Prediction of uterine dehiscence via machine learning by using lower uterine segment thickness and clinical features

Mervenur Kement, MD; Cihan Emre Kement, PhD; Mahmut Kuntay Kokanali, MD; Melike Doganay, MD

BACKGROUND: With the global increase of cesarean delivery rates, the long-term effects of cesarean delivery have started to become clearer. One of the most prominent complications of cesarean delivery in recurrent pregnancies is uterine rupture. Assessing the risk of uterine rupture by accurately predicting dehiscence is very important to prevent untimely operations and/or maternal and fetal complications.

OBJECTIVE: This study aimed to assess whether machine learning can be used to predict uterine dehiscence by using patients’ ultrasonographic findings, clinical findings, and demographic data as features. Hence, we investigated a potential method for preventing uterine rupture and its maternal and fetal complications.

STUDY DESIGN: The study was conducted on 317 patients with term (>37 weeks’ gestation) singleton pregnancies and no medical complications or medications that may affect uterine wound healing. Demographics, body mass indices, smoking and drinking habits, clinical features, past pregnancies, number and history of abortions, interdelivery period, gestational week, number of previous cesarean deliveries, fetal presentation, fetal weight, tocography data, transabdominal ultrasonographic measurement of lower uterine segment full thickness and myometrium thickness, and lower uterine segment findings during cesarean delivery were collected and analyzed using machine learning techniques. Logistic regression, multilayer perceptron, support vector machine, random forest, and naive Bayes algorithms were used for classification. The dataset was evaluated using 10-fold cross-validation. Correct classification rate, F-score, Matthews correlation coefficient, precision-recall curve area, and receiver operating characteristic area were used as performance metrics.

RESULTS: Among the machine learning techniques tested in this study, the naive Bayes algorithm showed the best predictive performance. Among the various combinations of features used for prediction, the essential features of parity, gravidity, tocographic contraction, cervical dilation, dilation and curettage, and sonographic thickness of lower uterine segment myometrium yielded the best results. The second-best performance was achieved with sonographic full thickness of lower uterine segment added to the base features. The base features alone could classify patients with 90.5% accuracy, whereas adding the myometrium measurement increased the classification performance by 5.1% to 95.6%. Adding the full thickness measurement to the base features raised the classification performance by 4.8% to 95.3% in terms of correct classification rate.

CONCLUSION: The naive Bayes algorithm can correctly classify uterine dehiscence with a correct classification rate of 0.953, an F-score of 0.952, and a Matthews correlation coefficient value of 0.641. This result can be interpreted as indicating that by using clinical features and lower uterine segment ultrasonography findings, machine learning can be used to accurately predict uterine dehiscence.

Key words: cesarean delivery, contraction pattern, lower uterine segment ultrasound, machine learning, uterine dehiscence, uterine rupture

The concern for uterine rupture can lead to patients with contractions and previous cesarean delivery undergoing preterm cesarean delivery. Repeated cesarean delivery not only increases the risk of uterine rupture, but also neonatal complications such as bleeding, hysterectomy, thromboembolism, and respiratory distress.
An earlier study by Rozenberg et al.5 predicted uterine dehiscence and rupture. AJOG Global Reports
increased risk of rupture.6

"The lower uterine segment (LUS) ultrasound measurements have shown that uterine defects are directly related to thin LUS and patients with LUS thickness of <2.5 mm, whereas it was 0.7% in patients with LUS thickness of ≥3.5 mm. Other studies also found correlation between thin LUS and increased risk of rupture.6−14 Conversely, meta-analyses indicated that LUS alone is insufficient for predicting uterine rupture while concurring that LUS thickness of <2 mm poses a risk for rupture and dehiscence.5,16"

In this study, we aimed to assess whether uterine dehiscence can be predicted via machine learning (ML) methods by using LUS measurements, clinical findings, and patient demographics as features. Moreover, we aimed to investigate which of these features are more relevant to the predictions attempted by ML algorithms. Furthermore, we compared the effects of using LUS myometrium thickness and LUS full thickness as features, and the impact of using these 2 features together on the prediction performance of ML techniques.

**Materials and Methods**

**Study design and participants**

The study was approved by the Clinical Research Ethics Committee of the Ankara City Hospital, University of Health Sciences (approval number: E2-20-108) in January 2021. Singleton pregnancies with previous cesarean delivery that applied to the Department of Obstetrics and Gynecology Labor Unit of the Ankara City Hospital, Ankara, Turkey between February and July of 2021 were included in this study. The study was conducted on 317 patients with term (>37 weeks’ gestation) singleton pregnancies. Both patients with scheduled cesarean delivery and patients who applied through the Emergency Unit were included in the study. Patients with medical complications or medications that can affect uterine wound healing were not included in the study. Information regarding demographics, body mass indices (BMIs), smoking and drinking habits, clinical features, past pregnancies, number and history of abortions, interdelivery period, gestational week, number of previous cesarean deliveries, fetal presentation, fetal weight, toco- graphic data, transabdominal ultrasonographic measurement of LUS full thickness and myometrium thickness, and lower uterine segment findings during cesarean delivery were collected. A total of 24 patients were excluded from the study because LUS could not be screened properly because of bladder adhesions. Transabdominal sonographic examination was carried out with a full urinary bladder to allow good imaging of the LUS. All examinations were performed on a GE Voluson S10 series 2−4 MHz probe (GE Healthcare, Chicago, IL) by a single sonographer. The patients’ labor and delivery outcomes were reviewed, and any evidence of rupture or dehiscence was noted. Nine additional patients were excluded from the study because of previous incisions that were not transverse.

**Statistical analysis**

Statistical analysis was performed with Mann–Whitney U, chi-square, and Fisher exact tests. P<.05 was considered significant. To increase the performance of ML methods, in the data preprocessing stage we eliminated the features that were found to be less relevant to uterine defects. For determining the features to be omitted, we used statistical significance and the information gain and gain ratio metrics.18 Of the 22 features collected, we selected 10 features on the basis of their statistical significance, information gain ranking, and gain ratio ranking. Because different ML algorithms perform better than others in different applications, logistic regression, multilayer perceptron, support vector machine, random forest, and naive Bayes algorithms19 were all...
trained and tested in this study to assess which algorithm outperforms the others. Because the dataset is relatively small and the prevalence of uterine dehiscence and rupture in general is relatively low, the data were not divided into training and validation sets. Instead, the performance of the algorithms was measured by using 10-fold cross-validation. Multiple performance criteria were selected, including the correct classification rate (CCR), F-score, Matthews correlation coefficient (MCC), precision-recall curve (PRC) area, and receiver operating characteristic (ROC) area to accurately measure and compare the performances.

**Results**

The mean age and mean BMI of the patients were found to be 28.74±4.93 and 29.22±4.41, respectively; 11.1% of the patients reported to be smokers. The mean number of miscarriages was 0.30±0.67, whereas 14 (4.4%) of the patients had dilation and curettage (d&c). Gestational age was ≤39 weeks for 77.9% of patients and >39 for 22.1%.

A total of 183 (57.7%) patients had contractions recorded on cardiotocography (CTG), and the mean value in Montevideo units (MVU) was 121.72±33.59. Full LUS measurements were <1 mm for 4 (1.3%), 1 to 2 mm for 45 (14.2%), and ≥2 mm for 268 (84.5%) of the patients. Estimated fetal weight was ≤3000 g for 90 (28.4%) and >3000 g for 227 (71.6%) of the patients. Lastly, intraoperative dehiscence was observed in 23 (7.3%) of the patients. All of the dehiscence cases observed were asymptomatic.

**Table 1** shows the quantitative variables’ means, standard deviations, and median values in case of dehiscence and no dehiscence. Statistical significance was observed for age, number of previous cesarean deliveries, number of years passed since the first cesarean delivery, and effacement features (P<.001, P<.001, P<.001, and P=.034, respectively). The mean age and mean number of previous cesarean deliveries of the patients with uterine dehiscence were significantly higher than those of patients with no uterine dehiscence. Furthermore, patients with no dehiscence had their first cesarean delivery on average 5.80±3.59 years before, whereas patients with dehiscence had their first cesarean delivery 9.78±4.06 years before on average. Moreover, the extent of effacement was also significantly larger in patients with dehiscence (24.35±21.07) than in patients with no dehiscence (15.37±20.33).

Qualitative patient variables are presented in Table 2 in case of intraoperative dehiscence and no intraoperative dehiscence. Gravidity, parity, CTG-detected contractions, dilation, LUS full measurement, and LUS myometrium measurement were found to be statistically significant (P<.001, P<.001, P=.038, P=.034, P<.001, and P<.001, respectively). Among the group of patients with dehiscence, 69.6% were found to be gravidas >3, and 60.9% were found to be paras >2. Moreover, 78.3% had contractions demonstrated by CTG, 60.9% had dilation, 82.6% had LUS full measurement between 1 and 2 mm, and 95.7% had LUS full measurement of <1 mm.

Gain ratio and information gain results of the features are shown in Figures 1 and 2, respectively. Only the features with highest gain ratio and information gain scores were included in the graphics. Furthermore, the features in Figures 1 and 2.
and the features that were found to be statistically significant did not overlap because P value, information gain, and gain ratio are different metrics, each of which provides insight about the relationship between the features and uterine dehiscence.

The performance of the ML methods in terms of CCR (accuracy), F-score, MCC, PRC area, and ROC area is shown in Table 3. The features included in this model are LUS myometrium thickness, parity, gravidity, contraction, dilation, d&c, number of years since the previous cesarean delivery, gestational week, and estimated fetal weight. Of the features in Figures 1 and 2, LUS full thickness was excluded because including it with LUS myometrium thickness, with which it is highly correlated, degraded the performance of the ML algorithms (this issue is called multicollinearity in the literature).

When LUS full thickness was replaced with LUS myometrium thickness in the model, the performance of the ML methods changed, as shown in Table 4. It can be observed that the overall performance of the methods decreased slightly when LUS full thickness was used instead of LUS myometrium thickness.

### Table 2

Qualitative patient characteristics

| Features                  | Intraoperative dehiscence |   |   |   |
|---------------------------|---------------------------|---|---|---|
|                           | No            | %   | n  | %   | P value |
| Smoking                   | No            | 264 | 89.8 | 21 | 91.3 | 1.00 |
|                           | Yes           | 30  | 10.2 | 2  | 8.7  | 1.00 |
| Gravidity                 | ≤3            | 238 | 81   | 7  | 30.4 | <.001|
|                           | >3            | 56  | 19   | 16 | 69.6 | <.001|
| Parity                    | ≤2            | 268 | 91.2 | 9  | 39.1 | <.001|
|                           | >2            | 26  | 8.8  | 14 | 60.9 | <.001|
| D&C                       | No            | 283 | 96.3 | 20 | 87   | .072 |
|                           | Yes           | 11  | 3.7  | 3  | 13   | .072 |
| Gestational week          | ≤39           | 228 | 77.6 | 19 | 82.6 | .573 |
|                           | >39           | 66  | 22.4 | 4  | 17.4 | .573 |
| Years since previous CD   | ≤2            | 69  | 23.5 | 3  | 13   | .250 |
|                           | >2            | 225 | 76.5 | 20 | 87   | .250 |
| Contraction               | No            | 129 | 43.9 | 5  | 21.7 | .038 |
|                           | Yes           | 165 | 56.1 | 18 | 78.3 | .038 |
| Cervical dilation         | No            | 181 | 61.6 | 9  | 39.1 | .034 |
|                           | Yes           | 113 | 38.4 | 15 | 60.9 | .034 |
| Presentation              | Cephalic      | 287 | 97.6 | 23 | 100  | 1.00 |
|                           | Breech        | 5   | 1.7  | 0  | 0    | 1.00 |
|                           | Transverse    | 2   | 0.7  | 0  | 0    | 1.00 |
| LUS full (mm)             | <1            | 1   | 0.3  | 3  | 13   | <.001|
|                           | 1–2           | 26  | 8.8  | 19 | 82.6 | <.001|
|                           | >2            | 267 | 90.9 | 1  | 4.3  | <.001|
| LUS myometrium (mm)       | <1            | 26  | 8.8  | 22 | 95.7 | <.001|
|                           | 1–2           | 127 | 43.2 | 1  | 4.3  | <.001|
|                           | >2            | 141 | 48   | 0  | 0    | <.001|
| Estimated fetal weight (g)| ≤3000        | 82  | 27.9 | 7  | 30.4 | .821 |
|                           | >3000         | 212 | 72.1 | 16 | 69.6 | .821 |

CD, cesarean delivery; D&C, dilation and curettage; LUS, lower uterine segment.

Kement. Predicting uterine dehiscence via machine learning by using lower uterine segment thickness and clinical features. Am J Obstet Gynecol Glob Rep 2022.
To assess the importance of LUS sonography in predicting uterine dehiscence, the ML algorithms were run without either of the LUS measurements. The resulting performance of the ML methods with the features of parity, gravidity, contraction, dilation, years since the previous cesarean delivery, gestational week, and estimated fetal weight is shown in Table 5. It can be observed that the performance of the ML algorithms significantly degraded when LUS measurements were not taken into account.

**Comment**

**Principal findings**

In this study, the mean value of the dehiscence group’s age was found to be higher than that of the no-dehiscence group. No significant difference was found between the groups in BMI, number of miscarriages, number of miscarriages after cesarean delivery, and d&c. The number of previous cesarean deliveries was found to be higher in the dehiscence group. Contractions in MVUs and monitorization duration were not found to be significantly different between the groups, whereas effacement percentage was found to be higher in the dehiscence group. This was an expected result given that effacement is correlated with the thinning of the cervix. The number of years since the first cesarean delivery was also found to be significantly higher in the dehiscence group. At first glance, this might seem counterintuitive given that more time passed implies better-healed uterine scars. However, it also implies that the patient is older, which affects the wound healing process.

**TABLE 3**
Performance of the machine learning algorithms in terms of different metrics with lower uterine segment myometrium thickness included as a feature

| Method                | CCR   | F-score | MCC   | PRC   | ROC   |
|-----------------------|-------|---------|-------|-------|-------|
| Logistic regression   | 0.943 | 0.941   | 0.544 | 0.966 | 0.949 |
| Multilayer perceptron | 0.953 | 0.953   | 0.655 | 0.963 | 0.965 |
| Support vector machine| 0.956 | 0.955   | 0.659 | 0.937 | 0.816 |
| Naive Bayes           | 0.956 | 0.956   | 0.672 | 0.965 | 0.950 |
| Random forest         | 0.946 | 0.946   | 0.594 | 0.959 | 0.944 |

CCR, correct classification rate; MCC, Matthews correlation coefficient; PRC, precision-recall curve; ROC, receiver operating characteristic.  

* Best performance among the algorithms in terms of the metric specified.

Kement. Predicting uterine dehiscence via machine learning by using lower uterine segment thickness and clinical features. Am J Obstet Gynecol Glob Rep 2022.
No significant difference was observed between the 2 groups in terms of smoking, d&c, gestational week, years since the last cesarean delivery, presentation, and estimated fetal weight. It was observed that most patients in the dehiscence group were gravidas > 3 and paras > 2. The prevalence of contractions recorded on CTG and dilation was higher in the dehiscence group. Most patients in the dehiscence group had myometrium thickness of < 1 mm, whereas most patients in the no-dehiscence group had > 2 mm. Estimated fetal weight did not significantly differ between the 2 groups.

Three rounds of ML simulations were conducted with 3 different sets of features. It was observed that when LUS myometrium thickness or LUS full thickness were added to the features, the predictive performance of the ML algorithms increased significantly. In terms of accuracy, the naive Bayes algorithm performed 4.8% and 5.1% better when LUS full thickness and LUS myometrium thickness were used, respectively. The increases were 4% and 4.4%, 24.2% and 27.3%, 4.6% and 4.1%, 17.4% and 17% in terms of F-score, MCC, PRC area, and ROC area, respectively. Among the attempted ML methods, the naive Bayes yielded the best classification performance, followed closely by the support vector machine.

**Results in the context of what is known**

In the study by Lipschuetz et al. that attempted to predict vaginal birth after cesarean delivery (VBAC) with ML, the ROC value was found to be 0.745 with the data collected from the first antenatal visit, and 0.793 with the added delivery unit admission features. The ROC value in our study was found to be as high as 0.966. There are 3 main factors that result in this difference. First, our study attempted to predict uterine defects as opposed to VBAC. Second, Lipschuetz et al. used parity, age, gestational week, previous delivery newborn weights, previous VBACs, dilation, and presentation as features. In this work we showed that except parity, these features have little to no effect on the performance of the ML algorithms (Figures 1 and 2). Lastly, we used boosting methods (gradient boosting) to increase the performance of the ML algorithms.

### TABLE 4

| Method                  | CCR  | F-score | MCC  | PRC  | ROC  |
|-------------------------|------|---------|------|------|------|
| Logistic regression     | 0.950| 0.948   | 0.610| 0.970 | 0.965|
| Multilayer perceptron   | 0.940| 0.938   | 0.528| 0.969 | 0.966|
| Support vector machine  | 0.953| 0.951   | 0.628| 0.932 | 0.794|
| Naive Bayes             | 0.953| 0.952   | 0.641| 0.970 | 0.954|
| Random forest           | 0.934| 0.932   | 0.478| 0.963 | 0.946|

CCR, correct classification rate; MCC, Matthews correlation coefficient; PRC, precision-recall curve; ROC, receiver operating characteristic.

### TABLE 5

| Method                  | CCR  | F-score | MCC  | PRC  | ROC  |
|-------------------------|------|---------|------|------|------|
| Logistic regression     | 0.924| 0.901   | 0.160| 0.912 | 0.730|
| Multilayer perceptron   | 0.909| 0.895   | 0.136| 0.904 | 0.702|
| Support vector machine  | 0.927| 0.893   | 0.001| 0.865 | 0.500|
| Naive Bayes             | 0.905| 0.912   | 0.399| 0.924 | 0.780|
| Random forest           | 0.918| 0.901   | 0.172| 0.899 | 0.678|

CCR, correct classification rate; MCC, Matthews correlation coefficient; PRC, precision-recall curve; ROC, receiver operating characteristic.

Kement. Predicting uterine dehiscence via machine learning by using lower uterine segment thickness and clinical features. Am J Obstet Gynecol Glob Rep 2022.
**Clinical implications**
This study showed that ultrasonographic measurement of LUS can be used to aid the assessment of the risk of uterine rupture. Hence, the number of untimely cesarean deliveries and their complications can be reduced.

**Research implications**
Our study demonstrated that the performance of ML algorithms is sufficient to justify their use in clinical applications. With a bigger dataset, more complex and fine-tuned models (such as deep learning and reinforcement learning) can be used to achieve greater predictive performance.

**Strengths and limitations**
This study was conducted on a smaller dataset compared with other (nonmedical) applications of ML. Furthermore, it suffered from imbalance (7.3% dehiscence vs 92.7% no-dehiscence), a typical limitation of medical datasets. However, completeness of data and high number of features (22) collected in the dataset helped to increase the performance of the ML algorithms.

**Conclusions**
ML methods can be used to predict uterine dehiscence and hence help assess possible rupture. Thus, the complications caused by untimely cesarean deliveries can be limited. Ultrasoundographic measurement of the LUS significantly increases the prediction performance of the ML algorithms.

**REFERENCES**

1. Mozurkewich EL, Hutton EK. Elective repeat cesarean delivery versus trial of labor: a meta-analysis of the literature from 1989 to 1999. Am J Obstet Gynecol 2000;183:1187-97.
2. Solheim KN, Esakof TF, Little E, Chang YW, Sparks TN, Caughhey AB. The effect of cesarean delivery rates on the future incidence of placenta previa, placenta accreta, and maternal mortality. J Matern Fetal Neonatal Med 2011; 24:1341-6.
3. Jastrow N, Demars S, Challet N, et al. Lower uterine segment thickness to prevent uterine rupture and adverse perinatal outcomes: a multicenter prospective study. Am J Obstet Gynecol 2016;215:604.e1-6.
4. Cui X, Wu S. Ultrasonic assessment has high sensitivity for pregnant women with previous cesarean section. J Obstet Gynaecol Can 2005;27:674-81.
5. Cheung VYT. Sonographic measurement of the lower uterine segment thickness in women with previous caesarean section. J Obstet Gynaecol Can 2005;27:674-81.
6. Sen S, Malik S, Salhan S. Ultrasonographic evaluation of lower uterine segment thickness in patients of previous cesarean section. Int J Gynaecol Obstet 2004;87:215-9.
7. Gizzo S, Zambon A, Saccardi C, et al. Effectiveness of hysteroscopically assisted myomectomy and abdominal hysterectomy. Am J Obstet Gynecol 2000;183:1187-97.
8. Mozurkewich EL, Hutton EK. Elective repeat cesarean delivery versus trial of labor: a meta-analysis of the literature from 1989 to 1999. Am J Obstet Gynecol 2000;183:1187-97.
9. Swift BE, Shah PS, Farine D. Sonographic lower uterine segment thickness after prior caesarean section to predict uterine rupture: a systematic review and meta-analysis. Acta Obstet Gynecol Scand 2019;98:330-41.
10. Asakura H, Nakai A, Ishikawa G, Suzuki S, Araki T. Prediction of uterine dehiscence by measuring lower uterine segment thickness prior to the onset of labor evaluation by transvaginal ultrasonography. J Nippon Med Sch 2000;67:352-6.
11. Kok N, Wiersma IC, Opmeer BC, De Graaf IM, Moll BW, Paikrt E. Sonographic measurement of lower uterine segment thickness to predict uterine rupture during a trial of labor in women with previous caesarean section: a meta-analysis. Ultrasound Obstet Gynecol 2013;42:132-9.
12. Sanz Lorenzo O, Farina A, Pula G, et al. Sonographic evaluation of the lower uterine segment thickness in women with a single previous cesarean section. Minerva Ginecol 2013;65:551-5.
13. Kushtagi P, Carepalli S. Sonographic assessment of lower uterine segment at term in women with previous cesarean delivery. Arch Gynecol Obstet 2011;283:455-9.
14. Bujold E, Jastrow N, Simonneau J, Brunet S, Gauthier RJ. Prediction of complete uterine rupture by sonographic evaluation of the lower uterine segment. Am J Obstet Gynecol 2009;201:320.e1-6.
15. Lipschutz M, Guedalia D, Rotteneich A, et al. Prediction of vaginal birth after cesarean deliveries using machine learning. Am J Obstet Gynecol 2020;222. 613.e1-12.
16. Lower uterine segment thickness measurements and clinical features dataset. 2022. Available at: https://doi.org/10.17632/9hr8rxd2r2d.1. Accessed September 1, 2022.
17. Harris, E. Information Gain Versus Gain Ratio: A Study of Split Method Biases. AI&M 2002.