Enhancing and Personalising Endometriosis Care with Causal Machine Learning *

Ariane Hine[1[0000−0002−7447−9371]], Thais Webber[2[0000−0002−8091−6021]], and Juliana Bowles[1,3[0000−0002−5918−9114]]

[1] School of Computer Science, University of St Andrews, KY16 9SX St Andrews, UK
{aah4,jkfb}@st-andrews.ac.uk
[2] School of Computer Science and Digital Tech., Aston University, B4 7ET, UK
t.webber@aston.ac.uk
[3] Software Competence Centre Hagenberg, Softwarepark, 4232 Hagenberg, Austria
juliana.bowles@scch.at

Abstract. Endometriosis poses significant challenges in diagnosis and management due to the wide range of varied symptoms and systemic implications. Integrating machine learning into healthcare screening processes can significantly enhance and optimise resource allocation and diagnostic efficiency, and facilitate more tailored and personalised treatment plans. This paper discusses the potential of leveraging patient-reported symptom data through causal machine learning to advance endometriosis care and reduce the lengthy diagnostic delays associated with this condition. The goal is to propose a novel personalised non-invasive diagnostic approach that understands the underlying causes of patient symptoms and combines health records and other factors to enhance prediction accuracy, providing an approach that can be utilised globally.

Keywords: Female reproductive health · Endometriosis · Artificial Intelligence · Prediction models · Diagnosis · Menstrual health

1 Introduction

Endometriosis is a prevalent, chronic, inflammatory condition affecting approximately 10% of individuals assigned female at birth during their reproductive years and beyond. Historically characterised as a gynaecological disease [7], recent research has revealed its systemic implications [16]. Endometriosis can affect virtually all organs in the human body, with symptoms ranging from asymptomatic cases to severe, life-altering conditions [14, 26].

Common symptoms include menstrual irregularities, heavy menstrual flow (menorrhagia), painful periods (dysmenorrhea), pain during sexual intercourse (dyspareunia), chronic pelvic pain unrelated to menstruation, tenderness, adnexal mass (growths near the uterus/ovaries), infertility or subfertility, depression, anxiety, abdominal bloating, nausea, and restricted mobility [10, 26]. In

* Bowles is partially supported by the Austrian Funding Council (FWF) under Meitner M 3338-N.

more advanced cases, patients often experience bowel and bladder-related symptoms such as painful bowel movements (dyschezia), loss of bladder control (dysuria), blood in the urine or stool during menstruation, painful urination, and chronic fatigue [24]. Given the multitude of symptoms, overlapping conditions, and the complexity of the disease, the root cause of endometriosis has not yet been conclusively determined [1, 10]. It has been hypothesised that one cause may be retrograde menstruation, where menstrual blood flows backwards into the pelvis during menstruation. Some of the menstrual blood contains endometrial tissue, which is believed to implant within a woman's abdomen, leading to patches of endometriosis [18]. This is not the sole cause of the condition, however, as 90% of all menstruating women are thought to experience this phenomenon. On the contrary, other sources suggest that endometriosis may start at birth, with symptoms not triggering until puberty [23]. There are many varying theories on the source of the condition, with no one definitive cause. The aetiology of endometriosis remains medically undetermined, complicated by its complex multifactorial nature. This complexity likely contributes to the challenges in understanding and diagnosing the condition, as multiple variables seem to influence its development and progression in patients [4, 15].

Further complexity arises when patients with endometriosis, whether suspected or diagnosed, have comorbidities that contribute to pelvic-related symptoms. These additional conditions must be considered within a comprehensive diagnostic algorithm. Common comorbidities include Irritable Bowel Syndrome (IBS), Interstitial Cystitis, and Adenomyosis [22], all of which present with similar pelvic symptoms. Their presence can complicate the clinical scenario, making accurate diagnosis more challenging and often resulting in significant delays in identifying the condition and starting appropriate treatment and management. This complexity emphasises the need for thorough diagnostic protocols and technological support for clinical decisions that account for conditions with overlapping symptoms. Integrating such considerations into diagnostic tools is crucial for enhancing diagnostic accuracy, supporting and ensuring that the diagnostic process reflects the multifaceted approach taken by clinicians in practice.

We believe that leveraging multiple diverse sources of health data (including patient-reported data) and cutting-edge machine learning (ML) techniques for analysis can reduce diagnostic delays by accurately identifying the underlying causes of patient symptoms [11]. Causal machine learning, which focuses on understanding and identifying the causal relationships between variables rather than just correlations, holds particular promise in this area. By applying causal ML techniques, we can move beyond traditional correlation-based models to uncover the underlying mechanisms of endometriosis. This can lead to more accurate diagnostic tools that not only predict the likelihood of the disease but also provide insights into the potential causes of patient symptoms.

The goal of the EndoML Project is to develop a novel personalised non-invasive diagnostic approach that utilises diverse health data, including patient-reported information, and applies advanced ML techniques to reduce diagnostic delays. By understanding the underlying causes of patient symptoms, combining health records with other relevant factors, and enhancing prediction accuracy, we aim to create a globally applicable diagnostic tool. Even if the EndoML Project has been specifically developed for targeting endometriosis, and the data collected focuses on that, the underlying (causal) machine learning approach can potentially shed light on more generic solutions to address the diagnosis of other conditions. Overall, such approaches can optimise healthcare resource allocation, for example, reducing unnecessary tests and procedures, improving the diagnostic process through timely and accurate identification of the condition (in this case, endometriosis), and facilitating the creation of shareable computational methods for personalised treatment.

This paper is structured as follows. Section 2 provides the context of our work and the need for digital solutions to address diagnostic delays in general, and endometriosis in particular. We discuss related work and research leveraging AI/ML techniques, including the shortcomings of existing attempts for algorithmic-based diagnosis of endometriosis. Section 3 presents the EndoML Project and the technical details underlying the proposed causal ML model, focusing on the prediction flow and introduction to the wide data collection through a comprehensive survey. Section 4 presents data collection approach and findings from initial data analysis, including data preprocessing and descriptive statistics. Finally, Section 5 concludes the paper with our aims for future work.

2 Endometriosis Care and Diagnosis Delays

One of the most challenging aspects clinicians face when seeing patients is to ensure the correct diagnosis of their conditions. This can be challenging when symptomatic manifestations can vary considerably between individuals, particularly in complex conditions such as endometriosis [1]. There are complex care pathways and diagnosis scenarios for endometriosis with delays evident from both patients' and clinicians' viewpoints (refer to Fig. 1 and Table 1). It is important to acknowledge that the list of delays is not exhaustive and is subject to further investigation [1, 11].

From clinicians' viewpoint, the complex nature of endometriosis, presenting varied symptoms and potential overlap with other conditions, brings challenges such as the lack of effective guidelines and integrated holistic health history. For instance, in this context, the average time to diagnosis in the UK is reported to range from 4 to 11 years [1]. From patients' viewpoint [1], for some individuals, endometriosis is accompanied by a barrage of symptoms that significantly impair their quality of life on a daily basis, including pain (through a variety of manifestations), and often infertility. Other patients may exhibit few to no symptoms, and as a result, diagnosis may occur unexpectedly during a secondary, unrelated medical procedure, leading to further tests to better understand the state of the patient's endometriosis. Some asymptomatic individuals are likely to remain undiagnosed but the same holds true for symptomatic sufferers, given the intricacies involved in the diagnostic process and potential misinterpretation

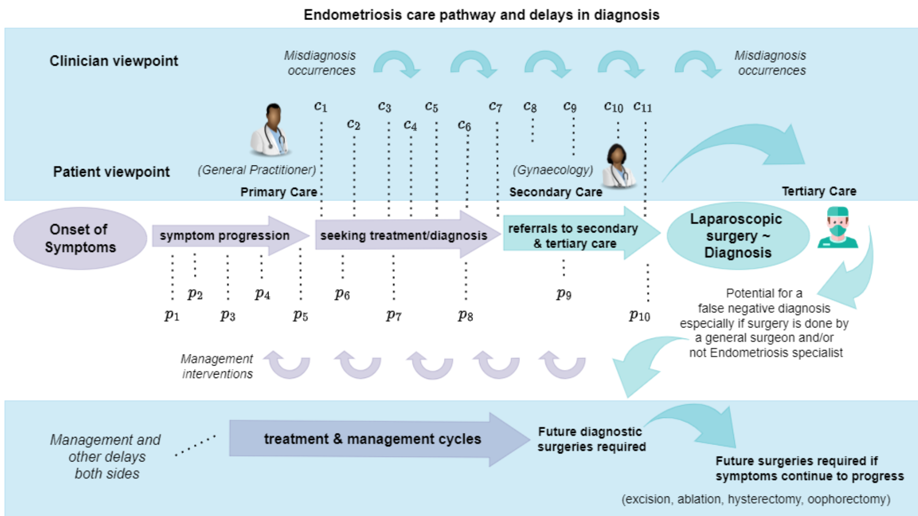


Fig. 1: Complex scenario of delays in endometriosis care (refer to Table 1).

Patient viewpoint

p_1	Lack of awareness
p_2	Normalisation of symptoms
p_3	(Limited) Access to care
p_4	Fear of judgement/being a burden
p_5	Downplaying symptoms
p_6	Difficulty recalling & communicating symptom history
p_7	Hormonal birth control masking symptoms
p_8	Long waiting list times for referrals from Primary Care
p_9	Reluctance to have invasive procedures performed
p_{10}	Long waiting list times in Secondary and Tertiary Care (Surgical interventions)

Clinician viewpoint

c_1	Lack of guidelines & holistic health history
c_2	Poor recognition of symptoms
c_3	Normalisation of symptoms
c_4	Somatisation of symptoms
c_5	Patient has non-specific symptom profile
c_6	Need to test for differential conditions
c_7	Poor collaboration between healthcare providers
c_8	Non-diagnostic imaging tests
c_9	First-line management of symptoms vs. diagnostic testing
c_{10}	Other factors masking symptoms
c_{11}	Reluctance to perform invasive procedures

Table 1: Contrasting viewpoints on diagnostic delays (in Fig. 1).

of clinical signs (refer to the clinicians' viewpoint misdiagnoses occurrences in Fig. 1).

In summary, the significant diagnosis delay is not solely due to clinical complexities but also stems from broader systemic challenges within healthcare delivery and patient awareness [6]. Both clinician and patient sources of delays are subject to societal-induced biases stemming from the historical minimisation, normalisation, and somatisation of sufferers' pain [1]; whether this delay directly or indirectly influences the outcome. Both are also impacted by the reluctance to perform invasive procedures, resulting in a longer waiting time for laparoscopic (gold standard) diagnosis.

Laparoscopic surgery is invasive, costly, and presents risks of potential damage to internal organs during surgical exploration. This is probably why, from the onset of symptoms and their progressive presentation, clinicians traditionally start the diagnosis process with physical examinations, periodic symptom tracking and management, followed by imaging modalities, tests (or assessments) for differential conditions, but mostly considering first-line management iterations modified over time based on individual's response or circumstance to suppress or alleviate the symptoms [22] (refer to management interventions illustration on Fig. 1). Moreover, even laparoscopic surgery has the potential for false negative results as illustrated on the right side of Fig. 1, culminating in more cycles of management interventions and further potential diagnostic surgeries. This risk is particularly high if surgery is performed by a general surgeon or a non-specialist in endometriosis. Such inaccuracies can lead to further patient interventions and more costly actions, as the underlying condition remains unidentified and inadequately managed (or insufficiently treated).

There is hence a need for digital solutions, algorithms, and tools that address the burden surrounding diagnostic difficulties, can support and reduce diagnostic delays and improve the quality of life for both confirmed and suspected endometriosis sufferers, especially those who are severely affected. Ultimately, these solutions would be highly beneficial when integrated into primary, secondary, and tertiary healthcare settings, potentially avoiding risky and costly invasive procedures [23, 27]. A personalised non-invasive diagnostic approach can become an essential tool to effectively triage patients for referrals to tertiary care, such as gynaecology, ensuring that those in need of specialised care are identified and treated more promptly.

2.1 Machine learning in endometriosis diagnosis and treatment

Machine learning (ML) has the potential to transform traditional approaches in healthcare diagnosis and treatment [27]. However, the integration of ML in healthcare remains challenging due to stringent requirements for data quality, privacy, and regulatory compliance [17]. As in any critical domain, developing healthcare AI/ML tools requires meticulous attention to detail and adherence to rigorous standards. In the last decade, research on algorithmic detection and treatment of endometriosis has gained attention due to the growing awareness and understanding of the condition among healthcare professionals, the general public, and increased patient advocacy [3, 8].

ML is set and suitable to optimise and increase efficiency in many areas of medical diagnostics and treatment, especially for conditions with complex or unknown etiologies such as endometriosis [23]. These areas range from improving the core understanding of a condition's manifestations, progression, and patient phenotypes to enhancing research and development towards improving processes and guidelines used by clinicians in their daily practice. This continuous improvement is supported by a feedback loop inherent in ML applications; the more data and results these systems integrate and analyse, the more refined and effective they increasingly become. This cyclical process of learning and adapting helps drive advancements in both theoretical knowledge and practical applications, effectively closing the loop of improvement. Fig. 2 lists our understanding of the varied focus and impact of AI/ML tools throughout the years in the context of endometriosis such as topics mentioned in [1, 11, 22, 25, 27], among other sources.

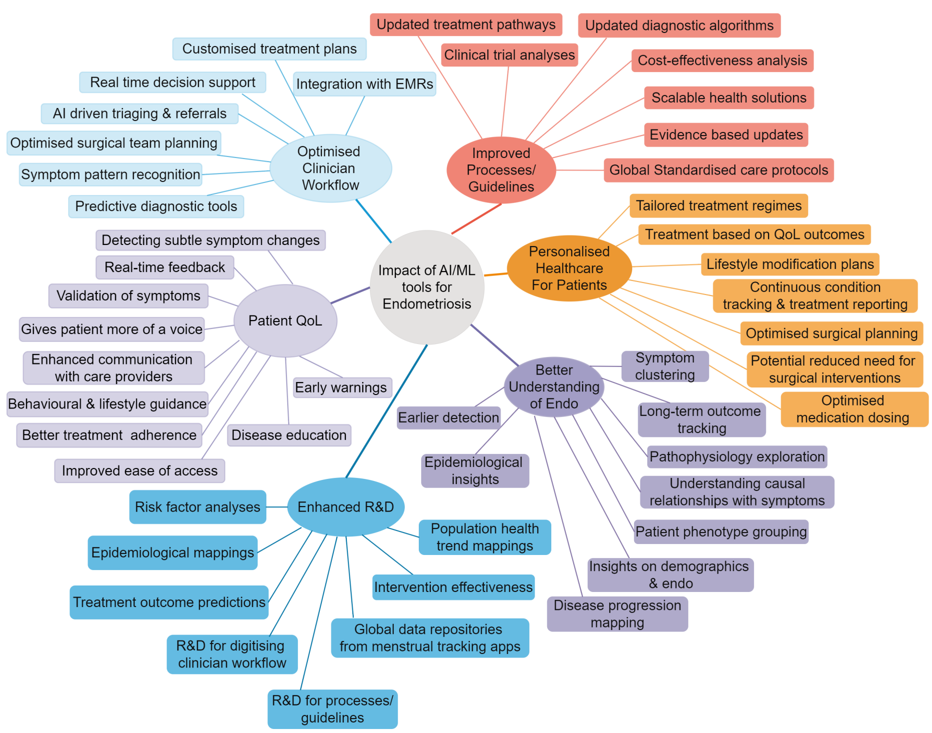


Fig. 2: Focus and impact of AI/ML algorithms and tools for endometriosis.

It is essential to remark there are a few relevant points related to the adequate choice of methods and approaches that focus on symptom information and a better understanding of endometriosis. A common theme within the published research on endometriosis diagnosis is the use of ML methods such as Regression analysis [5, 8], which look for data correlations without understanding deeper causal relationships between variables. Some approaches can have significant drawbacks, for instance, limited or inappropriate cohort selection (which can result in skewed outcomes and limited generalisability) and selection bias (due to strict inclusion/exclusion criteria, which can result in under-representation of certain subsets), failing to capture the disease's full heterogeneity [3, 5, 8]. These examples of limitations underscore the need for more sophisticated techniques that can handle diverse and complex data while providing deeper insights into causal mechanisms [19, 21].

In summary, capturing the full diversity of the endometriosis population is essential for developing accurate diagnostic tools that can be applied globally. By including a more representative sample of patients, these tools can better reflect the variety of symptoms and disease manifestations, leading to more precise and effective diagnostics and treatments. This comprehensive approach is relevant for improving patient outcomes and ensuring that diagnostic tools are effective across different demographics and healthcare settings [1].

While statistical prediction models for diagnosis offer considerable potential [20], it is crucial to advance our understanding of endometriosis and its nuanced behaviour through more advanced algorithmic techniques; moving beyond mere analysis of simple correlations [11]. In contrast to traditional ML methods (e.g., Decision trees, Gradient boosting, and AdaBoost) commonly applied in the literature [13, 20], causality-based techniques [19, 21] aim to uncover the diseases that cause symptoms, not just the statistical relationships to symptoms, which may be erroneous or unable to capture the complex existing causal relationships.

Causal ML specifically seeks to model and test hypothetical interventions, determining the direct effect of one variable over another through do-calculus performed on causal Bayesian networks [19]. This allows researchers to experiment with "what-if" scenarios, with insights on how altering certain factors could hypothetically impact the presence and severity of symptoms [11]. In a diagnostic context, the modelled scenario would be as follows: 'what if a patient did not have endometriosis; would their symptoms still be expected to persist?'. This shift towards prioritising causality over correlation is an important advancement in endometriosis research, where multiple overlapping symptoms and factors complicate diagnosis and treatment. Consequently, by distinguishing between correlation and causation, causal ML can help identify risk factors and implement effective early interventions, rather than merely associating symptoms with the disease. One predictor approach could prioritise interventions that identify which symptoms would persist if the condition were absent in a patient, hence accurately determining which symptoms are directly caused by endometriosis. Additionally, this approach holds promise to increase model accuracy and offer more targeted, effective diagnostic interventions.

3 The EndoML Project and Data Collection

Incorporating algorithmic tools to assist clinicians in diagnosing and prioritising treatment based on symptom profiles can decrease the number of supplementary investigative appointments and diagnostic tests, thereby alleviating diagnostic delays [1, 4, 9, 18, 27]. A significant portion of the delay in diagnosing endometriosis stems from healthcare professionals' uncertainty about a definitive diagnosis, which is particularly concerning given the condition's progressive nature [2].

The EndoML Project aims to minimise the need for invasive diagnostic procedures, such as laparoscopic surgery, blending patient-generated data from multiple sources and (causal) machine learning (ML) techniques. In addition, establishing the most likely underlying cause of a patient's symptoms can increase clinicians' confidence in promptly referring patients to specialised care teams, promoting efficient healthcare provisioning and optimised resource allocation. There is a noticeable gap in clinical workflows that ML tools could strengthen to enhance diagnostic efficiency, for instance, by automating patient data analysis, recognising patterns and abnormalities in patient data, and providing informed clinical decision support.

The ultimate goal is to develop an ML tool that is easily interpretable, and adaptable to new and evolving knowledge, providing comprehensive insights into variables and their relationships, rather than solely focusing on correlations. The premise is that causal ML techniques can result in more accurate diagnostic tools that can also provide valuable insights into the potential causes of a patient's symptoms [19, 21]. In practice, causal ML modelling involves key steps:

- **Data collection:** patient data must be collected, including variables related to the condition of interest, potential causal and risk factors, and treatment outcomes. The data collection phase should be comprehensive, aiming to integrate diverse and detailed patient information. It is important to balance data input from clinical assessments with self-reported patient data to capture a full picture of the condition's impact. This dual approach facilitates a deeper understanding of the condition from both medical and patient-oriented perspectives. Fig. 3 depicts the aspects related to data collection, with both clinician data, patient self-reported data, and their potential overlap. It stresses the characteristics, benefits, challenges and limitations related to different data collection perspectives. This overview presents the various data collection methods, illustrating examples of clinician data and potential patient self-reports, along with the common variables each source offers. It also highlights a few key advantages and disadvantages of each method, providing insights to consider when selecting data sources.

- **Graph construction:** Directed Acyclic Graph (DAG) construction, where potential causal variables are identified based on the collected data, domain knowledge and prior research. This step is crucial as it lays the groundwork for understanding the underlying mechanisms of the condition. This step also involves the integration of expert knowledge into the DAG structure to

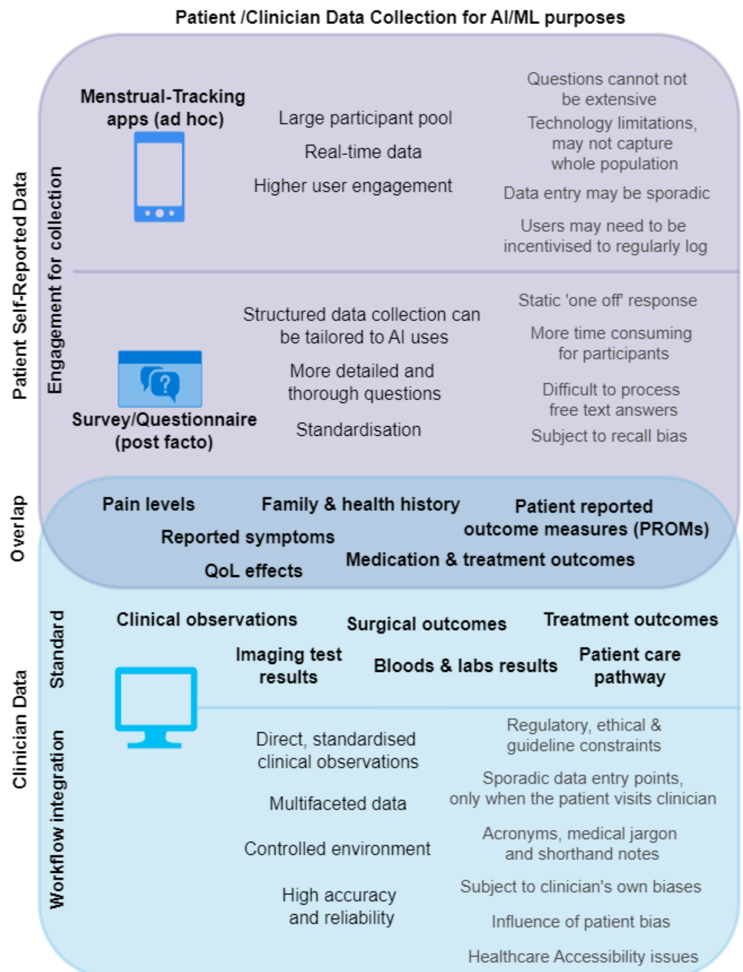


Fig. 3: Comparing clinical and patient self-reported data.

ensure that it accurately represents the hypotheses about causal pathways and reflects both direct and indirect relationships among variables. Through iterative refinement and updating, the DAG becomes a dynamic tool upon which we perform simulated interventions based on established causal links.

– **Causal Inference:** causal inference techniques are employed to estimate the causal effects of variables on the outcome of interest. Structures such as causal Bayesian networks [21] can model these relationships. This step involves the examination of hypothetical scenarios through counterfactual reasoning, which helps predict what would happen to the outcome if a variable were altered while controlling for other factors. This provides a solid foundation for making informed decisions about treatment strategies and diagnostic outcomes when used in a clinical setting.

– **Evaluation:** causal ML model must undergo rigorous evaluation and results interpretation. Appropriate metrics, such as accuracy, precision, recall, and F1-score, are used to evaluate the model's ability to make effective and reliable predictions for new data. Cross-validation techniques are also employed to validate the model's performance and to prevent overfitting. Involving clinicians in the validation process is crucial to ensure the practical applicability and clinical relevance of the model. This step incorporates expert feedback to fine-tune the model.

– **Updating:** causal ML model is continuously updated with new data and findings. As more patient data becomes available, the model parameters and structures are revised to incorporate the latest information. The updating process ensures that the model remains current and accurate over time. Regular updates also involve re-evaluating the model's performance metrics and making necessary adjustments to improve prediction accuracy. This step is vital for maintaining the model's relevance and for adapting to new information. In addition to routine model updates, ongoing surveys and novel data collection methods should be explored to constantly enrich the dataset. Future efforts may focus on integrating diverse health data streams, including real-time patient monitoring and digital health records, to continuously refine and enhance the predictive capabilities. These initiatives are crucial for capturing a broader spectrum of patient experiences and for tailoring interventions more effectively.

– **Deployment:** once validated, the model can be deployed in a clinical setting, where it can assist healthcare providers in making informed diagnostic and treatment decisions. The deployment phase involves integrating the model into existing healthcare systems, training clinicians on its use, and establishing protocols for its application in clinical practice.

– **Maintenance:** post-deployment, the causal ML model must be continuously monitored for performance and accuracy, leveraging ongoing feedback from clinicians and patients to identify areas for improvement. Regular maintenance will ensure the model adapts to new data and evolves according to clinical practice and emerging research findings, sustaining its effectiveness and utility in real-world settings.

Finally, Fig. 4 shows the proposed prediction modelling flow for endometriosis. The model input is global patient-collected survey data, which undergoes a series of stages, from data preparation and cleaning to their utilisation within the causal ML algorithms via the causal graph. The prediction modelling flow leads to probabilities as output that can be leveraged to estimate the likelihood (either positive or negative) of endometriosis or differential diagnoses in new

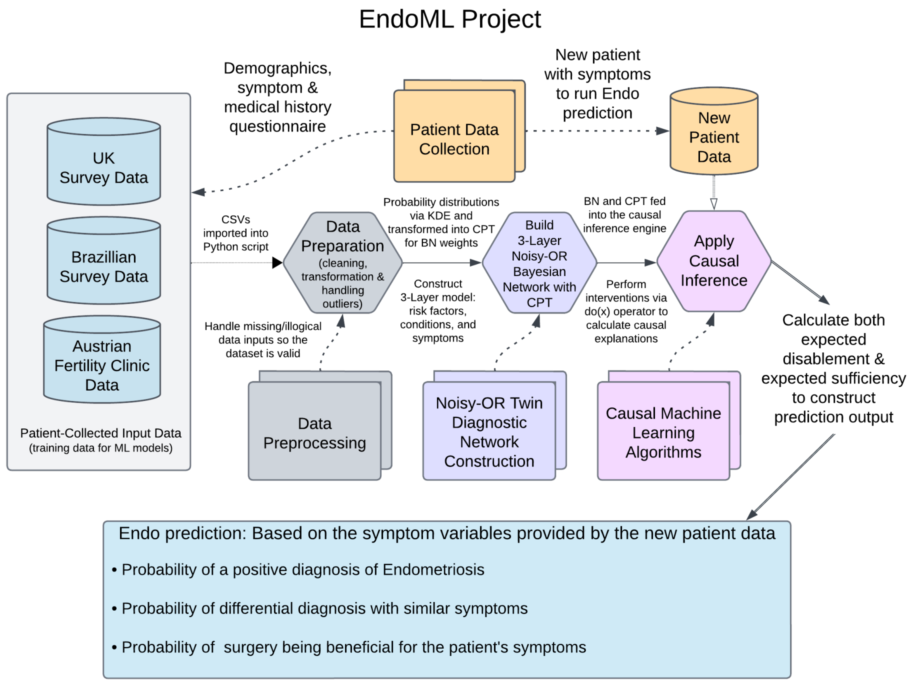


Fig. 4: Overview of EndoML Project and its proposed prediction modelling flow.

patients exhibiting symptoms. Future refinements could involve tailoring output predictions to further identify the most effective treatment pathways to match the specific manifestations of symptoms in patients.

Causal machine learning emerges as a promising research direction, upgrading existing Bayesian Network-based solutions by incorporating causal reasoning into the modelling process [21]. The hypothesis is that we can overcome the limitations of traditional statistical-based ML approaches, which are primarily correlative, integrating causal inference methodologies. Unlike traditional methods, the Noisy-OR Twin Bayesian Networks [19, 21] allow us to model complex causal relationships among disease and symptom variables, providing a deeper understanding of patient-reported data. The term “noisy-OR” refers to the extension of traditional Bayesian Networks (BNs) by introducing uncertainty via ‘noisy’ nodes and logical OR operations. The term ‘twin’ refers to the two different versions of the world that the model encapsulates, the factual world and the counterfactual world (where ‘interventions’ are performed). Introducing additional uncertainty and complexity allows us to model more realistic scenarios with incomplete or uncertain data [21]. Furthermore, causal inference uncovers these causal relationships by simulating interventions on diseases via the mathematical ‘`do(x)`’ operator [21]. These interventions demonstrate the strength of the relationship between a condition being ‘cured’ or ‘switched off’ and the likelihood of a symptom also being 'switched off' as a result [19, 21]. This logic leads to typically stronger predictions that account for temporal knowledge and underlying mechanisms of associations between variables, facilitating more accurate and reliable predictions, thus possibly increasing chances of clinician adoption.

It should be emphasised that the EndoML Project prioritises a patient-centric approach, highlighting the importance of understanding and addressing the unique experiences and needs of individuals affected by endometriosis. Utilising patient-reported symptom data allows the patient to feel empowered by proactively participating in their healthcare journey, despite one of the most prominent challenges being data quality. Examples are the impact of recall bias and the reduced ability to control confounders in causal ML techniques [12, 21]. The potential benefits of the approach still outweigh these risks, as healthcare providers can customise interventions based on individual symptom profiles for faster and more effective treatment outcomes.

Our data collection methodology involves a thorough questionnaire applied across three countries (UK, Brazil, Austria) based on the EHP-30 survey and influenced by the intake form used in an Austrian Fertility clinic. This survey covers an array of variables spanning demographics, menstrual history, family background, fertility, sexual health, surgical history, and contraceptive usage, to cover the various symptoms experienced by women. The selection of a developed country with a national healthcare system (UK), a developed country with a predominant reliance on private healthcare (Austria), and a developing country (Brazil) was strategic. This choice aimed at capturing socioeconomic and geographical differences to include diversity in the input data and, hence, build a potentially stronger predictor. It is important to note that although we selected social media support groups in the countries of interest (UK and Brazil), we encouraged the distribution of this survey link outside of these groups to gain a larger respondent pool and capture a larger subset of the population. From the UK Survey collection, we reached a total of 475 entries with 227 complete valid responses from the UK specifically but we are aware of respondents from outwith the UK (except from Brazil/South America).

Fig. 5 details the geographical distribution of respondents of the UK Survey (2024) per country. Although a vast majority of respondents were from the UK, a large proportion of responses (91) came from the United States. Australia was the third most frequent country of response with 14 valid questionnaires being submitted. This distribution indicates not only a wide interest in the topic but also reflects the effectiveness of our outreach strategy in engaging a diverse international audience even though this was not the original intended population. In total, including the UK, the survey reached 28 countries so far which highlights the global interest in this work and the willingness of sufferers to share their experience to improve endometriosis research. Once collected, the survey data is cleaned and treated, automatically analysed by a Python script, which generates probability distributions for each variable, through Kernel Density Estimation (KDE) for continuous variables, and builds custom predictions for distinct variables. The resulting distributions are then transformed into a Conditional

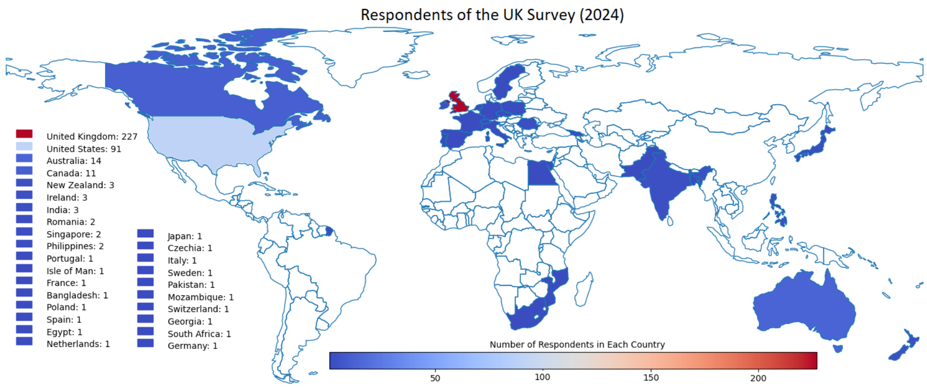


Fig. 5: Geographical distribution of respondents of the UK Survey per country.

Probability Table (CPT) to weight the BN (please refer to Fig. 4). Free text inputs may undergo analysis via NLP techniques to uncover relevant responses that can also add knowledge to the ML prediction model. However, in practical applications, addressing challenges during data preprocessing becomes crucial, especially when managing the diverse quality of free text inputs and formulating robust strategies to tackle ambiguous or context-dependent language [12].

Finally, the output of the causal inference algorithm, determined by the expected disablement and expected sufficiency calculations [19, 21], is instrumental in determining whether a patient's symptoms likely indicate endometriosis, based on diverse data (i.e., from the three countries) and the causal ML model predictions. Expected disablement quantifies the estimated degree of impairment caused by endometriosis-associated symptoms. This involves the causal inference algorithm's analysis of collected data, considering the severity of symptoms and their impact on the individual's well-being. Expected sufficiency reflects the likelihood that the observed symptoms suggest a positive endometriosis diagnosis by evaluating symptom patterns and their associations. Both calculations can be used interchangeably to predict an individual's probability of diagnosis [21].

Therefore, in addition to upgrading existing Bayesian network-based solutions, this approach focuses on collecting diverse data from multiple sources. This strategy is essential for designing a highly accurate diagnostic tool capable of understanding more comprehensively the intricacies of endometriosis. The data diversity ensures that the resulting tool remains applicable, shareable, and usable across various geographical regions and demographic groups. This also brings the benefit of bias mitigation that may arise from limited or homogeneous datasets [2], thereby enhancing the generalisability and reliability of our diagnostic tool, increasing the probability of adoption into clinical workflows.

4 Preliminary Insights

The EndoML Project data collection and preprocessing have commenced in preparation for the application of the causal ML algorithms. Initial data analysis focuses on understanding the demographic and clinical characteristics of the survey respondents, as well as identifying key patterns and variances in their reported symptoms and medical history. This preliminary analysis serves as a foundational step towards building robust ML models that can accurately diagnose and provide personalised treatment recommendations for endometriosis. Thus, in this section, we present findings from initial data analysis, including descriptive statistics and visual data representations.

4.1 Data collection & survey key areas

We orchestrated and implemented an extensive online survey to collect diverse data concerning endometriosis and women's health perceptions. The EndoML Project Survey was distributed through online social media support channels in two countries (UK and Brazil), whilst in the Austrian fertility clinic a consolidated version was utilised as an intake form for new patients. Furthermore, although the survey questions remain consistent across all data collection clusters, translation is necessary to accommodate participants who may be more comfortable responding in their native language, thereby increasing accessibility and reducing potential misinterpretations.

The survey includes diverse variables such as demographics, menstrual history, family background, fertility, sexual health, surgical history, and contraceptive use, alongside gathering individuals' perceptions on different aspects of the condition. The goal is to integrate a comprehensive dataset with key variables of the condition, providing detailed and high-quality information to inform the causal ML model. Following is the summary of these key areas covered by the survey and the insights they offer:

- Demographics
  - *Coverage:* age, country, ethnicity, education, occupation.
  - *Insights:* identifies demographic patterns and highlight potential biases.

- Diagnosis & History
  - *Coverage:* age at diagnosis, diagnostic methods, family history.
  - *Insights:* highlight current diagnostic delays and genetic predispositions.

- Symptoms Descriptions
  - *Coverage:* presence and severity of symptoms like menstrual irregularities, pelvic pain, and pain during intercourse.
  - *Insights:* essential for identifying symptom patterns and variations, which aids in making accurate predictions.

– Cycle & Bleeding
  - *Coverage:* menstrual cycle regularity, duration, bleeding intensity.
  - *Insights:* indicators of menstrual abnormalities, disease/condition presence, severity and progression.

– Pain Characteristics
  - *Coverage:* type, frequency, and intensity of pain during menstruation, ovulation, and intercourse.
  - *Insights:* differentiates between pain severity levels, manifestations, pain sensations, essential for predictive accuracy.

– Surgeries
  - *Coverage:* surgical history, types, and outcomes.
  - *Insights:* details surgical interventions and their effect on symptom relief and disease progression.

– Bowel & Bladder
  - *Coverage:* symptoms like pain during urination (or defecation); frequency, severity.
  - *Insights:* on the multi-system impact of endometriosis, not only reproductive symptoms; it can assist in identifying advanced stages (III/IV).

– Pregnancies
  - *Coverage:* pregnancy history, conception difficulties, outcomes.
  - *Insights:* highlights impacts on fertility and pregnancy outcomes; it can be used to help guiding treatment plans.

– Birth Control
  - *Coverage:* use of birth control, types, effectiveness.
  - *Insights:* evaluates symptom management effectiveness through birth control methods.

The diverse responses enabled by the UK Survey (2024) allow us to capture the nuances of the heterogeneous nature of endometriosis, with significant variance in symptom profiles and women's experiences. This variability is critical for developing a robust model that provides personalised diagnostic and treatment recommendations.

4.2 Data preprocessing & descriptive statistics

Building on the understanding gained from the diverse questions and responses, this section presents descriptive statistics derived from the collected survey data. A natural first step (please refer to Fig. 4) is to prepare the data, performing tasks such as cleaning (e.g., removing empty responses, correcting data entry errors or inconsistencies), transforming (i.e., coding responses, standardising data, etc.) and handling outliers (i.e., excluding or adjusting them).

An example of a common error made by respondents, among others, was entering their weight in pounds instead of kilograms as requested. To account for this, we established a threshold within the data to identify instances where respondents likely misunderstood the input format. Then, we manually converted these presumed pound values to their kilogram equivalents. This issue likely occurred because the questionnaire link was shared beyond the UK, leading to numerous responses from participants non-UK residents.

Following, we present some results from the UK Survey data, the characteristics of the average respondent, the variance observed in the data distributions of some variables and a preliminary data interpretation. Initial data analysis has laid the groundwork for our causal ML algorithms, which aim to leverage the gathered insights to enhance endometriosis diagnosis and treatment.

Average Respondent. In terms of demographic characteristics, the average respondent age (considering a total of 371 valid responses out of 475) was 33.3 years old, with a standard deviation of 8.38, indicating a diverse age range among participants. Additionally, the distribution of educational backgrounds varied, with the majority of respondents holding a professional degree or equivalent qualification. These findings underscore the heterogeneous nature of the survey sample and emphasise the importance of considering individual differences in developing personalised diagnostic and treatment strategies for endometriosis.

Variance in Data Distributions. A preliminary analysis revealed significant variance in key variables related to endometriosis, reflecting the diverse experiences of individuals with the condition. For instance, the distribution of diagnosis time (in months) exhibited a wide range, with some individuals experiencing markedly delayed diagnoses (refer to the histograms in Fig. 6).

Similarly, the distribution of the number of days bleeding during a menstrual period showed considerable variability, with some respondents reporting very long durations whilst others reporting none at all (Fig. 7a); also distribution of age at menarche or first period (Fig. 7b).

These data distributions highlight the complexity of endometriosis and underscore the need for tailored approaches to diagnosis and management. To visually represent these variations, we present histograms, bar charts, and box plots, depicting the distribution and variance of some key variables, providing insights into the heterogeneity of experiences among individuals with endometriosis.

Overall, the descriptive statistics and variance analysis presented here offer valuable insights into the symptomatology and diversity of experiences associated with endometriosis. These findings serve as a foundation for further research and the development of personalised diagnostic and treatment algorithms aimed at improving outcomes for individuals affected by this condition.

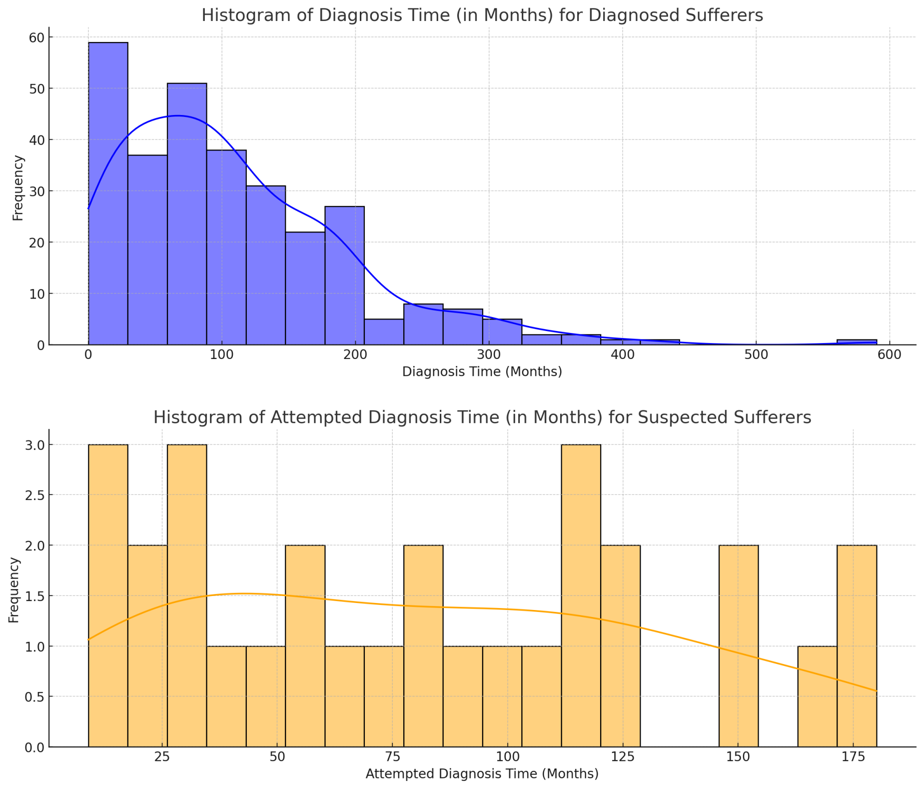
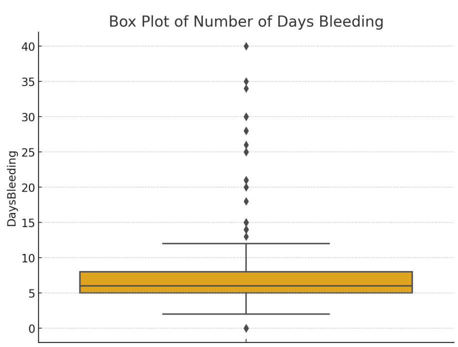
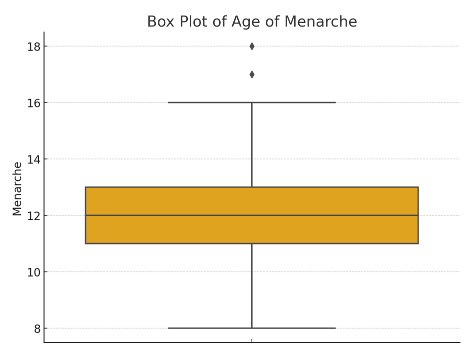


Fig. 6: Distribution of diagnosis time (in months) experienced by suspected and diagnosed endometriosis sufferers.



(a) Menstrual cycle length.



(b) Age at menarche.

Fig. 7: Distribution of menstrual cycle length and age of first period (menarche) experienced by endometriosis sufferers.

Preliminary Data Interpretation. We illustrate in this paper some insights through visual representations. For instance, histograms depict the distribution of diagnosis time in months for both diagnosed and suspected sufferers (refer back to Fig. 6). It is evident from these representations the significant delay experienced by both groups, which emphasises a concerning gap in healthcare provision quality. Preliminary analysis reveals significant variance also in other key variables, indicating the diverse experiences of individuals with endometriosis. For example, the duration of the menstrual cycle and the age at menarche (Fig. 7) show wide ranges, suggesting that endometriosis can manifest differently among individuals. In addition, the prevalence and variation of common symptoms highlight the importance of personalised varied approaches.

The insights derived from the UK Survey data serve as valuable input for developing the causal ML approach, allowing us, for example, to understand why certain diagnostic outcomes occur and why certain interventions are more effective. Our aim is to develop more tailored and effective diagnostic outcomes for individuals with endometriosis.

Another histogram (Fig. 8) illustrates the distribution of a key variable, the menstrual pain intensity, providing insights into the severity experienced by respondents. Additional box plots (Fig. 9) present the distribution of menstrual versus non-menstrual pain intensity (severity), providing a comprehensive overview of the dataset in relation to the variable ‘pain’.

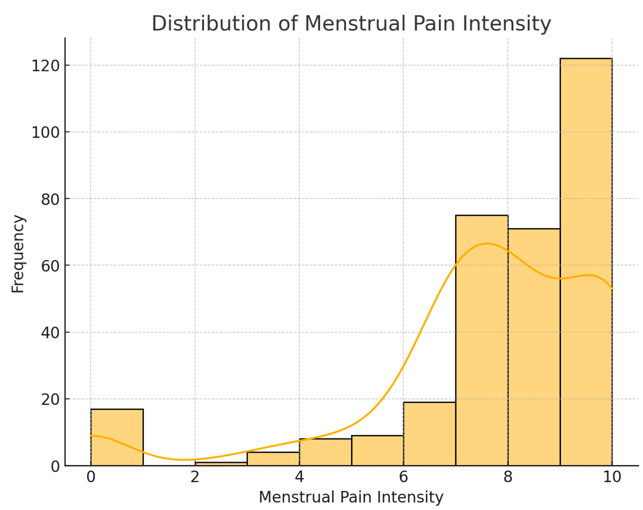
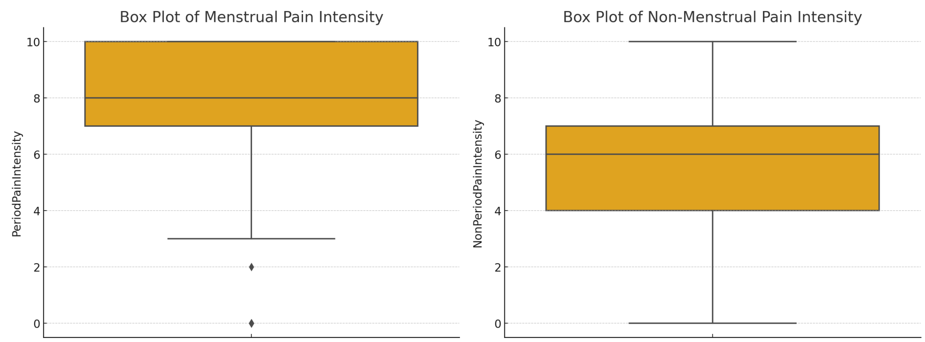


Fig. 8: Intensity of menstrual pain experienced by endometriosis sufferers.

Fig. 10 presents a bar chart displaying the frequency of comorbidities commonly linked with endometriosis, shedding light on its potential impact on other health conditions and the diagnostic challenges that may occur due to comor-



(a) Pain during menstruation.

(b) Pain outwith menstruation.

Fig. 9: Distribution of pain intensity during and outwith menstruation.

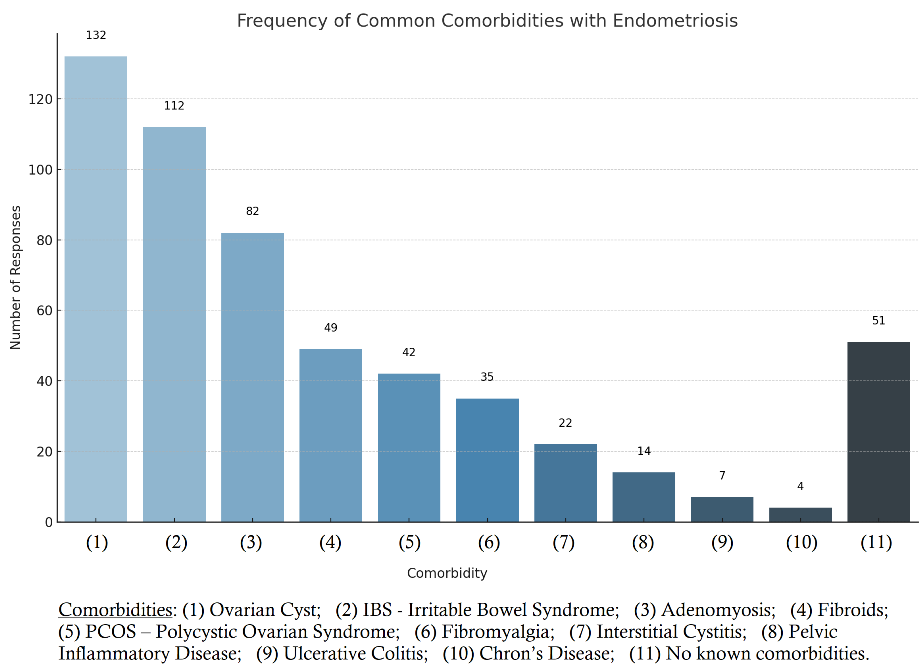


Fig. 10: Most frequent comorbidities reported by endometriosis sufferers.

bidities. Furthermore, Fig. 11 shows a bar chart highlighting the different types of pain experienced by individuals with suspected or diagnosed endometriosis who participated in the UK Survey, demonstrating the multifaceted nature of the condition. The data reveals common pain descriptors among sufferers: 281 respondents characterised their pain as 'stabbing' and 271 as 'aching'. The term

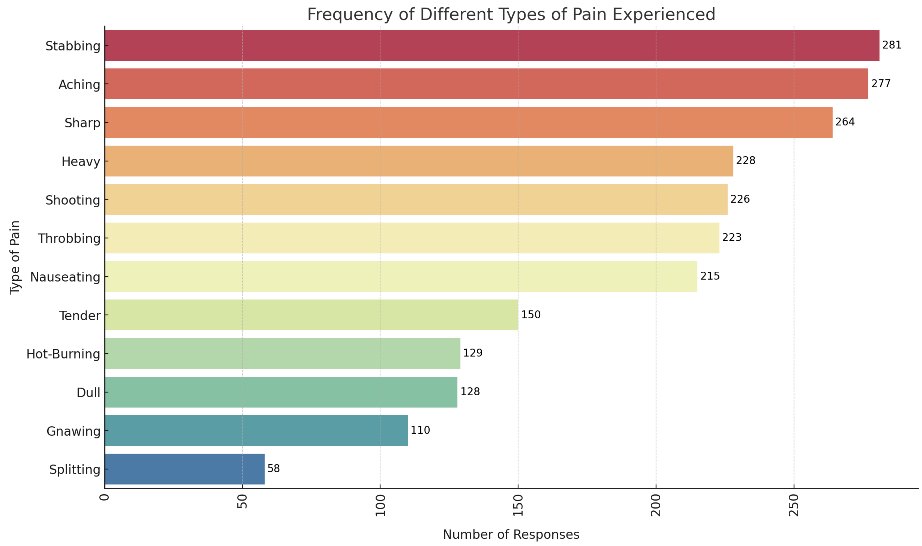


Fig. 11: Types of pain experienced by endometriosis sufferers.

'splitting' was least frequently used, with only 58 responses, suggesting it may be a less common descriptor. This variation in pain descriptions not only reflects the different pain manifestations in endometriosis but also highlights that not all individuals experience pain in the same way [5, 10].

The visual representations provided complement our descriptive statistics, enhancing our understanding of the data and reinforcing the need for personalised approaches in diagnosing and treating endometriosis. They also serve as valuable tools for communication, enabling stakeholders' understanding of the multifaceted nature and diverse manifestations of the condition. The data highlights that personalised approaches are vital for managing the heterogeneity of endometriosis. Future research should focus on integrating machine learning with clinical outcomes to develop dynamic models that can predict individual treatment responses.

The next steps in our analysis process will involve the implementation of the Noisy-OR twin diagnostic network and counterfactual algorithm in order to make predictions on new users' probability of endometriosis, based on the knowledge gained from the diverse data sources and the expertise of teams of clinicians. There is a challenge ahead to ensure the EndoML predictor is accurate considering the complexities of real-world data. We understand that data quality is critical to the solution, as any errors or biases can affect the model's performance. We may refine our data integration approach to gather more varied and targeted information, enhancing the completeness of our dataset and strengthening our model's predictive power.

5 Conclusion

Causal machine learning algorithms for diagnostics show promise in providing accurate predictions that could in the future be used to aid clinicians in understanding the root cause of patient symptoms in complex cases. Earlier identification of symptoms may reduce diagnostic delay and indirectly also promote a more sustainable use of resources. Understanding the most likely cause of symptoms facilitates the development of tailored healthcare frameworks with worldwide benefits, particularly in the case of conditions with varied manifestations, as with endometriosis. Ultimately, the goal is to reduce the pain and suffering of patients, moving beyond traditional management approaches and the recurrent use of medications and surgeries, delaying diagnosis and effective treatment. By refining diagnostic accuracy, we can offer a more efficient, less invasive diagnostic pathway that seeks to identify the root causes of symptoms caused by conditions like endometriosis, leading to better patient outcomes and quality of life.

Innovations in health informatics, like causal ML, offer potential for optimal diagnosis and treatment pathways, however, challenges such as data quality, biases from self-reported patient data, and clinical adoption of AI-based tools need addressing. Despite these obstacles, future research should prioritise the development of user-friendly diagnostic tools based on causal decision-making algorithms for both clinicians and patients. In this context, 'user-friendly' means providing interfaces that simplify the complex outputs of causal decision-making algorithms into actionable insights. This ensures accessibility without requiring deep technical knowledge. This approach can thus facilitate further potential exploitation of underlying AI/ML algorithms in critical scenarios.

Implementing such tools for endometriosis — a condition marked by complex symptoms and diagnostic challenges — establishes a precedent for transforming the diagnostic processes for other multifactorial diseases in similarly causality-focused ways. Such advancements could significantly improve the diagnostic process and patient experience across various conditions, demonstrating the broad relevance and applicability of these algorithmic tools.

Acknowledgement

We thank Thomas Ebner from the Kinderwunsch Zentrum, Kepler Universitätsklinikum, Linz, for valuable insights into the processes in the Austrian healthcare system. We thank the team from the School of Nursing, University of São Paulo, Brazil, led by Lislaine Aparecida Fracolli with Ana Luiza Vilela Borges, Carla Marins Silva and Marlise de Oliveira Pimentel Lima, for ongoing discussions on the Brazilian perspective of endometriosis care.

References

1. Agarwal, S.K., Chapron, C., Giudice, L.C., Laufer, M.R., Leyland, N., Missmer, S.A., Singh, S.S., Taylor, H.S.: Clinical diagnosis of endometriosis: a call

to action. American journal of obstetrics and gynecology **220**(4), 354–e1 (2019). https://doi.org/10.1016/j.ajog.2018.12.039

2. Ballweg, M.L.: Impact of endometriosis on women's health: comparative historical data show that the earlier the onset, the more severe the disease. Best practice & research Clinical obstetrics & gynaecology **18**(2), 201–218 (2004). https://doi.org/10.1016/j.bpobgyn.2004.01.003
3. Blass, I., Sahar, T., Shraibman, A., Ofer, D., Rappoport, N., Linial, M.: Revisiting the risk factors for endometriosis: A machine learning approach. Journal of Personalized Medicine **12**(7), 1114 (2022). https://doi.org/10.3390/jpm12071114
4. Burton, C., Iversen, L., Bhattacharya, S., Ayansina, D., Saraswat, L., Sleeman, D.: Pointers to earlier diagnosis of endometriosis: a nested case-control study using primary care electronic health records. British Journal of General Practice **67**(665), e816–e823 (2017). https://doi.org/10.3399/bjgp17x693497
5. Conroy, I., Mooney, S.S., Kavanagh, S., Duff, M., Jakab, I., Robertson, K., Fitzgerald, A.L., Mccutchan, A., Madden, S., Maxwell, S., et al.: Pelvic pain: what are the symptoms and predictors for surgery, endometriosis and endometriosis severity. Australian and New Zealand Journal of Obstetrics and Gynaecology **61**(5), 765–772 (2021). https://doi.org/10.1111/ajo.13379
6. Cox, H., Henderson, L., Andersen, N., Cagliarini, G., Ski, C.: Focus group study of endometriosis: Struggle, loss and the medical merry-go-round. International journal of nursing practice **9**(1), 2–9 (2003). https://doi.org/10.1046/j.1440-172x.2003.00396.x
7. Eskenazi, B., Warner, M.L.: Epidemiology of endometriosis. Obstetrics and gynecology clinics of North America **24**(2), 235–258 (1997). https://doi.org/10.1016/s0889-8545(05)70302-8
8. Fauconnier, A., Drioueche, H., Huchon, C., Du Cheyron, J., Indersie, E., Candau, Y., Panel, P., Fritel, X.: Early identification of women with endometriosis by means of a simple patient-completed questionnaire screening tool: a diagnostic study. Fertility and Sterility **116**(6), 1580–1589 (2021). https://doi.org/10.1016/j.fertnstert.2021.07.1205
9. Goldstein, A., Cohen, S.: Self-report symptom-based endometriosis prediction using machine learning. Scientific Reports **13**(1), 5499 (2023). https://doi.org/10.1038/s41598-023-32761-8
10. Gordon, H.G., Mooney, S.S., Conroy, I.C., Grover, S.R.: When pain is not the whole story: Presenting symptoms of women with endometriosis. Australian and New Zealand Journal of Obstetrics and Gynaecology **62**(3), 434–438 (2022). https://doi.org/10.1111/ajo.13482
11. Hine, A.A., Kuster Filipe Bowles, J., Webber, T.: The need for a non-invasive technology for endometriosis detection and care. Caring is sharing-exploting the value in data for health and innovation (2023). https://doi.org/10.3233/shti230073
12. Iroju, O.G., Olaleke, J.O.: A systematic review of natural language processing in healthcare. International Journal of Information Technology and Computer Science **8**, 44–50 (2015). https://doi.org/10.5815/ijitcs.2015.08.07
13. Kleczyk, E.J., Peri, A., Yadav, T., Komera, R., Peri, M., Guduru, V., Amirtharaj, S., Huang, M.: Predicting endometriosis onset using machine learning algorithms. BMC Women's Health (preprint under review) (2020). https://doi.org/10.21203/rs.3.rs-135736/v1
14. Kruk, M.E., Gage, A.D., Arsenault, C., Jordan, K., Leslie, H.H., Roder-DeWan, S., Adeyi, O., Barker, P., Daelmans, B., Doubova, S.V., et al.: High-quality health systems in the sustainable development goals era: time for a revolution. The

Lancet global health **6**(11), e1196–e1252 (2018). https://doi.org/10.1016/s2214-109x(18)30386-3

15. Malvezzi, H., Marengo, E.B., Podgaec, S., Piccinato, C.d.A.: Endometriosis: current challenges in modeling a multifactorial disease of unknown etiology. J. Transl. Med. **18**, 1–21 (2020). https://doi.org/10.1186/s12967-020-02471-0
16. Markham, S.M., Carpenter, S.E., Rock, J.A.: Extrapelvic endometriosis. Obstetrics and gynecology clinics of North America **16**(1), 193–219 (1989)
17. McKee, M., Wouters, O.J.: The Challenges of Regulating Artificial Intelligence in Healthcare: Comment on "Clinical Decision Support and New Regulatory Frameworks for Medical Devices: Are We Ready for It?-A Viewpoint Paper". International journal of health policy and management **12** (2023). https://doi.org/10.34172/ijhpm.2022.7261
18. Nnoaham, K.E., Hummelshoj, L., Kennedy, S.H., Jenkinson, C., Zondervan, K.T.: World endometriosis research foundation women's health symptom survey consortium. developing symptom-based predictive models of endometriosis as a clinical screening tool: results from a multicenter study. Fertil Steril **98**(3), 692–701 (2012). https://doi.org/10.1016/j.fertnstert.2012.04.022
19. Pearl, J.: Causality. Cambridge university press (2009). https://doi.org/10.1017/CBO9780511803161
20. Pergialiotis, V., Pouliakis, A., Parthenis, C., Damaskou, V., Chrelias, C., Papantoniou, N., Panayiotides, I.: The utility of artificial neural networks and classification and regression trees for the prediction of endometrial cancer in postmenopausal women. Public Health **164**, 1–6 (2018). https://doi.org/10.1016/j.puhe.2018.07.012
21. Richens, J.G., Lee, C.M., Johri, S.: Improving the accuracy of medical diagnosis with causal machine learning. Nature communications **11**(1), 3923 (2020). https://doi.org/10.1038/s41467-020-17419-7
22. Rogers, P.A., D'Hooghe, T.M., Fazleabas, A., Gargett, C.E., Giudice, L.C., Montgomery, G.W., Rombauts, L., Salamonsen, L.A., Zondervan, K.T.: Priorities for endometriosis research: recommendations from an international consensus workshop. Reproductive sciences **16**, 335–346 (2009). https://doi.org/10.1177/1933719108330568
23. Rolla, E.: Endometriosis: advances and controversies in classification, pathogenesis, diagnosis, and treatment. F1000Research **8** (2019). https://doi.org/10.12688/f1000research.14817.1
24. Sankaravadivel, V., Thalavaipillai, S.: Symptoms based endometriosis prediction using machine learning. Bulletin of Electrical Engineering and Informatics **10**(6), 3102–3109 (2021). https://doi.org/10.11591/eei.v10i6.3254
25. Shafrir, A.L., Farland, L., Shah, D., Harris, H., Kvaskoff, M., Zondervan, K., Missmer, S.: Risk for and consequences of endometriosis: a critical epidemiologic review. Best practice & research Clinical obstetrics & gynaecology **51**, 1–15 (2018). https://doi.org/10.1016/j.bpobgyn.2018.06.001
26. Simoens, S., Dunselman, G., Dirksen, C., Hummelshoj, L., Bokor, A., Brandes, I., Brodszky, V., Canis, M., Colombo, G.L., DeLeire, T., et al.: The burden of endometriosis: costs and quality of life of women with endometriosis and treated in referral centres. Human reproduction **27**(5), 1292–1299 (2012). https://doi.org/10.1093/humrep/des073
27. Sivajohan, B., Elgendi, M., Menon, C., Allaire, C., Yong, P., Bedaiwy, M.A.: Clinical use of artificial intelligence in endometriosis: A scoping review. Npj Digital Medicine **5**(1), 109 (2022). https://doi.org/10.1038/s41746-022-00638-1