


BMJ Open Using machine learning to identify quality-of-care predictors for emergency caesarean sections: a retrospective cohort study

Betina Ristorp Andersen,¹ Ida Ammitzbøll,¹ Jesper Hinrich,² Sune Lehmann,² Charlotte Vibeke Ringsted,³ Ellen Christine Leth Løkkegaard ¹, Martin G Tolsgaard⁴

To cite: Andersen BR, Ammitzbøll I, Hinrich J, *et al.* Using machine learning to identify quality-of-care predictors for emergency caesarean sections: a retrospective cohort study. *BMJ Open* 2022;**12**:e049046. doi:10.1136/bmjopen-2021-049046

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2021-049046>).

Received 13 January 2021
Accepted 01 February 2022



© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Professor Ellen Christine Leth Løkkegaard;
Ellen.Christine.Leth.
Loekkegaard@regionh.dk

ABSTRACT

Objectives Emergency caesarean sections (ECS) are time-sensitive procedures. Multiple factors may affect team efficiency but their relative importance remains unknown. This study aimed to identify the most important predictors contributing to quality of care during ECS in terms of the arrival-to-delivery interval.

Design A retrospective cohort study. ECS were classified by urgency using emergency categories one/two and three (delivery within 30 and 60 min). In total, 92 predictor variables were included in the analysis and grouped as follows: 'Maternal objective', 'Maternal psychological', 'Fetal factors', 'ECS Indication', 'Emergency category', 'Type of anaesthesia', 'Team member qualifications and experience' and 'Procedural'. Data was analysed with a linear regression model using elastic net regularisation and jackknife technique to improve generalisability. The relative influence of the predictors, percentage significant predictor weight (PSPW) was calculated for each predictor to visualise the main determinants of arrival-to-delivery interval.

Setting and participants Patient records for mothers undergoing ECS between 2010 and 2017, Nordsjællands Hospital, Capital Region of Denmark.

Primary outcome measures Arrival-to-delivery interval during ECS.

Results Data was obtained from 2409 patient records for women undergoing ECS. The group of predictors representing 'Team member qualifications and experience' was the most important predictor of arrival-to-delivery interval in all ECS emergency categories (PSPW 25.9% for ECS category one/two; PSPW 35.5% for ECS category three). In ECS category one/two the 'Indication for ECS' was the second most important predictor group (PSPW 24.9%). In ECS category three, the second most important predictor group was 'Maternal objective predictors' (PSPW 24.2%).

Conclusion This study provides empirical evidence for the importance of team member qualifications and experience relative to other predictors of arrival-to-delivery during ECS. Machine learning provides a promising method for expanding our current knowledge about the relative importance of different factors in predicting outcomes of complex obstetric events.

Strengths and limitations of this study

- The application of a penalised and bias corrected machine-learning approach which is advantageous in explorative studies for identifying patterns in large and complex data set.
- The statistical approach included penalised regularisation that imposes a penalty on the model by shrinking the contribution of extreme and less important predictors.
- Given the exploratory nature of this approach and that we only considered caesarean sections in one hospital, our findings need to be reproduced to determine which are context-specific and which are generalisable.
- Category one and two emergency caesarean sections were analysed together in order to compare our results to the international literature.

INTRODUCTION

In Western Europe and North America, the safety of childbirth is generally high.¹ Maternal and fetal mortality during childbirth is infrequent,¹ leading to an increased focus on the prevention of maternal and fetal morbidity by improving quality of care.² In developed countries, emergency caesarean sections (ECS) are associated with up to five times increased risk of maternal morbidity compared with vaginal delivery.³

An ECS is a time-sensitive, unplanned surgical procedure often performed in pregnancies, where vaginal delivery was first planned or attempted. To increase safety for the mother and the unborn child, monitoring the decision-to-delivery interval (DDI) has become an important part of obstetric audit.⁴ Key factors to reduce DDI suggested by prior studies include clinical urgency category,⁵⁻⁷ type of anaesthesia used,^{7,8} cervical dilatation,⁷ previous delivery by caesarean section,⁶

maternal obesity⁹ and whether the ECS was performed during the day or nighttime.^{6 7} Additionally, communication skills between team members,⁵ seniority of the surgeon^{6 7} and midwifery staff level⁸ have been found to correlate to quality of care.

Each study has provided information about single predictors of quality of care in ECS. However, to our knowledge, no studies have compared and determined the relative importance of these predictors in determining critical ECS outcomes. Advances in machine learning and statistics, such as the use of penalised regression models, are particularly well-suited for analysing large and complex data sets to identify, to find patterns and make predictions that are often misrepresented by traditional statistical techniques.¹⁰

DDI is an internationally recognised quality indicator that represents a series of events during four time intervals: Interval I, the decision by the obstetrician to transfer to the operating room; Interval II, patient arrival in the operating room until induction of anaesthesia; Interval III, from anaesthesia induction to surgical incision; and Interval IV, from surgical incision to delivery of the infant.¹¹ Intervals II–IV (arrival-to-delivery) are characterised by significant complexity due to the multidisciplinary team, which includes anaesthesia staff and scrub nurses, in addition to obstetricians.

The aim of our study was to identify the most important factors contributing to quality of care during ECS based on assessments of the arrival-to-delivery (AD) interval using a machine learning approach.

MATERIAL AND METHODS

Data from 2892 patient records of mothers undergoing ECS between 1 January 2010 to 16 March 2017 at the Department of Gynecology and Obstetrics, Nordsjællands Hospital, Denmark, were included in a retrospective cohort study.

Caesarean sections were categorised based on urgency into four categories¹²: ‘Immediate threat to the life of the woman or the fetus (category one)’; ‘Maternal or fetal compromise not immediately life-threatening (category two)’; ‘Needing early delivery but no maternal or fetal compromise (category three)’; ‘Delivery at a time suitable to the patient and maternity team (category four)’.¹² Our study only included category one, two and three ECS, in which urgency and DDI are considered of importance.

The ECS team aims to deliver the baby as quickly as possible without subjecting the mother to unnecessary risk. In Denmark, the standard goal is to deliver the baby within 15 min (following the decision by the obstetrician to pursue ECS) for category one and 30 min for category two ECS.¹³ In the USA and Great Britain, time recommendations for DDI intervals are 30 min.^{11 13–15} To compare our results to the international standard, we analysed data from categories one and two together.

Predictor variables

Demographic data collected from patient records included maternal age, gestational age at delivery and parity.

Predictor variables were selected, based on a literature review^{5–9 11 13 16–21} (online supplemental table 1). We searched the MEDLINE database and Google Scholar for publications describing quality of care predictors for ECS using terms as ‘quality-of-care’, ‘Emergency-Caesarean-sections’, ‘decision-to-delivery’, ‘Anaesthesia’ and ‘Obesity’.^{5–9 11 13 16–21} We hand searched references in key publications to help inform the content and selections of predictor variables in our study.

Predictors were grouped under the following headings: ‘Maternal objective predictors’, N=21; ‘Maternal psychological’ predictors, N=4; ‘Fetal predictors’, N=7; ‘Indication for ECS’, N=11; ‘Emergency category of ECS’, N=5; ‘Type of anaesthesia’, N=5; and ‘Procedural predictors’ N=2. Team members responsible for patient care were identified by name and title. The group ‘Team member qualifications and experience’ included educational characteristics extracted from the Danish Health Authorities and years of employment in the department. The qualifications for the individual team member were updated yearly. For this data set, a physician in his or her first year of training was identified as a resident, whereas a trainee in the second to fifth year of training was identified as a senior resident. A designation of obstetrician consultant indicated a physician who is engaged full-time in caring for obstetric patients, whereas a gynaecological consultant referred to a physician who is responsible for outpatient gynaecologic patients and surgical procedures during the day but takes the nighttime call on labour and delivery.

To assess the importance of experience, we registered the number of ECS performed by a team member during the year before the ECS study case. Additionally, the presence of an extended team was included in the group ‘Team member qualifications and experience’, N=37 (online supplemental table 1).

Primary outcome

AD interval represents DDI II–IV. We chose AD as our primary outcome to focus on aspects of care that can be improved in an OR setting. This outcome variable excludes DDI Interval I that represents the time from the decision to perform ECS until transfer to the operating room (labour ward and transfer).¹¹ Local hospitals vary considerably in terms of standard protocols for Interval I. Moreover, the complex multidisciplinary operating room teams do not play a role in Interval I. AD interval was measured in minutes. Intervals above 90 min were excluded from analysis (N=64) as values above this threshold were likely to be recording errors.

Statistical methods

We were interested in identifying the most important predictor variables out of a large set of 92 factors related to each ESC and the associated operating room team.

Due to the large number of predictor variables relative to the total number of ECSs, we applied a machine learning method with predictor selection. Specifically, we use linear regression with elastic net regularisation^{22 23} and split the data 80–20 into development and test observations, the latter intended to simulate newly acquired observations. Estimated model parameters were based solely on the development set observations. This set was analysed using leave-one-out cross-validation/jackknife sampling²⁴ to estimate predictor variable importance. For each jackknife sample, fivefold cross validation²⁵ was used to estimate the optimal Elastic Net parameter. See Statistical Appendix (SA) for full details.

We estimated the importance of predictor variables using *percentage significant predictor weight* (PSPW).²⁶ The PSPW is simply a linear scaling of the significant predictor weights estimated by the model, such that the sum of all absolute PSPW values is 100. Larger absolute PSPW implies greater influence.

Patient and public involvement

The research question explored in this study is highly relevant to all future mothers undergoing ECS. The study provides insight into the most important predictors for quality-of care and could inform future quality improvement measures. The study is based on prior studies of patient-perceived quality of care in the simulated and clinical setting.^{26 27} Patients were not involved in the establishment of the present retrospective cohort study.

RESULTS

The study included data from 2892 patient records. Sixty-four ECSs were excluded because of an AD interval >90 min, 150 ECSs were discarded due to data inconsistency (zero or negative AD interval, impossible or clearly erroneous predictor variable combinations, no category label) and 135 due to multiple pregnancy (which could influence analysis in a non-obvious way). Of the remaining 2543 ECS, 134 were categorised as category four (planned caesarean section) and therefore withdrawn from further analysis. Of the remaining 2409 ECSs, 165 fell into category one, 1022 in category two and 1222 in category three.

Descriptive data for included mothers (maternal age, body mass index (BMI), parity, prior caesarean section (CS), gestational age and fetal weight) are shown in table 1.

All ECSs were performed by at least two surgeons, however, there were 709 ECSs with no name registered for ‘surgeon 2’, for whom the competency level could not be determined. The pairwise surgeons’ competencies and the corresponding AD interval are illustrated in table 2. In ECS category one/two, 46% were performed by an attending gynaecologist or obstetrician as surgeon 1 and 54% were performed by a resident. In category three, 42.3% were performed by an attending and 57.7% were performed by a resident.

The ECSs were analysed in four groups: only category one, only category two, only category three and category one/two. The estimated Elastic Net models outperformed the null models and captured patterns in the data not accounted for by the baseline (median outcome on the development set) or random variation (permutation testing) in all four groups. In the majority of the ECS groups (3 out of 4), the Elastic Net models more accurately predicted ($R^2 > 0$) the test observations than using their true (but unknown) mean outcome ($R^2 = 0$). This means that the patterns captured by the model can be generalised to unseen data examples (new ECS), as simulated by holding out 20% of the ECS for testing.

When modelling only ECS category one, the model did not outperform the true mean outcome. This group had the lowest number of observations, that is, N=140 for development and N=25 for the test set.

The test set performance for the ECS category two was only slightly improved and was not analysed further, that is, AD Root Mean Square Error (RMSE=6.01 $R^2=0.05$). The performance in ECS emergency category three was RMSE=9.05, $R^2=0.11$, and combined ECS emergency category one/two RMSE=5.61, $R^2=0.24$.

Prediction results

The most important group of predictors associated with AD interval was ‘Team member qualifications and experience’ for all emergency categories (PSPW 25.9% for category one/two and PSPW 35.5% for category three).

Table 1 Descriptive data of characteristics by emergency caesarean section category

	Emergency category		
	Mean (Range) Category one/two	Mean (Range) Category three	Mean (Range) No information about category
Maternal age	31.2 (16–48)	31.4 (15–47)	31.0 (20–44)
Maternal body mass index (kg/m ²)	24.8 (15.2–49.8)	25.4 (15.4–51.6)	25.4 (18.0–40.9)
Maternal parity	0.55 (0–6)	0.56 (0–10)	0.45 (0–4)
Number of prior caesarean section	0.23 (0–4)	0.3 (0–4)	0.20 (0–2)
Gestational age (weeks+days)	39+5 (24+3–44+0)	38+6.3 (25+6–43+1)	39+4.1 (28+1–42+3.0)
Fetal weight (g)	3439.35 (495–5640)	3353.64 (685–5530)	3573.27 (1140–5800)

Table 2 Surgeon pairs and arrival-to-delivery interval in emergency category one/two and three

Surgeon pairs qualifications 'surgeon' + 'surgeon 2'	Emergency category				
	One–three totals		Mean (SD) Minutes	Three	
	N (%)	N (%)		N (%)	Mean (SD) Minutes
Gynaecologist +missing	127 (5.3)	80 (6.7)	18.3 (0.9)	47 (3.8)	26.1 (1.6)
Gynaecologist +gynaecologist	8 (0.3)	2 (0.2)	15.5 (0.5)	6 (0.5)	26.8 (4.7)
Gynaecologist +obstetrician	8 (0.3)	5 (0.4)	20.4 (3.0)	3 (0.2)	23.3 (4.6)
Gynaecologist +resident	185 (7.7)	112 (9.4)	19.3 (0.7)	73 (6.0)	27.0 (1.2)
Obstetrician +missing	262 (10.9)	119 (10.0)	20.1 (1.0)	143 (11.7)	28.6 (0.7)
Obstetrician +gynaecologist	15 (0.6)	2 (0.2)	25.0 (7.0)	13 (1.0)	32.7 (2.9)
Obstetrician +obstetrician	70 (2.9)	23 (1.9)	18.2 (1.3)	47 (3.8)	29.3 (1.5)
Obstetrician +resident	391 (16.3)	204 (17.2)	18.8 (0.5)	187 (15.3)	28.8 (0.7)
Senior resident +missing	113 (4.7)	61 (5.2)	18.9 (0.8)	52 (4.3)	30.0 (1.3)
Senior resident +gynaecologist	92 (3.8)	56 (4.7)	17.8 (0.9)	36 (2.9)	29.2 (1.7)
Senior resident +obstetrician	137 (5.7)	76 (6.4)	19.8 (0.9)	61 (5.0)	31.2 (1.2)
Senior resident +resident	325 (13.5)	183 (15.4)	19.2 (0.5)	142 (11.6)	29.1 (0.8)
Resident +missing	68 (2.8)	33 (2.8)	17.0 (1.0)	35 (3.0)	28.2 (1.3)
Resident +gynaecologist	136 (5.6)	57 (4.8)	19.1 (0.8)	79 (6.5)	27.0 (0.9)
Resident +obstetrician	249 (10.3)	79 (6.7)	19.6 (0.7)	170 (13.9)	30.3 (0.6)
Resident +resident	223 (9.3)	95 (8.0)	19.6 (0.8)	128 (10.5)	30.6 (1.0)
Total number	2409 (100)	1187 (100)		1222 (100)	

The second most important predictor group was 'Indication for ECS' in category one/two ECS (PSPW 24.9%) and 'Maternal Objective predictors' in category three ECS (PSPW 24.2%). The relative importance measured by PSPW is shown in figures 1 and 2 for ECS categories one/two and three, respectively.

Individual predictor weights are shown in online supplemental table 2.

Surgical and delivery difficulties were strong predictors of longer AD intervals in both ECS category one/two and three.

In emergency category one/two, the level of seniority of Surgeon 2 was a strong predictor of AD interval. Having a gynaecologist or obstetrician as Surgeon 2 predicted shorter AD, whereas having a trainee as the assistant predicted longer AD. In ECS category three, the level of seniority of Surgeon 1 was the strongest predictor. Having a gynaecologist or obstetrician as Surgeon 1 predicted shorter AD, whereas having a trainee as Surgeon 1 predicted longer AD. Also, the presence of three or more surgeons was a strong predictor of longer AD.

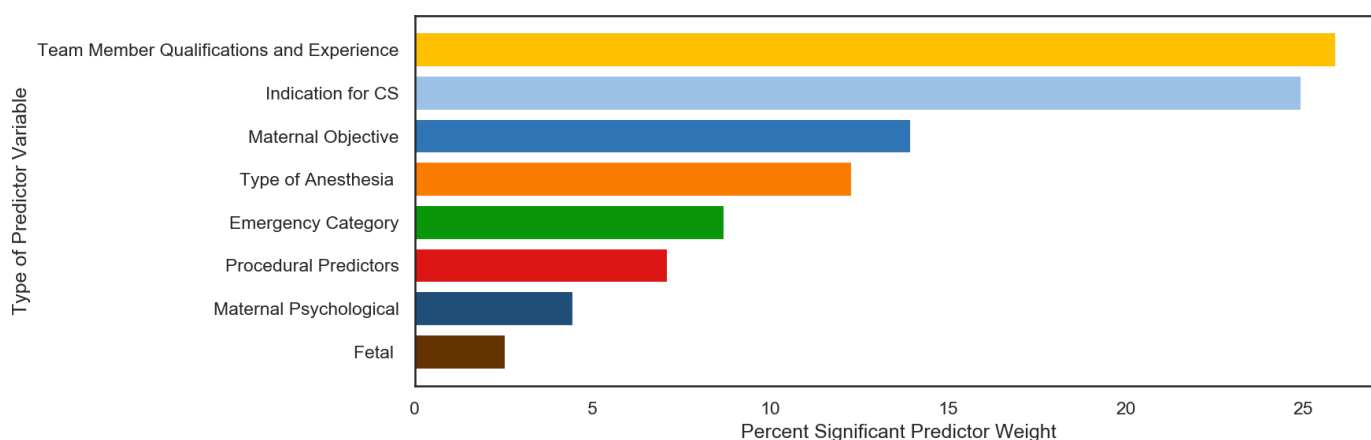


Figure 1 Percentage significant predictor weight (PSPW) for emergency category one/two. Arrival-to-delivery interval PSPW for all predictor groups emergency caesarean section category one/two: 'Team member qualifications and experience'= 25.9%, 'Indication for CS'= 24.9, 'Maternal objective'=14%, 'Type of anaesthesia'=12.3%, 'CS emergency category'=8.7%, procedural predictors=7.1%, 'Maternal psychological'=4.5%, 'Fetal' predictors=2.6%. CS, caesarean section.

Figure 2. Percentage significant predictor weight (PSPW) for Emergency Cesarean Sections, category three.

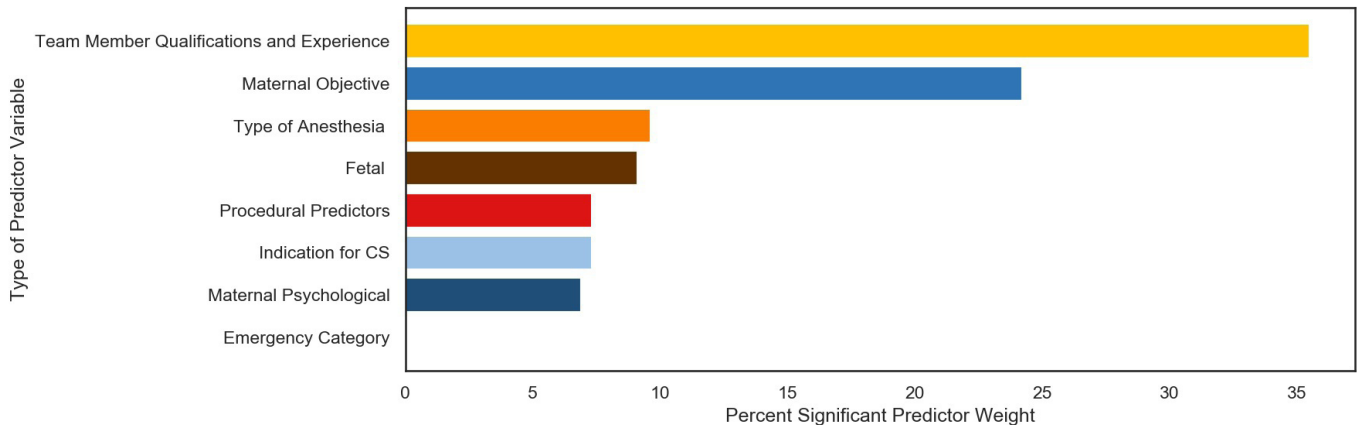


Figure 2 Percentage significant predictor weight (PSPW) for emergency category three. Arrival-to-delivery interval PSPW for all predictor groups emergency caesarean section category three: ‘Team member qualifications and experience’=35.5%, ‘Maternal objective’=22.2%, ‘Type of anaesthesia’=9.6%, ‘Fetal’ predictors=9.1%, procedural predictors=7.3%, ‘Indication for CS’=7.3, ‘Maternal psychological’=6.9%, emergency category=0%. CS, caesarean section.

Regarding the other team members who were not surgeons, having two or more members of a specialty, that is, anaesthesia nurse, anaesthesiologist or circulating nurse, corresponded to a higher Percentage Significant Predictor Weight (PSPWs) than their PSPWs corresponding to their seniority; that is, the number of non-surgeon team members was more important than their experience levels.

‘Last year CS-experience’ corresponded with relatively low PSPWs in general (<1.0%), the highest weights were for Surgeon 1 and the anaesthesiologist, both associated with shorter AD intervals.

Indication for CS related to high emergency threats, that is, placental abruption, fetal stress, uterine rupture and cord prolapse, were important predictors of shorter AD interval in ECS category one/two, whereas maternal request/exhaustion and planned elective CS predicted longer AD intervals in all ECS categories.

Placenta previa was a strong predictor of prolonged AD interval in ECS category three (PSPW 6.4%) but had less importance in ECS category one/two (PSPW 1.1%). BMI had low importance in all emergency categories, whereas prior CS predicted prolonged AD intervals in ECS category three.

‘Fetal predictors’, such as fetal weight or descent of fetal head had little importance in ECS with high emergency pressure, whereas in ECS category three, fetal descent to the pelvic floor was associated with a longer AD interval (PSPW 4.2%).

The category/group ‘Type of anaesthesia’ was associated with high PSPW. Epidural anaesthesia was associated with a shorter AD interval in all ECS emergency categories. Change of anaesthesia, that is, from epidural or spinal to general anaesthesia, predicted a longer AD interval.

DISCUSSION

In this study, we explored the association between numerous factors, as predictors of the AD interval in ECS.

Our use of machine learning made it possible to analyse many predictors simultaneously and, in turn, to identify those that were of greatest relative importance.

Our model demonstrated that ‘Team member qualifications and experience’ was by far the most important group of predictors associated with AD interval in all ECS emergency categories (one/two and three). ‘Maternal objective predictors’ and ‘Type of anaesthesia’ also emerged as critical factors in all ECS categories, whereas ‘Procedural predictors’ appeared to be a less important predictor (medium weight). In ECS category one/two, ‘Indication for ECS’ demonstrated high predictor weight, whereas ‘Fetal predictors’ was a low weight factor. By contrast, in emergency category three, the group of ‘Fetal predictors’ was shown to be of increased importance.

A strength of our study was the use of a machine-learning approach to analyse handling a large and complex data set. A penalised regression approach in the statistical analysis helped identify the most important associations within a large set of predictor variables (online supplemental appendix 1). However, concerns exist including issues of correlated features and overfitting. To mitigate the effect of correlated features, we used jackknife bias-correction, and to avoid overfitting we used cross-validation. This method has been described and used in a previous exploratory study,²⁶ in which the primary objective was to navigate and explore through a large and complex data set. Given the exploratory nature of this approach, our findings need to be reproduced to determine which are context-specific and which are generalisable.

In this study the outcome was the AD interval, which includes three smaller intervals (DDI II: arrival in the operating room until induction of anaesthesia; DDI III: induction of anaesthesia to surgical incision; DDI VI: surgical incision to delivery of the infant).¹¹ Our data did not allow subanalyses in DDI II, DDI II or DDI IV as



any effects may be dwarfed due to variability and lack of power.

The importance of individual team member's qualifications in healthcare has been demonstrated in prior studies,^{27–30} that is, availability of a trained midwife, having obstetricians and anaesthetists on-site near the labour ward^{5 8} and experience of the surgeon.⁶ In our study, we included the entire multidisciplinary team of surgeons, anaesthetists, nurses and midwives. Although the surgeons contributed most to the model, our inclusion of *all* professionals on the team expands on existing knowledge and emphasises that all team members have an impact on AD intervals during ECS.

In our clinical context, only 5% of the included ECS were performed solely by obstetrics/gynaecology specialists. Former studies report longer operating time if a resident participates in an operation.^{31–33} However, in one study focusing *exclusively* on caesarean sections, resident participation had little impact on OR-time, and resident participation was recommended even in difficult ECS.³¹ In our study with extensive resident involvement in ECS, we did not find that this factor predicted a longer AD interval. Although the study was not designed to directly assess the impact of resident involvement, our findings did not suggest any detrimental effects of resident participation in ECS.

We measured individual experience by how many ECS they had participated in the preceding year for each professional. Individual competency had little impact on AD in our statistical model across all ECS categories. This finding is somewhat surprising given that surgeons are thought to continuously improve their skills through repeated and deliberate practice.^{34–37} Our study focused exclusively on emergency CS. Hence, the variability in the context and composition of teams may outweigh factors related to variations in individual skill levels.

Procedural predictors, such as surgical and delivery difficulties, had high predictor weight in all emergency categories and were associated with longer AD intervals. Surgical difficulties encompass a woman's previous history and anatomy, including previous abdominal or pelvic surgeries, intraabdominal adhesions and prior CS.^{38 39} Delivery difficulties include situations such as an impacted fetal head requiring an assistant to push the head of the child upwards before delivery or a transition to breech delivery.⁴⁰ The reason for surgical and delivery difficulties includes a combination of individual competences, team-competences and maternal characteristics.^{5 38–41} Our study adds to existing knowledge by demonstrating the relatively high importance of delivery and surgical difficulties compared with other predictors for AD.

In our study, ECS indication was an important predictor in category one/two ECS. Indications such as maternal request and planned CS were associated with longer AD interval, and more urgent indications for ECS as 'placental abruption', 'fetal stress', 'rupture of a uterus scar' or 'cord prolapse' were associated with a

shorter AD interval. Other studies have reported similar associations.^{7 42} However, our results add to our existing knowledge by demonstrating their relative importance as compared with maternal objective predictors and the applied type of anaesthesia.

In the group of maternal predictors, we found a strong association between prior caesarean section and placenta previa with longer AD interval. Our finding of prolonged AD interval in cases of prior CS is supported by other studies.⁴³ Placenta previa is associated with adverse maternal outcomes, such as excessive bleeding and postpartum hysterectomy.⁴⁴ Consequently, in low urgency ECS, the team will choose to prepare for excessive bleeding before incision resulting in a prolonged AD interval. Interestingly, BMI, and maternal age demonstrated little importance in our model. Other studies have found a prolonged decision-to-delivery interval in association with maternal age⁴⁴ and obesity.^{17 44–46} Local guidelines for prophylactic epidural analgesia during delivery in obese women may have influenced our findings.

Clinical implications

The findings in our study provide new information about the importance of team member qualifications and experience in a statistical model, which enabled us to include many predictor variables representing different aspects of ECS. Although we included multiple known predictors, such as maternal characteristics and psychological predictors, fetal predictors and anaesthesia-related predictors, the group of predictors including team member qualifications and experience was found to be of greater importance than we had anticipated. Our findings strengthen the importance of team competencies during a complex/multifaceted event, such as an emergency caesarean section. Training of individual competencies and/or team-performance should have a higher priority when the aim is to reduce arrival-to-delivery interval. Resources should be prioritised for skills-training and team-training to ensure low AD interval, whenever possible. Finally, using statistical models such as those developed in our study for providing feedback to the ad-hoc team during and after managing an ECS may provide valuable real-time information. Such feedback can be used to guide and improve future performances of ad-hoc surgical teams during clinical practice or as input for team training interventions in a simulated setting.

Research implications

Our use of machine learning to analyse the data provided the opportunity to uncover patterns in a complex obstetric data set. Our approach was intended to provide a hierarchy of variables that were most relevant to the outcome. Even though machine learning is used in several areas of healthcare, it is often criticised for being a 'black box' approach and the results are often met with scepticism.⁴⁷ In our study, we sought transparency in several ways. First, the predictors included in our statistical model were based on prior studies, which had demonstrated associations

using traditional statistics. Moreover, we only included context-relevant predictors, carefully evaluated by a panel of expert obstetricians, to ensure that all analysed predictors were meaningful to AD. Second, we analysed single predictors instead of pooling the predictors, which also facilitated transparency of the model findings. Third, after the models provided information about important predictors, specialists within our research group (anaesthetists, obstetricians, gynaecologists) performed an additional check to ensure plausibility. Finally, to ensure transparency and provide the opportunity to verify our results, we have included a statistical appendix (online supplemental appendix 1).

CONCLUSIONS

This study provides empirical evidence for the importance of team member qualifications and experience relative to other predictors affecting AD during ECS. Machine learning-based analyses provide a promising method for expanding our current knowledge about the relative importance of the different predictors of complex and multifaceted obstetric events. Moreover, this methodology may help gain insight into other aspects of obstetric care quality.

Author affiliations

¹Department of Gynecology and Obstetrics, Nordsjællands Hospital & Department of Clinical Medicine, University of Copenhagen, Hillerød, Capital Region, Denmark

²Cognitive Systems, Department of Applied Mathematics and Computer Science, Technical University of Denmark, Lyngby, Denmark

³Faculty of Health, Aarhus Universitet, Aarhus, Denmark

⁴Copenhagen Academy of Medical Education and Simulation, Rigshospitalet, København, Capital Region, Denmark

Acknowledgements We thank Jesper Friis Petersen for setting up the database.

Contributors BRA contributed to the conception and design of the work, and to data acquisition and interpretation, data analysis and drafted the paper. IA contributed to data acquisition and interpretation and assisted in drafting the paper. MGT and ECLL contributed to the conception and design of the work and to data acquisition and interpretation, data analysis and assisted in drafting the paper. JH and SL contributed to data analysis and interpretation and assisted in the drafting of the paper. CVR contributed to the conception and design of the work and assisted in drafting the paper. All authors contributed to the critical revision of the paper and approved the final manuscript for publication. All authors have agreed to be accountable for all aspects of the work. BRA acts as a gaurantor of the study.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by Danish Patient Safety Authority (ID# 3-3013-1732/1). Data was collected from patient records.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement No data are available.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines,

terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Ellen Christine Leth Løkkegaard <http://orcid.org/0000-0003-4149-5663>

REFERENCES

- 1 Knight M, Nair M, Tuffnell D, eds. *on behalf of MBRRACE-UK. Saving lives, improving mothers' care - surveillance of maternal deaths in the UK 2012-14 and lessons learned to inform maternity care from the UK and Ireland confidential enquiries into maternal deaths and morbidity 2009-14*. Oxford: National perinatal epidemiology unit, University of oxford, 2016.
- 2 Brace V, Kernaghan D, Penney G. Learning from adverse clinical outcomes: major obstetric haemorrhage in Scotland, 2003-05. *BJOG* 2007;114:1388-96.
- 3 Sandall J, Tribe RM, Avery L, et al. Short-term and long-term effects of caesarean section on the health of women and children. *Lancet* 2018;392:1349-57.
- 4 Rosser J. Confidential enquiry into stillbirths and deaths in infancy (CESDI). highlights of the 6th annual report. Part 4: record keeping. *Pract Midwife* 2000;3:11.
- 5 Chauleur C, Collet F, Furtos C, et al. Identification of factors influencing the decision-to-delivery interval in emergency caesarean sections. *Gynecol Obstet Invest* 2009;68:248-54.
- 6 Livermore LJ, Cochrane RM. Decision to delivery interval: a retrospective study of 1,000 emergency caesarean sections. *J Obstet Gynaecol* 2006;26:307-10.
- 7 Kolås T, Hofoss D, Oian P. Predictions for the decision-to-delivery interval for emergency caesarean sections in Norway. *Acta Obstet Gynecol Scand* 2006;85:561-6.
- 8 Cerbinskaite A, Malone S, McDermott J, et al. Emergency caesarean section: influences on the decision-to-delivery interval. *J Pregnancy* 2011;2011:1-6.
- 9 Väänänen AJ, Kainu JP, Eriksson H, et al. Does obesity complicate regional anesthesia and result in longer decision to delivery time for emergency cesarean section? *Acta Anaesthesiol Scand* 2017;61:609-18.
- 10 Tibshirani R. Regression shrinkage and selection via the LASSO. *Journal of the Royal Statistical Society: Series B* 1996;58:pp. Vol.:267-88. Vol.1.
- 11 Gupta S, Naitani U, Madhanmohan C, et al. Evaluation of decision-to-delivery interval in emergency cesarean section: a 1-year prospective audit in a tertiary care hospital. *J Anaesthesiol Clin Pharmacol* 2017;33:64-70.
- 12 Lucas DN, Yentis SM, Kinsella SM, et al. Urgency of caesarean section: a new classification. *J R Soc Med* 2000;93:346-50.
- 13 Griffiths M. What is the acceptable decision-to-delivery interval for an emergency caesarean section? *BJOG* 2016;123:476.
- 14 RCOG. Nice guideline. cesarean section, 2011. Available: <https://www.nice.org.uk/guidance/cg132/resources/caesarean-section-pdf-35109507009733>
- 15 RCOG. Classification of urgency of caesarean section – a continuum of risk (good practice No. 11), 2010. Available: <https://www.rcog.org.uk/en/guidelines-research-services/guidelines/good-practice-11/>
- 16 Afolabi BB, Lesi FEA, Merah NA. Regional versus general anaesthesia for caesarean section. *Cochrane Database Syst Rev* 2006;4:CD004350.
- 17 Turner E, Warshak CR. Cesarean for non-reassuring fetal status: effect of obesity on decision to delivery interval. *J Perinatol* 2019;39:814-9.
- 18 Girsén AI, Osmundson SS, Naqvi M, et al. Body mass index and operative times at cesarean delivery. *Obstet Gynecol* 2014;124:684-9.
- 19 Rossouw JN, Hall D, Harvey J. Time between skin incision and delivery during cesarean. *Int J Gynaecol Obstet* 2013;121:82-5.
- 20 Greenberg MB, Daniels K, Blumenfeld YJ, et al. Do adhesions at repeat cesarean delay delivery of the newborn? *Am J Obstet Gynecol* 2011;205:380.e1-380.e5.
- 21 Palmer E, Ciechanowicz S, Reeve A, et al. Operating room-to-incision interval and neonatal outcome in emergency caesarean

- section: a retrospective 5-year cohort study. *Anaesthesia* 2018;73:825–31.
- 22 Zou H, Hastie T. Regularization and variable selection via the elastic net. *J Royal Statistical Soc B* 2005;67:301–20.
 - 23 Pedregosa F. Scikit-learn: machine learning in python. *the Journal of Machine Learning Research* 2011;12:2825–30 <https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf>
 - 24 Afanador NL, Tran TN, Buydens LMC. An assessment of the jackknife and bootstrap procedures on uncertainty estimation in the variable importance in the projection metric. *Chemometrics and Intelligent Laboratory Systems* 2014;137:162–72.
 - 25 Arlot S, Celisse A. A survey of cross-validation procedures for model selection. *Stat Surv* 2010;4:40–79.
 - 26 Andersen BR, Hinrich JL, Rasmussen MB, et al. Social ties between team members affect patient satisfaction: a data-driven approach to handling complex network analyses. *Adv Health Sci Educ Theory Pract* 2020;25:581–606.
 - 27 Andersen BR, Rasmussen MB, Christensen KB, et al. Making the best of the worst: care quality during emergency cesarean sections. *PLoS One* 2020;15:e0227988.
 - 28 Shalhoub J, Marshall DC, Ippolito K. Perspectives on procedure-based assessments: a thematic analysis of semistructured interviews with 10 UK surgical trainees. *BMJ Open* 2017;7:e013417-2016-013417.
 - 29 Norcini J, Talati J. Assessment, surgeon, and society. *International Journal of Surgery* 2009;7:313–7.
 - 30 Norcini JJ, Blank LL, Duffy FD, et al. The mini-CEX: a method for assessing clinical skills. *Ann Intern Med* 2003;138:476–81.
 - 31 Bougie O, Zuckerman SL, Switzer N, et al. Influence of resident involvement in obstetrics and gynaecology surgery on surgical outcomes: systematic review and meta-analysis. *J Obstet Gynaecol Can* 2018;40:1170–7.
 - 32 Mourad M, Gupta S, Rebarber A, et al. Effect of resident participation on outcomes in high-order cesarean deliveries. *Obstet Gynecol* 2015;126:21S–6.
 - 33 McGee J, Winick-Ng J, McClure JA, et al. Resident trainees increase surgical time: a comparison of obstetric and Gynaecologic procedures in academic versus community hospitals. *J Obstet Gynaecol Can* 2020;42:430–8.
 - 34 Thinggaard E. Take-home training in laparoscopy. *Dan Med J* 2017;64:B5335.
 - 35 Ericsson KA. Towards a science of the acquisition of expert performance in sports: Clarifying the differences between deliberate practice and other types of practice. *J Sports Sci* 2020;38:159–76.
 - 36 Leonard PSJ, Crouse DL, Boudreau JG, et al. Provider volume and maternal complications after caesarean section: results from a population-based study. *BMC Pregnancy Childbirth* 2020;20:37.
 - 37 Ericsson KA, Harwell KW. Deliberate practice and proposed limits on the effects of practice on the acquisition of expert performance: why the original definition matters and recommendations for future research. *Front Psychol* 2019;10:2396.
 - 38 Radhakrishnan G, Yadav G, Vaid N. Factors affecting “decision to delivery interval” in emergency caesarean sections in a tertiary care Hospital: a cross sectional observational study. *Int J Reprod Contracept Obstet Gynecol* 2013;2:651–6.
 - 39 Pearson GA, MacKenzie IZ. Factors that influence the incision-to-delivery interval at caesarean section and the impact on the neonate: a prospective cohort study. *Eur J Obstet Gynecol Reprod Biol* 2013;169:197–201.
 - 40 Fasubaa OB, Ezechi OC, Orji EO, et al. Delivery of the impacted head of the fetus at caesarean section after prolonged obstructed labour: a randomised comparative study of two methods. *J Obstet Gynaecol* 2002;22:375–8.
 - 41 Gedikbasi A, Akyol A, Asar E, et al. Re-laparotomy after cesarean section: operative complications in surgical delivery. *Arch Gynecol Obstet* 2008;278:419–25.
 - 42 Wong TC, Lau CQ, Tan EL, et al. Decision-to-delivery intervals and total duration of surgery for caesarean sections in a tertiary General Hospital. *Singapore Med J* 2017;58:332–7.
 - 43 Balachandrar K, Melov SJ, Nayyar R. The risk of adverse maternal outcomes in cases of placenta praevia in an Australian population between 2007 and 2017. *Aust N Z J Obstet Gynaecol* 2020;60:890–5.
 - 44 Lewandowska M, Sajdak S, Więckowska B, et al. The influence of maternal BMI on adverse pregnancy outcomes in older women. *Nutrients* 2020;12:2838.
 - 45 Doi L, Williams AJ, Marryat L, et al. Cohort study of high maternal body mass index and the risk of adverse pregnancy and delivery outcomes in Scotland. *BMJ Open* 2020;10:e026168.
 - 46 Mourad M, Silverstein M, Bender S, et al. The effect of maternal obesity on outcomes in patients undergoing tertiary or higher cesarean delivery. *J Matern Fetal Neonatal Med* 2015;28:989–93.
 - 47 Bini SA. Artificial intelligence, machine learning, deep learning, and cognitive computing: what do these terms mean and how will they impact health care? *J Arthroplasty* 2018;33:2358–61.