Bayesian Diagnostic Model Development for Endometriosis Using Markov Chain Monte Carlo Methods (MCMC) and Synthetic Data

June 2026

Ludiah Bagakas

Department of Mathematics and Statistics
Cleveland State University
2121 Euclid Ave
Cleveland, Ohio 44115
United States

Corresponding Author: Ludiah Bagakas
Email: l.bagakas@alumni.csuohio.edu
Telephone: +1 (216) 533-1088

Abstract

Endometriosis is notoriously difficult to diagnosis. Patients wait an average of seven years to receive a formal diagnosis [1]. In this study, Bayesian logistic regression is applied to synthetic data to develop a predictive model which can aid in early detection of endometriosis. Comparisons are made across six candidate models with informative normal and weakly informative Cauchy priors. The Markov Chain Monte Carlo and Metropolis-Hastings algorithm is applied to these general linear models to predict the odds of possessing this gynecological condition.

Keywords: endometriosis, Bayesian inference, general linear model, Markov Chain Monte Carlo, Metropolis-Hastings, normal prior, Cauchy prior

Introduction

Endometriosis is a chronic inflammatory condition, where tissue similar to the uterine lining grows outside of the uterus resulting in inflammation, scarring, and severe pain. It is characterized by a broad spectrum of symptoms, some of which include chronic pain, gastrointestinal issues, fatigue, irregular menstruation, and infertility.

This condition has been associated with a reduced quality of life, and the delay in its detection prolongs patients' access to necessary treatments for symptom management. As of now, the only definitive way of diagnosing this condition is through a laparoscopy; minimally invasive surgery [2]. When endometrial lesions are detected, they are generally removed and biopsied to confirm diagnosis. The removal of these lesions is expected to alleviate symptoms. One major drawback of this technique is that when endometriosis is not found, patients are burdened with the vulnerability of recovering from a painful, unwarranted, and costly procedure. Women with presumed diagnoses avoid this procedure out of fear of a similar outcome. These drawbacks highlight several motivations to develop a diagnostic model to infer the probability of endometriosis being present.

Methodology

Data Collection

Since endometriosis research is widely considered to be under-researched and under-funded, relevant data sets are hard to access. Consequently, this study utilizes synthetic data simulated to resemble a real-life experiment. This Kaggle data set consists of 10,000 observations reflecting common characteristics and symptoms of endometriosis [3]. This study's data is strictly utilized for training purposes and should not be generalized to its respective population.

EDA

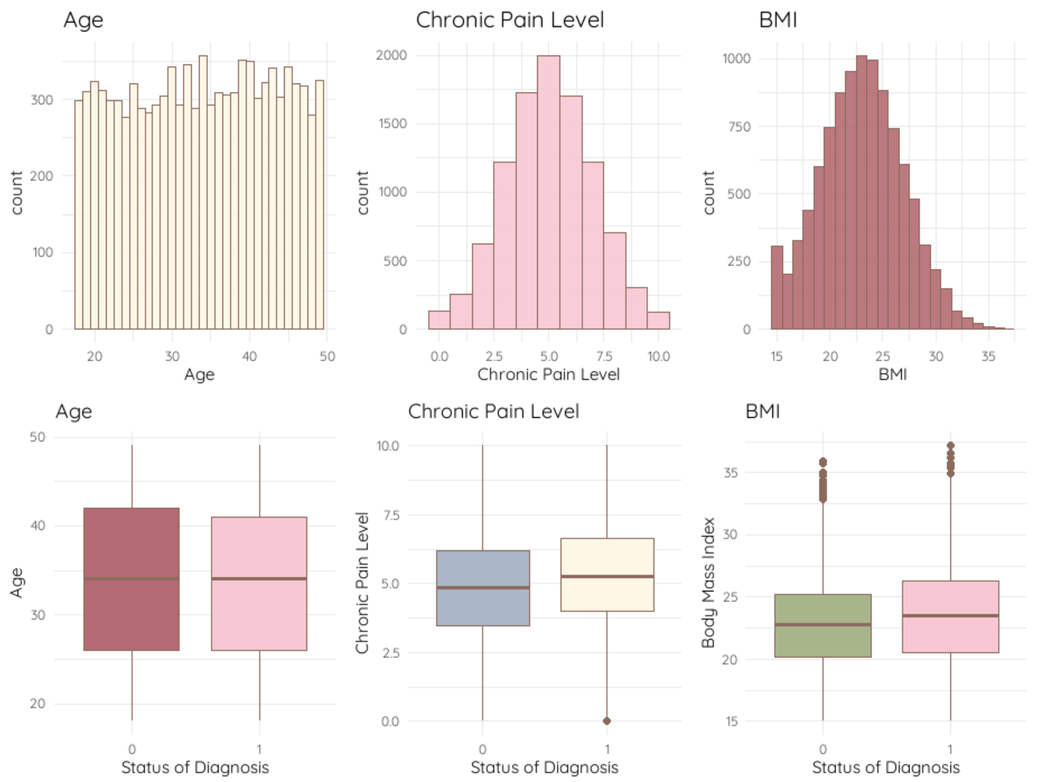
In this study, six variables are used to predict the binary outcome of whether or not endometriosis is present (0 = No, 1 = Yes). These predictors are highlighted in *Table 1*.

Table 1: Description of Predictors

Variable	**Type**	**Description**
Age	Discrete	Age of individual (18-50).
Menstrual Irregularity	Binary	Indicates whether or not
Chronic Pain Level	Continuous	Self-reported pain severity level (scale from 0-10).
Hormone Level Abnormality	Binary	Indicates presence of abnormal hormone levels (0 = No , 1 = Yes).
Infertility	Binary	Indicates whether or not an individual experiences infertility (0 = No, 1 = Yes).
BMI	Continuous	Infertility |Body Mass Index (18-40 kilograms per square meter).

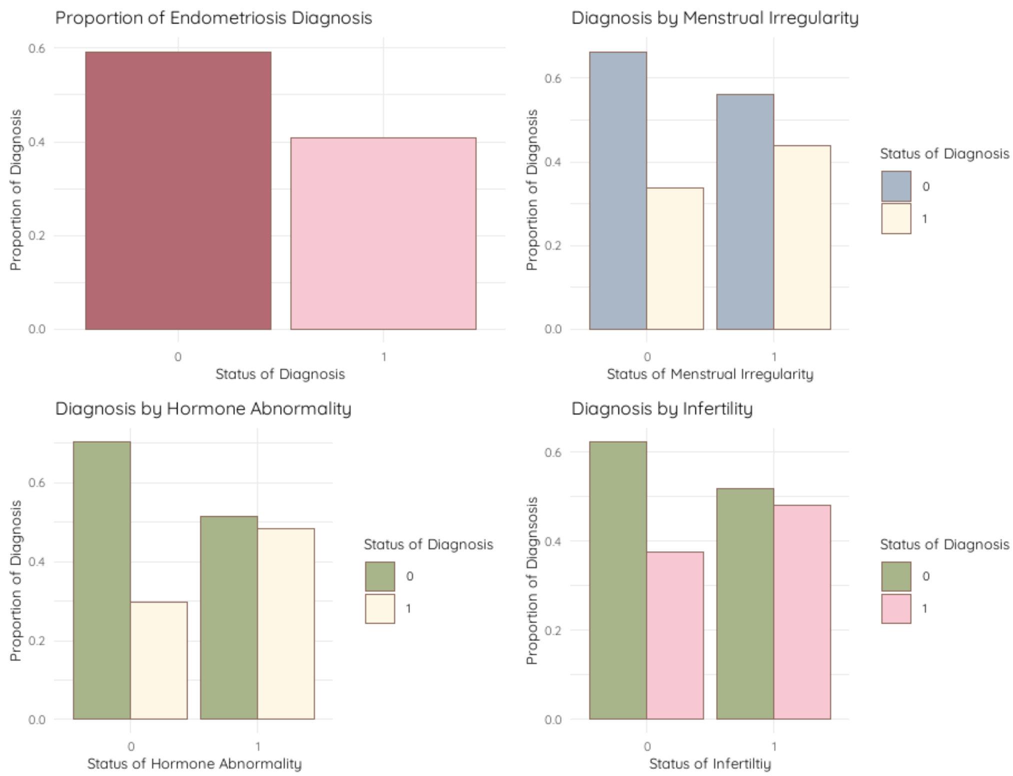
According to the graphs of quantitative variables in *Figure 1,* The distribution of ages appears to follow a normal, yet slightly static distribution centered around a mean of 34 years old. Chronic pain levels are distributed normally around a mean of 5, while body mass indices are slightly skewed right with a mean and median of about 23. The skewness of the distribution for BMI values can be explained by several high outliers. This feature is expected, as there is a high prevalence of obesity in the United States. Based on the box plots in *Figure 1*, a diagnosis of endometriosis seems to be associated with higher chronic pain levels and BMIs. Since endometriosis is more common in women of reproductive age, and less common in women experiencing and transitioning into menopause, we would expect this condition to affect a narrower range (smaller variance) of individuals.

Figure 1: Histograms & Boxplots of Quantitative Predictors



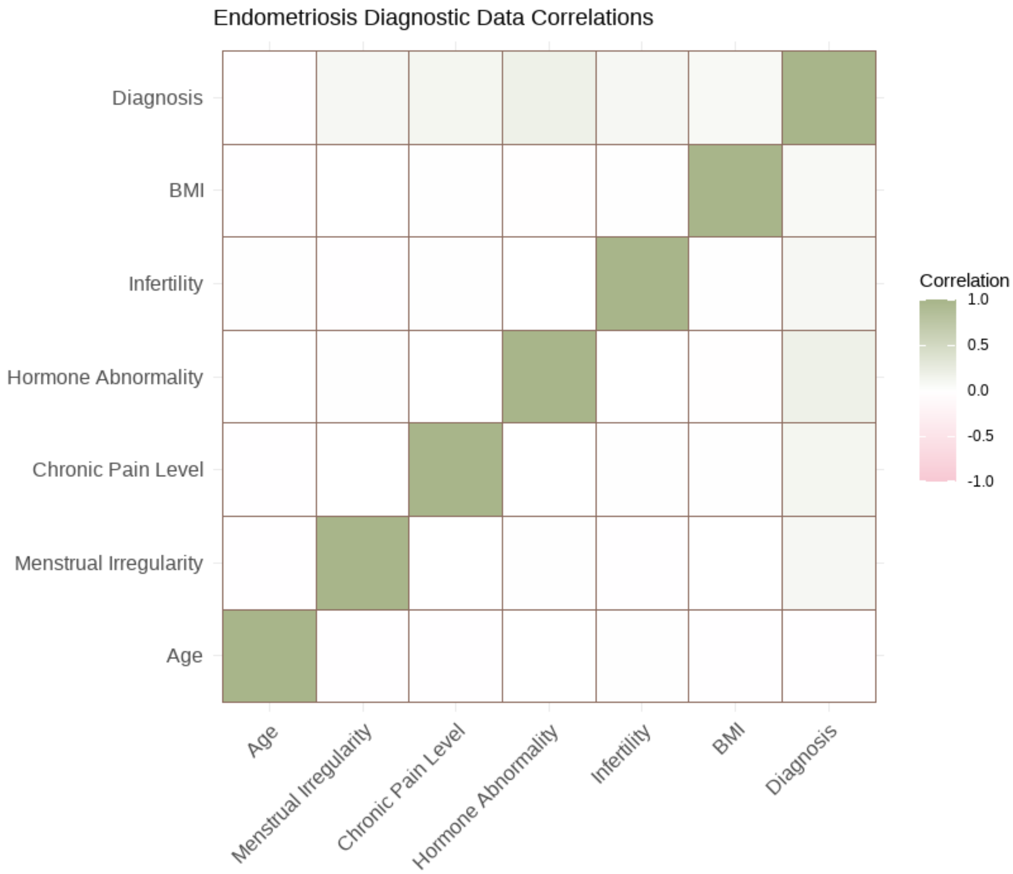
Based on the bar charts in *Figure 2*, endometriosis appears to be more prevalent in individuals who experience menstrual irregularity, hormone abnormality, and infertility. These binary variables are expected to share a positive relationship with the odds of an endometriosis diagnosis.

Figure 2: Bar Charts of Binary Variables



The correlation heat map in *Figure 3*, indicates that the strongest positive association exists between the presence of endometriosis and hormone level abnormalities. Other significant predictors that share a significant positive relationship with the response are chronic pain level, infertility and menstrual irregularity respectively. These explanatory variables are expected to dominate the regression model.

Figure 3: Correlation Plots



Bayesian Logistic Regression

The response variable follows a binomial distribution $y_i \sim Bin(n_i, \mu_i)$, where n_i is known. We select a logit transformation $log(\frac{u_i}{1-u_i})$ to link the linear predictors to the mean of the response [4].

We choose a model:

$\log(\frac{\pi}{1-\pi}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6$, where
$\beta_0\ \beta_1\ \beta_2\ \beta_3\ \beta_4, \beta_5, \beta_6 \sim Normal(0, \sigma)$

The data distribution for data *y* is:

$$p(y|\beta) = \prod_{i=1}^{n} \binom{n_i}{p_i} (\frac{e^{n_i}}{1+e^{n_i}})^{y_i} (\frac{1}{1+e^{n_i}})^{n_i - y_i}$$

Normal Informative Priors

Parameterization of the explanatory variables were considered using normal prior distributions $\beta_i \sim N(\beta, \Sigma_\beta)$ for $\beta_0, \beta_1 ... \beta_6$. In this intuitive approach, prior information can be expressed directly in terms of these parameters [5].

The intercept parameter β_0 is selected based on the belief that the "typical" women age 18-50 has a 5%-20% chance of having endometriosis. The logit transformation is applied to this range and the prior mean is established as the center of this range.

$$\beta_0(centered) = \frac{log(\frac{0.05}{1-0.05}) + log(\frac{0.20}{1-0.20})}{2}$$

The prior standard deviation of β_0 is set as the half way distance between this centered value and the endpoints of this range.

$$\Sigma_{\beta_0} = \frac{\beta_{0(centered)} - logit(0.05)}{2} = \frac{logit(0.20) - \beta_{0(centered)}}{2}$$

For each proceeding coefficient ($\beta_1 ... \beta_6$), we calculate the expected log of odds for each unit increase in the explanatory variable. For instance, we would expect the odds of an endometriosis diagnosis to increase by 25% for each increased level of chronic pain. As a result, the prior mean on β_3 is log(1.25) and the standard deviation is the halfway distance between this value and the upper and lower bounds bounds of β_3 [5]. *Table 2* provides the prior mean and standard deviation selected for each coefficient.

Table 2: Prior Mean and Standard Deviation for Coefficients

Predictors	Prior Mean	Prior Standard Deviation
Intercept	-2.1653667	0.3895362
Age	0.0476551	0.0238275
Menstrual Irregularity	0.1311821	0.0655911
Chronic Pain Level	0.2027326	0.1013663
Hormone Level Abnormaility	0.2027326	0.1013663
Infertility	0.1311821	0.0655911
BMI	0.0476551	0.0238275

Before new data is observed from the likelihood, the expected range of values can be estimated. *Table 3* highlights these 95% prior quantile intervals, which reinforce the prior belief that a typical patient is unlikely to have endometriosis (β_0), but with each increased value in the coefficients, the odds of a positive diagnosis increase.

Table 3: 95% Prior Quantile Intervals

Predictors	**Lower Bound**	**Upper Bound**
Intercept	-2.9288435	-1.4018898
Age	0.0009540	0.0943562
Menstrual Irregularity	0.0026260	0.2597383
Chronic Pain Level	0.0040583	0.4014068
Hormone Level Abnormaility	0.0040583	0.4014068
Infertility	0.0026260	0.2597383
BMI	0.0009540	0.0943562

Cauchy Priors

We consider the Cauchy distribution to embrace a conservative approach to selecting priors. These weekly informative priors offer a balance between the unstable, extreme inferences obtained from highly informative and completely non-informative priors. To apply this model, binary independent variable are standardized to have a mean of zero and standard deviation of 1, while the remaining quantitative variables are standardized to have a mean of zero and a standard deviation of 0.5. This step is essential to compensate for the varying units of measure and possible ranges in the data [6].

The Cauchy assumes a Student t-distribution with mean 0, degrees-of-freedom set to 1, and a scale of 2.5 for all six coefficients [6].

$$\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6 \sim Cauchy(x_0 = 0, \gamma = 2.5)$$

Alternatively, we apply a weaker prior to the constant term. Since the baseline prevalence of endometriosis is not as high, the scale is increased from 2.5 to 10 [6].

$$\beta_0 \sim Cauchy(x_0 = 0, \gamma = 10)$$

Metropolis-Hastings

The joint posterior distribution for the logistic regression model can be expressed as

$$p(\beta, y) \propto \prod_i Bin(x_i|n_i, logit(X_i\beta)) \prod_i N(\beta_i|u_i, \sigma_i)$$

when normal priors are selected, and

$$p(\beta, y) \propto \prod_i Bin(x_i|n_i, logit(X_i\beta)) \prod_i Cauchy(\beta_i|x_i, \gamma_i)$$

when Cauchy priors are applied.

Since these distributions can not be obtained directly, MCMC methods are applied to estimate the posterior.

The Metropolis-Hastings algorithm relies on an arbitrary value $x^{(0)}$ from a proposed probability distribution $q(x)$ to initialize the process.

Initialize $x^{(0)} \sim q(x)$

For the next iteration, a new candidate value $x^{(cand)}$is simulated from another distribution $q(x^i|x^{(i-1)})$.

Propose $x^{(i)} \sim q(x^i|x^{(i-1)})$

The acceptance probability α is evaluated to determine whether or not the proposed value will become the current value in the next iteration. This probability is calculated using the ratio of the proposed posterior to the current posterior distribution.

$$\alpha = min(1, \frac{q(x^{(cand)}|x^{(i-1)})}{q(x^{(i-1)}|x^{(cand)})}$$

This process is repeated for iteration $i = 1,2, \dots$ until the distribution of Markov Chains converge to the posterior [7].

Results

Fitted Models

Six proposal models are developed to predict the odds of an endometriosis diagnosis given a series of potential symptoms and characteristics. The first four models utilize normal informative prior distributions, where prior parameters are intuitively chosen. The first model includes all input variables and can be expressed as the following:

Model 1:

$$log(\frac{\pi}{1-\pi}) = \beta_0 + \beta_1(Age) + +\beta_2(Menstrual.Irregularity) + \beta_3(Chronic.Pain.Level) + \beta_4(Hormone.Abnormality) + \beta_5(Infertility) + \beta_6(BMI)$$

In the fitted models, coefficients for age and BMI are close to zero indicating a low level of significance on the overall model. To further assess their significance, these predictors are transformed in the second model, so that age is scaled to every five years and BMI is analyzed categorically. *Table 4* shows the frequency table for the BMI categories.

Table 4: Frequency Table for BMI Categories

BMI Category	**Frequency**
Healthy Weight	0.5944791
Obese	0.5945416
Overweight	0.6132502
Underweight	0.6061215

The fitted equation for Model 2 is in the form:

Model 2:

$$log(\frac{\pi}{1-\pi}) = \beta_0 + \beta_1(Age/5) + \beta_2(Menstrual.Irregularity) + \beta_3(Chronic.Pain.Level) + \beta_4(Hormone.Abnormality) + \beta_5(Infertility) + \beta_6(Obese) + \beta_7(Overweight) + \beta_8(Underweight)$$

In Models 3 and 4, variables for age and BMI are dropped respectively taking on the form:

Model 3:

$$log(\frac{\pi}{1-\pi}) = \beta_0 + \beta_1(Menstrual.Irregularity) + \beta_2(Chronic.Pain.Level) + \beta_3(Hormone.Abnormality) + \beta_4(Infertility) + \beta_5(Obese) + \beta_6(Overweight) + \beta_7(Underweight)$$

Model 4:

$$log(\frac{\pi}{1-\pi}) = \beta_0 + \beta_1(Menstrual.Irregularity) + \beta_2(Chronic.Pain.Level) + \beta_3(Hormone.Abnormality) + \beta_4(Infertility)$$

Although these variables are removed to simplify the model and prevent noise, dropping these variables could result in information loss, especially since there is no evidence of collinearity present in the data. This is considered when fitting the two remaining models with Cauchy priors. Since there is significant information in the overweight and obese categories, the BMI variable is kept in Models 5 and 6.

Model 5:

$$log(\frac{\pi}{1-\pi}) = \beta_0 + \beta_1(Age) + +\beta_2(Menstrual.Irregularity) + \beta_3(Chronic.Pain.Level) + \beta_4(Hormone.Abnormality) + \beta_5(Infertility) + \beta_6(BMI), \beta_i \sim Cauchy(x_0, \gamma)$$

Model 6:

$$log(\frac{\pi}{1-\pi}) = \beta_0 + \beta_1(Menstrual.Irregularity) + \beta_2(Chronic.Pain.Level) + \beta_3(Hormone.Abnormality) + \beta_4(Infertility) + \beta_5(BMI), \beta_i \sim Cauchy(x_0, \gamma)$$

Table 5 provides the results of fitted Models 1-6. Based on these results, hormone abnormalities and chronic pain level appear to consistently hold the most predictive power across all six models.

Table 5: Posterior Coefficient Estimates for Models 1-6

Estimate	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**
Intercept	-2.144471	-1.231417	-1.832922	-1.720188	-0.404786	-0.404599
Age	-0.001714	-0.001642	ommitted	ommitted	-0.039812	ommitted
Menstrual Irregularity	0.132292	0.129262	0.335697	0.339813	0.206866	0.207461
Chronic Pain Level	0.109986	0.110095	0.125449	0.12447	0.505207	0.506221
Hormone Level Abnormaility	0.230077	0.22889	0.610487	0.603655	0.4406668	0.40658
Infertility	0.131111	0.131966	0.340058	0.339403	0.205028	0.204763
BMI	0.043092	-	-	ommitted	0.353553	0.3542265
Obese	-	0.067705	0.240297	-	-	-
Overweight	-	0.129528	0.348393	-	-	-
Underweight	-	0.026705	0.038178	-	-	-

Diagnostics

ROC Curves & Accuracy Metrics

Classifications are created so that posterior values greater than or equal to 0.4 are presumed to have a positive endometriosis diagnosis. Otherwise, a patient is classified as negative. ROC curves *(Figure 4)* and accuracy measures *(Table 6)* serve as preliminary performance indicators for model selection. With an area-under-the curve (AUC) ranging from 0.63-0.67, each model performs slightly better than random chance at different thresholds. Since this study is most concerned with identifying positive cases of endometriosis, it makes sense to maximize sensitivity. The models with the highest sensitivity are 3, 5 and 6. These models achieved the highest metrics for specificity, Cohen's Kappa, and AUC. Kappa values in these three models are approximately 0.22, indicating fair agreement between predicted and observed outcomes after accounting for random chance. Due to the results of this section, the remainder of the analysis is performed on Models 3, 5, and 6.

Figure 4: ROC Curves for Models 1-6

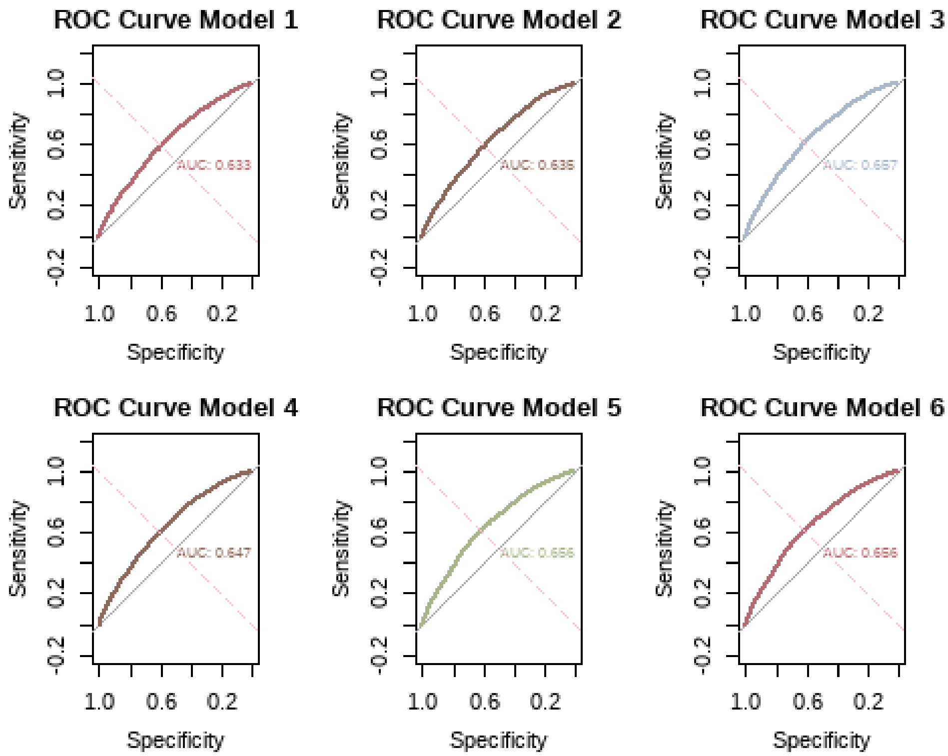


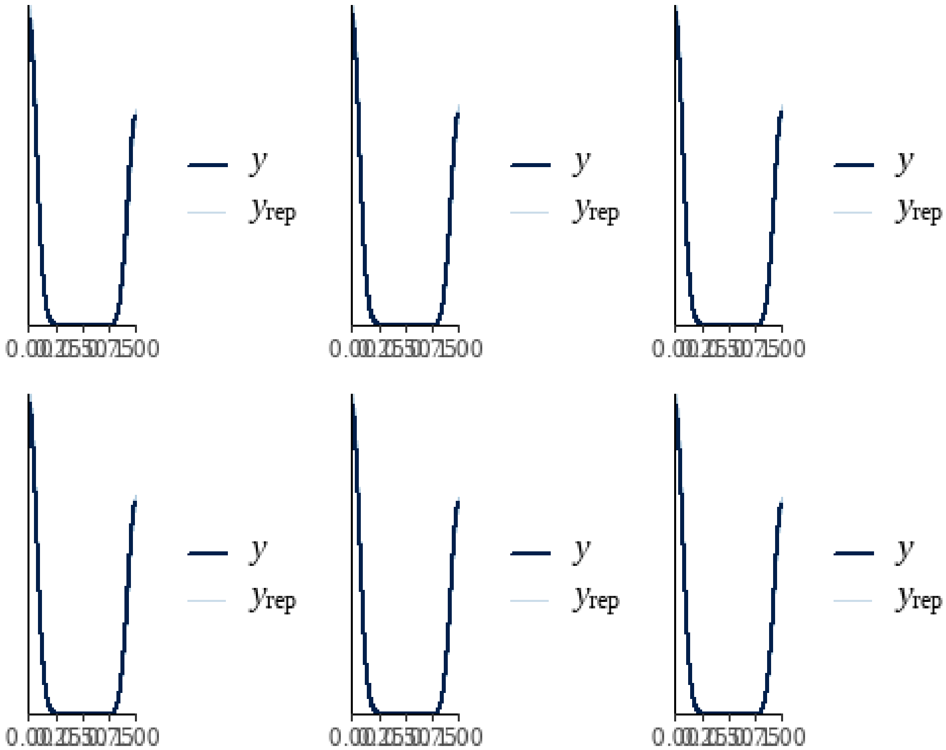
Table 6: Preliminary Performance Metrics

Model	**Balanced Accuracy**	**Sensitivity**	**Specificity**	**Kappa**
Model 1	0.5944791	0.6312822	0.5576761	0.1812516
Model 2	0.5945416	0.6452562	0.5438271	0.1804540
Model 3	0.6132502	0.6526109	0.5738895	0.2173342
Model 4	0.6061215	0.6496690	0.5625739	0.2032446
Model 5	0.6150781	0.6518755	0.5782807	0.2210765
Model 6	0.6156774	0.6513851	0.5799696	0.2223255

Posterior Predictive Checks

Posterior predictive checks are performed to evaluate the fit of the model to the observed data [5]. As shown in *Figure 5*, the dark blue line, which represents the observed data distribution closely overlaps the light blue line representing data simulated by the posterior distribution. Based on these features, all six models appear to be an adequate fit for newly observed data [8].

Figure 5: Posterior Predictive Checks



Stratified K-Fold Cross Validation, ELPD, & LOOIC

The accuracy of these three models are further assessed using stratified k-fold cross validation accuracy, expected log-predictive density (ELPD), and the leave-one-out cross validation information criterion. The aim is to maximize the expected log-predictive density (ELPD) and cross validation accuracy. Alternatively, the LOOIC must be minimized to select the strongest model [10]. In *Table 7*, Model 6 has the lowest LOOIC and highest ELPD suggesting higher out-of-sample predictive accuracy.

LOOIC

Table 7: ELPD & LOOIC for Models 3,5, and 6

Model	LOOIC	ELPD
Model 3	12785.17	-6392.585
Model 5	12780.23	-6390.113
Model 6	12779.38	-6389.692

For stratified k-fold cross validation, the data-set is divided into 80% for training and 20% validation with respect to its sample proportions. Ten equal size folds are created on training data and each model is iteratively applied to nine folds, while accuracy metrics are evaluated on the tenth fold of the training data. The process is repeated until each fold has been evaluated [7].

Table 8 contains the mean accuracy, sensitivity, and specificity for the three candidate models after applying cross validation on the training data. Only slight discrepancies are apparent in the performance metrics between the three models. The results for predictive density, LOOIC, and cross validation accuracy suggest that Model 6 slightly outperforms the rest. Thus, Model 6 is established as the final model.

Final Model:

$$log(\frac{\pi}{1-\pi}) = \beta_0 + \beta_1(Menstrual.Irregularity) + \beta_2(Chronic.Pain.Level) + \beta_3(Hormone.Abnormality) + \beta_4(Infertility) + \beta_5(BMI), \beta_i \sim Cauchy(x_0, \gamma)$$

Table 8: K-Fold CV Performance Metrics

Model	Balanced Accuracy	Sensitivity	Specificity
Model 3	0.6108165	0.6515880	0.5700450
Model 5	0.6099075	0.6459776	0.5738375
Model 6	0.6104180	0.6466010	0.5742350

Conclusions

Final assessments for Model 6 are evaluated on the test set and summarized in *Table 9* and *Figure 6*. An area of 0.67 under the ROC curve suggests that there is a 67% probability of endometriosis correctly being diagnosed across all thresholds. In developing this model, priority is placed on optimizing sensitivity, so that cases of endometriosis are being accurately detected. With an accuracy of about 63% and a sensitivity of about 67%, this Bayesian predictive model outperforms routine vaginal examinations by nearly 30% [11].

Figure 6: ROC Curve of Final Model

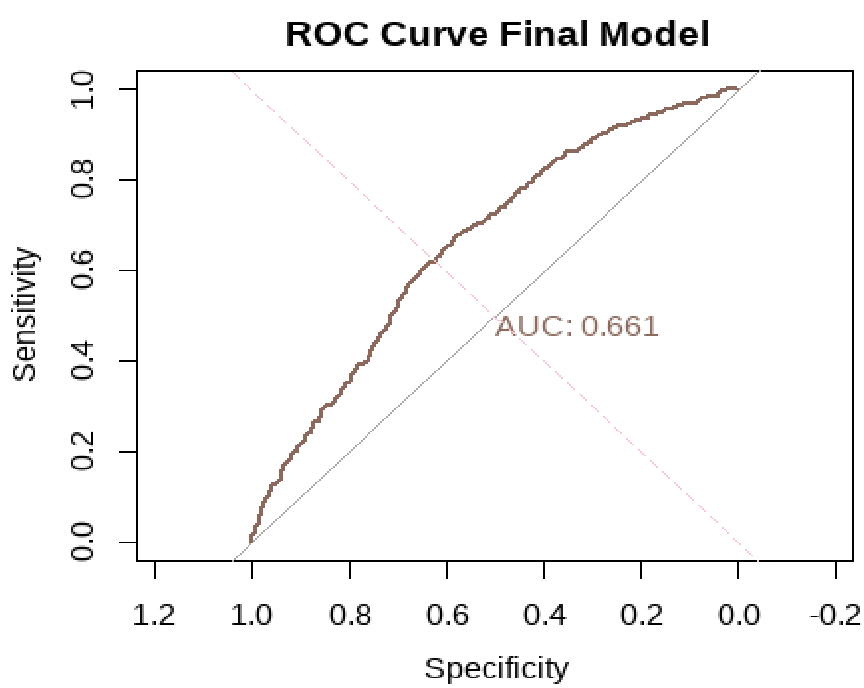


Table 9: Out-of-Sample Performance Metrics for Final Model

Metric	Value
Accuracy	0.6255732
Sensitivity	0.6654321
Specificity	0.5857143
AUC	0.6607107
LOOIC	2554.6130248
ELPD	-1277.3065124

Although the final model's accuracy only slightly exceeds random chance, this analysis serves as a starting point for further development of diagnostic algorithms for endometriosis.

Declarations

Funding

The author received no specific grant, funding, or financial support from any public commercial, not-for-profit funding agency for the research, authorship, or publication of this manuscript.

Conflicts of Interest

The author declares no conflicts of interest.

Author Contributions

Ludiah Bagakas conceived the study, conducted the statistical analyses, interpreted the results, interpreted the findings, and prepared the manuscript.

Data Availability

The data supporting the findings of this study are publicly available from the source cited in the References section.

References

[1]Endometriosis Australia. Diagnosis of Endometriosis. Available from: https://endometriosisaustralia.org/diagnosis/

[2] Cleveland Clinic. Endometriosis: Causes, Symptoms, Diagnosis & Treatment. Available from: https://my.clevelandclinic.org/health/diseases/10857-endometriosis

[3] Anietie M. Endometriosis Dataset. Kaggle. Available from: https://www.kaggle.com/datasets/michaelanietie/endometriosis-dataset

[4] Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. Bayesian Data Analysis. 3rd ed. Boca Raton (FL): CRC Press; 2013.

[5] Hitchcock DB. Bayesian Regression and Classification Notes. University of South Carolina.

[6] Gelman A, Jakulin A, Pittau MG, Su YS. A Weakly Informative Default Prior Distribution for Logistic and Other Regression Models. Annals of Applied Statistics. 2008.

[7] Jacobs Laboratory. Metropolis-Hastings Sampling Cheat Sheet. University of Rochester.

[8] Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. Bayesian Data Analysis. 3rd ed. Boca Raton (FL): CRC Press; 2013.

[9] Stan Development Team. Posterior and Prior Predictive Checks. In: Stan User's Guide. Available from: https://mc-stan.org/docs/stan-users-guide/posterior-predictive-checks.html

[10] Vehtari A, Gelman A, Gabry J. Practical Bayesian Model Evaluation Using Leave-One-Out Cross-Validation and WAIC. arXiv. 2015.

[11] Kido A. MRI in the diagnosis of endometriosis and related diseases. Japanese Journal of Radiology. 2022;40(3):213-227. doi:10.1007/s11604-021-01215-z.

Appendix

Model Summaries

Model Summary for Model 1 (STAN_GLM)

```
## stan_glm
##  family:       binomial [logit]
##  formula:      Diagnosis ~ .
##  observations: 10000
##  predictors:   7
## ------
##                            Median    MAD_SD  
## (Intercept)                -2.144471  0.149156
## Age                        -0.001714  0.002122
## Menstrual_Irregularity      0.132292  0.020286
## Chronic_Pain_Level          0.109859  0.009756
## Hormone_Level_Abnormality   0.230077  0.021659
## Infertility                 0.131111  0.020855
## BMI                         0.043092  0.004977
## 
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg
```

Model Summary for Model 2 (STAN_GLM)

```
## stan_glm
##  family:       binomial [logit]
##  formula:      Diagnosis ~ .
##  observations: 10000
##  predictors:   9
## ------
##                            Median    MAD_SD  
## (Intercept)                -1.237454  0.094006
## Age                        -0.000938  0.010150
## Menstrual_Irregularity      0.132483  0.020066
## Chronic_Pain_Level          0.109602  0.009697
## Hormone_Level_Abnormality   0.229167  0.020831
## Infertility                 0.131686  0.020924
## BMIObese                    0.069148  0.023216
## BMIOverweight               0.125917  0.020902
## BMIUnderweight              0.027054  0.021661
## 
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg
```

Model Summary for Model 3 (STAN_GLM)

```
## stan_glm
##  family:       binomial [logit]
##  formula:      Diagnosis ~ .
##  observations: 10000
##  predictors:   8
## ------
##                            Median    MAD_SD
## (Intercept)               -1.833895  0.071675
## Menstrual_Irregularity     0.341458  0.038214
## Chronic_Pain_Level         0.124841  0.010919
## Hormone_Level_Abnormality  0.607745  0.036987
## Infertility                0.341929  0.037274
## BMIObese                   0.244599  0.053959
## BMIOverweight              0.336316  0.036384
## BMIUnderweight             0.040202  0.045144
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg
```

Model Summary for Model 4 (STAN_GLM)

```
## stan_glm
##  family:       binomial [logit]
##  formula:      Diagnosis ~ .
##  observations: 10000
##  predictors:   5
## ------
##                            Median    MAD_SD
## (Intercept)               -1.720188  0.070910
## Menstrual_Irregularity     0.339813  0.037011
## Chronic_Pain_Level         0.124470  0.010865
## Hormone_Level_Abnormality  0.603655  0.035378
## Infertility                0.339403  0.036733
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg
```

Model Summary for Model 5 (STAN_GLM)

```
## stan_glm
##  family:       binomial [logit]
##  formula:      Diagnosis ~ .
##  observations: 10000
##  predictors:   7
## ------
##                             Median    MAD_SD
## (Intercept)                -0.404786  0.021006
## Menstrual_Irregularity      0.206866  0.020874
## Hormone_Level_Abnormality   0.406668  0.021150
## Infertility                 0.205028  0.021056
## Age                        -0.039812  0.041374
## Chronic_Pain_Level          0.505207  0.042337
## BMI                         0.353553  0.042837
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg
```

Model Summary for Model 6 (STAN_GLM)

```
## stan_glm
##  family:       binomial [logit]
##  formula:      Diagnosis ~ .
##  observations: 10000
##  predictors:   6
## ------
##                             Median    MAD_SD
## (Intercept)                -1.452601  0.130407
## Menstrual_Irregularity      0.207297  0.022588
## Hormone_Level_Abnormality   0.406565  0.021099
## Infertility                 0.204750  0.021349
## Chronic_Pain_Level          0.505425  0.041508
## BMI                         0.045472  0.005589
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg
```

Confusion Matrices

Confusion Matrix for Model 1

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##          0 3302 1504
##          1 2619 2575
##
##                Accuracy : 0.5877
##                  95% CI : (0.578, 0.5974)
##     No Information Rate : 0.5921
##     P-Value [Acc > NIR] : 0.8174
##
##                   Kappa : 0.1813
##
##  Mcnemar's Test P-Value : <2e-16
##
##             Sensitivity : 0.6313
##             Specificity : 0.5577
##          Pos Pred Value : 0.4958
##          Neg Pred Value : 0.6871
##              Prevalence : 0.4079
##          Detection Rate : 0.2575
##    Detection Prevalence : 0.5194
##       Balanced Accuracy : 0.5945
##
##        'Positive' Class : 1
##
```

Confusion Matrix for Model 2

```
conf.mat2

## Confusion Matrix and Statistics
## 
##           Reference
## Prediction    0    1
##          0 3220 1447
##          1 2701 2632
## 
##                Accuracy : 0.5852
##                  95% CI : (0.5755, 0.5949)
##     No Information Rate : 0.5921
##     P-Value [Acc > NIR] : 0.9213
## 
##                   Kappa : 0.1805
## 
##  Mcnemar's Test P-Value : <2e-16
## 
##             Sensitivity : 0.6453
##             Specificity : 0.5438
##          Pos Pred Value : 0.4935
##          Neg Pred Value : 0.6900
##              Prevalence : 0.4079
##          Detection Rate : 0.2632
##    Detection Prevalence : 0.5333
##       Balanced Accuracy : 0.5945
## 
##        'Positive' Class : 1
## 
```

Confusion Matrix for Model 3

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##          0 3398 1417
##          1 2523 2662
##
##                Accuracy : 0.606
##                  95% CI : (0.5963, 0.6156)
##     No Information Rate : 0.5921
##     P-Value [Acc > NIR] : 0.002381
##
##                   Kappa : 0.2173
##
##  Mcnemar's Test P-Value : < 2.2e-16
##
##             Sensitivity : 0.6526
##             Specificity : 0.5739
##          Pos Pred Value : 0.5134
##          Neg Pred Value : 0.7057
##              Prevalence : 0.4079
##          Detection Rate : 0.2662
##    Detection Prevalence : 0.5185
##       Balanced Accuracy : 0.6133
##
##        'Positive' Class : 1
##
```

Confusion Matrix for Model 4

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##          0 3331 1429
##          1 2590 2650
##
##                Accuracy : 0.5981
##                  95% CI : (0.5884, 0.6077)
##     No Information Rate : 0.5921
##     P-Value [Acc > NIR] : 0.1129
##
##                   Kappa : 0.2032
##
##  Mcnemar's Test P-Value : <2e-16
##
##             Sensitivity : 0.6497
##             Specificity : 0.5626
##          Pos Pred Value : 0.5057
##          Neg Pred Value : 0.6998
##              Prevalence : 0.4079
##          Detection Rate : 0.2650
##    Detection Prevalence : 0.5240
##       Balanced Accuracy : 0.6061
##
##        'Positive' Class : 1
##
```

Confusion Matrix for Model 5

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##          0 3424 1420
##          1 2497 2659
##
##                Accuracy : 0.6083
##                  95% CI : (0.5987, 0.6179)
##     No Information Rate : 0.5921
##     P-Value [Acc > NIR] : 0.0004962
##
##                   Kappa : 0.2211
##
##  Mcnemar's Test P-Value : <2e-16
##
##             Sensitivity : 0.6519
##             Specificity : 0.5783
##          Pos Pred Value : 0.5157
##          Neg Pred Value : 0.7069
##              Prevalence : 0.4079
##          Detection Rate : 0.2659
##    Detection Prevalence : 0.5156
##       Balanced Accuracy : 0.6151
##
##        'Positive' Class : 1
##
```

Confusion Matrix for Model 6

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##          0 3434 1422
##          1 2487 2657
##
##                Accuracy : 0.6091
##                  95% CI : (0.5995, 0.6187)
##     No Information Rate : 0.5921
##     P-Value [Acc > NIR] : 0.0002793
##
##                   Kappa : 0.2223
##
##  Mcnemar's Test P-Value : <2e-16
##
##             Sensitivity : 0.6514
##             Specificity : 0.5800
##          Pos Pred Value : 0.5165
##          Neg Pred Value : 0.7072
##              Prevalence : 0.4079
##          Detection Rate : 0.2657
##    Detection Prevalence : 0.5144
##       Balanced Accuracy : 0.6157
##
##        'Positive' Class : 1
##
```