(1), 12. [Crossref]

[46] Zhang, J., & Zhang, Z. M. (2023). Ethics and governance of trustworthy medical artificial intelligence. *BMC medical informatics and decision making, 23*(1), 7. [Crossref]

[47] Ameen, S., Wong, M. C., Yee, K. C., & Turner, P. (2022). AI and clinical decision making: the limitations and risks of computational reductionism in bowel cancer screening. *Applied Sciences, 12*(7), 3341. [Crossref]

[48] Taqa, Amer. (2025). AI-Assisted Diagnosis in Healthcare: Opportunities and Limitations.

[49] Hulsen, T. (2023). Explainable artificial intelligence (XAI): concepts and challenges in healthcare. *Ai, 4*(3), 652-666. [Crossref]

[50] Marey, A., Arjmand, P., Alerab, A. D. S., Eslami, M. J., Saad, A. M., Sanchez, N., & Umair, M. (2024). Explainability, transparency and black box challenges of AI in radiology: impact on patient care in cardiovascular radiology. *Egyptian Journal of Radiology and Nuclear Medicine, 55*(1), 183. [Crossref]

[51] Hatherley, J. J. (2020). Limits of trust in medical AI. *Journal of medical ethics, 46*(7), 478-481. [Crossref]

[52] Lewis, A. E., Weiskopf, N., Abrams, Z. B., Foraker, R., Lai, A. M., Payne, P. R., & Gupta, A. (2023). Electronic health record data quality assessment and tools: a systematic review. *Journal of the American Medical Informatics Association, 30*(10), 1730-1740. [Crossref]

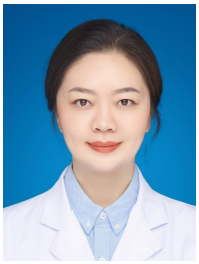
[53] Chen, F., Wang, L., Hong, J., Jiang, J., & Zhou, L. (2024). Unmasking bias in artificial intelligence: a systematic review of bias detection and mitigation strategies in electronic health record-based models. *Journal of the American Medical Informatics Association, 31*(5), 1172-1183. [Crossref]

[54] Al-Sahab, B., Leviton, A., Loddenkemper, T., Paneth, N., & Zhang, B. (2024). Biases in electronic health records data for generating real-world evidence: an overview. *Journal of healthcare informatics research, 8*(1), 121-139. [Crossref]

[55] López, D. M., Rico-Olarte, C., Blobel, B., & Hullin, C. (2022). Challenges and solutions for transforming health ecosystems in low-and middle-income countries through artificial intelligence. *Frontiers in Medicine, 9*, 958097. [Crossref]

[56] Nair, M., Svedberg, P., Larsson, I., & Nygren, J. M. (2024). A comprehensive overview of barriers and strategies for AI implementation in healthcare: Mixed-method design. *PLoS One, 19*(8), e0305949. [Crossref]

Jiayi Yang received her Bachelor's degree in Microelectronics Science and Engineering from the School of Microelectronics, Shanghai University in 2024. She is currently pursuing her Master's degree in the Integrated Circuit Design program at the School of Engineering, Hong Kong University of Science and Technology. Her research achievements include advancements in device modeling, particularly in the field of Gallium Nitride (GaN) high electron mobility transistors. Additionally, she has gained substantial experience in circuit design, specializing in SRAM construction and Power Management Integrated Circuits (PMICs). (Email: jyangeg@connect.ust.hk)



Qian Hong is an Associate Chief Physician in the Department of Obstetrics and Gynecology at Jiangxi Provincial Fifth People's Hospital. She specializes in gynecological endocrinology and the diagnosis and treatment of common obstetric and gynecological diseases. Dr.Hong holds several professional appointments, including Standing Committee Member of the Reproductive Health Branch of the Jiangxi Provincial Society of Preventive Medicine, Standing Committee Member of the Women's Health Care Branch of the Jiangxi Provincial Research Hospital Association and Standing Committee Member of the Gynecological Vaginal Surgery Professional Committee of the Jiangxi Provincial Maternal and Child Health Care and Eugenics Association. (Email: qian.h@163.com)