



OPEN

## Accurate prediction of birth implementing a statistical model through the determination of steroid hormones in saliva

Silvia Alonso<sup>1</sup>, Sara Cáceres<sup>1✉</sup>, Daniel Vélez<sup>2</sup>, Luis Sanz<sup>2</sup>, Gema Silvan<sup>1</sup>, Maria Jose Illera<sup>1</sup> & Juan Carlos Illera<sup>1</sup>

Steroidal hormone interaction in pregnancy is crucial for adequate fetal evolution and preparation for childbirth and extrauterine life. Estrone sulphate, estriol, progesterone and cortisol play important roles in the initiation of labour mechanism at the start of contractions and cervical effacement. However, their interaction remains uncertain. Although several studies regarding the hormonal mechanism of labour have been reported, the prediction of date of birth remains a challenge. In this study, we present for the first time machine learning algorithms for the prediction of whether spontaneous labour will occur from week 37 onwards. Estrone sulphate, estriol, progesterone and cortisol were analysed in saliva samples collected from 106 pregnant women since week 34 by enzyme-immunoassay (EIA) techniques. We compared a random forest model with a traditional logistic regression over a dataset constructed with the values observed of these measures. We observed that the results, evaluated in terms of accuracy and area under the curve (AUC) metrics, are sensibly better in the random forest model. For this reason, we consider that machine learning methods contribute in an important way to the obstetric practice.

In the obstetric field, one of the main objectives is the understanding of the specific physiology at the beginning of parturition and the hormonal interaction, as these concepts are crucial for the adaptation of the fetus to extrauterine life and labour.

The gestation period is considered at term between week 37–42<sup>1</sup>. The perinatal mortality rate increases from 2 to 3% in week 40 to 4–7% in week 42<sup>2</sup>, with post-term pregnancy a reason for the induction of labour<sup>3</sup>. Currently, there is no set time for induction, although it is recommended between week 41–42<sup>4</sup>. Only 4% of infants are born on the estimated date of delivery, which is calculated based on the date of the last menstruation plus 280 days, with an error of two weeks<sup>5</sup>. Childbirth involves the complex relationship between mother, fetus and placenta that implies a complex interaction of biomolecular, immunological and endocrine mechanisms, modulated by aetiology, ethnicity and gestational age<sup>6</sup>. It is a perfect coordination of events that include progressive effacement and dilation of the cervix, rupture of the amniotic membranes, and initiation and maintenance of effective uterine contractions, culminating in labour<sup>7</sup>.

The development of pregnancy is under hormonal control of the fetoplacental unit. Progestogens, estrogens, androgens and glucocorticoids are secreted during pregnancy and their interaction modulates different cellular and physiological mechanisms<sup>8</sup>.

Estrogen concentrations progressively increase during pregnancy<sup>9</sup> and are involved in the increase of the size of the uterus, the stimulation of the expression of oxytocin (OT) and progesterone (P4) receptors and the induction of synthesis of hepatic proteins. In late pregnancy, estrogens are involved in the activation of the labour mechanism<sup>8</sup>.

Estriol (E3) is the dominant estrogen during pregnancy. It is produced by the placenta from dehydroepiandrosterone sulphate (DHEAS) synthesised exclusively by the fetal adrenal gland, and is used as an indicator of the function of the fetoplacental unit<sup>10–12</sup>. Another important estrogen during pregnancy is estrone sulfate (E1SO4), which acts as a reserve for the peripheral formation of bioactive estrogenic forms<sup>13</sup>.

<sup>1</sup>Department of Physiology, School of Veterinary Medicine, University Complutense of Madrid, 28040 Madrid, Spain. <sup>2</sup>Department of Statistics and Operational Research, Faculty of Mathematics, University Complutense of Madrid, 28040 Madrid, Spain. ✉email: sacacere@ucm.es

In comparison, P4 is essential for the maintenance of pregnancy as it inhibits the contractility of the myometrium by decreasing the production of prostaglandins and the genetic expression of contraction associated proteins (CAPs)<sup>14</sup>.

Regarding glucocorticoids, cortisol (C) is of vital importance in pregnancy, since it is responsible for fetal lung maturation<sup>15</sup>. The increase in C levels at the end of pregnancy can serve as a signal for the fetus to induce estrogen synthesis prior to the onset of delivery<sup>16,17</sup>.

The endocrine system is a complex regulatory system with multiple levels of organisation (central nervous system, tissues, adrenal glands, cells) and time cycles (circadian oscillations, rapid responses, ultradian rhythms). This system displays nonlinear feedback responses involved in the interaction between its components and other parts of the body<sup>18</sup>. Statistical models have assisted in understanding how feedback mechanisms are key to homeostatic modulation<sup>18</sup>, and also the behaviour of the complex pituitary-adrenal-pituitary axis complex<sup>19</sup>.

Traditional logistic regression models are generally used to face a problem of predicting premature births<sup>20</sup>. However, the use of machine learning models has aroused growing interest and have been widely employed in bioinformatics<sup>21</sup>, as well as in different fields of health sciences<sup>22,23</sup>. The possibility of its use is also mentioned in the diagnosis of premature births and emerging obstetric diseases<sup>20</sup>. In this way, machine learning models, such as tree-based algorithms and neural networks, can be used to address problems related to the prediction of labour. In this study, random forest models were used to contrast the contribution of such a machine learning technique, compared to traditional logistic regression models.

Therefore, the aim of our study was to determine P4, C, E3 and E1SO4 levels by a non-invasive method, in saliva samples from healthy pregnant women during the third trimester of pregnancy (from week 27 to the time of delivery), in order to develop a mathematical model to predict the probability of a spontaneous birth.

## Results

**Validation parameters.** The sensitivity of the EIA technique was tested by means of low detection limit and calculated in 10 consecutive assays. Results were P4=0.25 nmol/L, C=0.14 nmol/L, E3=0.13 nmol/L and E1SO4=0.31 nmol/L. The accuracy of the EIA was tested by determining the recovery rates of known amounts of P4, C, E3 and E1SO4 spiked into saliva samples. The average range of recovery rates was as follows (mean  $\pm$  standard deviation): P4=94.29  $\pm$  3.85 to 98.71  $\pm$  5.37; C=94.29  $\pm$  3.85 to 97.15  $\pm$  2.68; E3=90.75  $\pm$  3.92 to 99.65  $\pm$  2.01; and E1SO4=90.08  $\pm$  3.29 to 93.27  $\pm$  0.90. Precision of P4, C, E3 and E1SO4 was determined by calculating the intra- and inter-assay coefficients of variation (CV%). P4=intra-assay: 3.2  $\pm$  0.98%, inter-assay: 5.3  $\pm$  0.89%; C=intra-assay: 4.8  $\pm$  1.02%, inter-assay: 7.4  $\pm$  1.23%; E3=intra-assay: 6.7  $\pm$  1.32%, inter-assay: 9.1  $\pm$  2.14%; and E1SO4=intra-assay: 2.9  $\pm$  0.84%, inter-assay: 6.5  $\pm$  1.36%. In order to determine the effects of saliva on standard curve, the standard curves with saliva samples were run in parallel with the standard dose-response curve. There was a good degree of parallelism between both standard curves for the hormones studied: P4, R=0.87; C, R=0.84; E3, R=0.89; and E1SO4, R=0.83.

**Hormonal concentrations.** Total means of each hormone together with a confidence interval based on one standard deviation were calculated from week 26–41 (Fig. 1 and supplementary Table 1). In addition, the evolution of mean values distinguished by the week of delivery is also shown. Among all the hormones analysed, P4 is the only one that increased gradually throughout the weeks of pregnancy during the third trimester (Fig. 1A). C and E3 results showed that concentrations increased significantly ( $p < 0.05$ ) from week 37 until giving birth (Fig. 1B,C). E1SO4 concentrations started increased exponentially from the 35<sup>th</sup> week until giving birth (Fig. 1D).

Regarding repeated measures correlation coefficient (rmcorr) between hormones, results demonstrated that all hormones are positively correlated (Table 1), with C as the one presenting the strongest correlation with the rest of the hormones while P4 as the one with the weakest correlation.

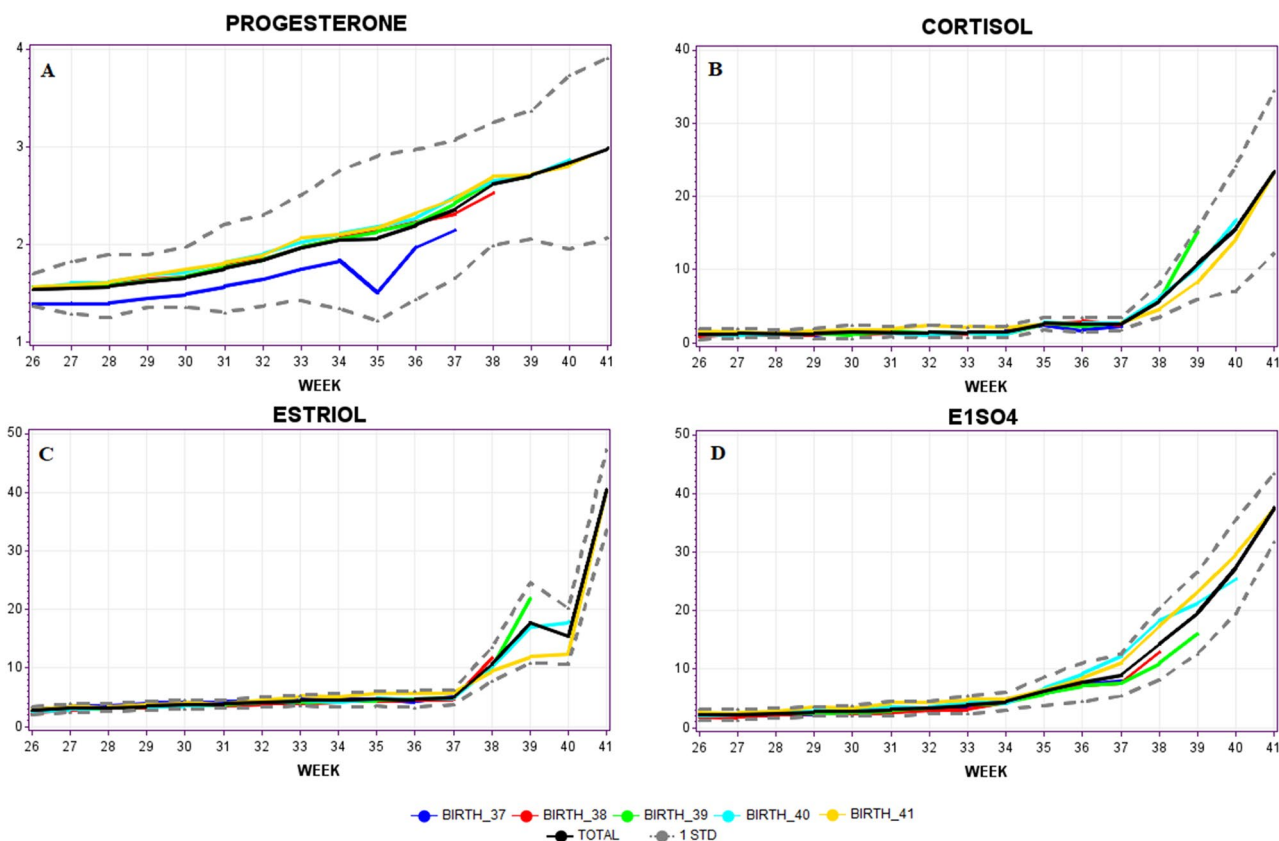
**Statistical model.** We have adjusted the statistical models in order to predict the probability of spontaneous birth in the following week, from gestation week 37 (37<sup>th</sup> week included), based on the hormonal levels collected since the 34<sup>th</sup> week (Fig. 1). It was decided to adjust two types of statistical models. On the one hand, we fit a multivariate logistic regression model, since it is a reference model in the biosanitary field, mainly due to its simple interpretation. Such a model allows the measurement of the influence on the response variable of the independent variables by means of the odds ratio associated with each of the estimated parameters. On the other hand, a random forest model has been adjusted, which is a model that loses the interpretability of the results, but in return, is very competitive from the predictive point of view<sup>25</sup>.

Both models have been implemented using SAS Institute Enterprise Miner software, which internally invokes the DMREG procedure to perform the logistic regression and HPFOREST procedure to perform the random forest, proposing accuracy as the metric to be maximised.

In the logistic regression model, the logit function has been considered as the linkage function. Also, a step-wise variable selection process has been used, proposing a significance value of 0.05 for the sl-entry and sl-stay parameter. To avoid possible overfitting effects in the adjustment of the models, we have applied a cross-validation strategy. The value obtained for the accuracy metric in this process have been 77.4%. The assessment over the test dataset has provided an accuracy of 69.07% (Table 2).

Table 3 shows the order in which the variables have been included in the model. No variables were removed in the adjustment process.

Table 4 shows the estimates of the model coefficients, the corresponding p-values and the odds ratios associated with variables selected.



**Figure 1.** Progesterone (A), cortisol (B), estriol (C) and estrone sulphate (D) mean concentrations during third trimester of pregnancy (from 34th week until delivery). Values corresponded to each group of women, distinguished by the week of delivery.

From 26–37 weeks		
v1	v2	$\rho_{v1v2}$
CORTISOL	ESTRIOL	0.374
CORTISOL	PROGEST	0.542
CORTISOL	E1SO4	0.408
ESTRIOL	PROGEST	0.251
ESTRIOL	E1SO4	0.365
PROGESTE	E1SO4	0.314

**Table 1.** Repeated measures correlation (rmcorr) of P4, C, E3, and E1SO4 concentrations in third trimester of gestation.

We can observe that the positive and negative signs of the coefficients associated with the hormonal variables are consistent with the biological functions. We think that negative values for E1SO4 and progesterone are related to longer pregnancies, while positive signs for estriol and cortisol are related to earlier deliveries.

In the random forest model, a maximum of 100 decision trees were considered. Regarding the value of the parameters associated to the trees, following are those considered by default in Enterprise Miner:

- Maximum depth = 50
- Number of training observations a node must have to consider splitting it = 30
- Minimum number of observations per leaf = 1
- Percent of the training dataset randomly selected to adjust each one of the trees = 60%. Seed used for generating random numbers = 12,345
- Number of variables by tree =  $\text{int}(\sqrt{\text{total number of variables}}) = \text{int}(\sqrt{22}) = 4$ .

Logistic regression				Random forest			
Actual	0	1	Total	Actual	0	1	Total
<b>Predicted</b>							
0	TN = 52	FN = 15	67	0	60	13	73
1	FP = 15	TP = 15	30	1	7	17	24
TOTAL	67	30	97	TOTAL	67	30	97
Accuracy	69.07%			Accuracy	79.38%		
PPV	50.00%			PPV	70.83%		
NPV	77.61%			NPV	82.19%		
AUC	0.707			AUC	0.819		

**Table 2.** Confusion matrices of logistic regression and random forest models over test dataset. TN = number of true negative cases, FN = number of false negative cases, FP = number of false positive cases, TP = number of true positive cases, PPV = positive predictive value (precision), NPV = negative predictive value, AUC = area under the ROC. Being:  $Accuracy = \frac{TN+TP}{TN+FN+FP+TP}$ ,  $PPV = PositivePredictiveValue(Precision) = \frac{TP}{TP+FP}$ ,  $NPV = NegativePredictiveValue = \frac{TN}{TN+FN}$

Step	Sense	Variable	Cross-validation accuracy (%)
1	Entry	ESTRIOL_VAR_WEEK	76.41
2	Entry	CORTISOL_FORTNIGHT	74.87
3	Entry	EISO4_VAR_FORTNIGHT	76.92
4	Entra	PRIMIPARA	73.33
5	Entry	AGE	76.92
6	Entry	PROGESTERONA_MONTH	77.44

**Table 3.** Evolution of training success rate and cross validation of stepwise logistic regression.

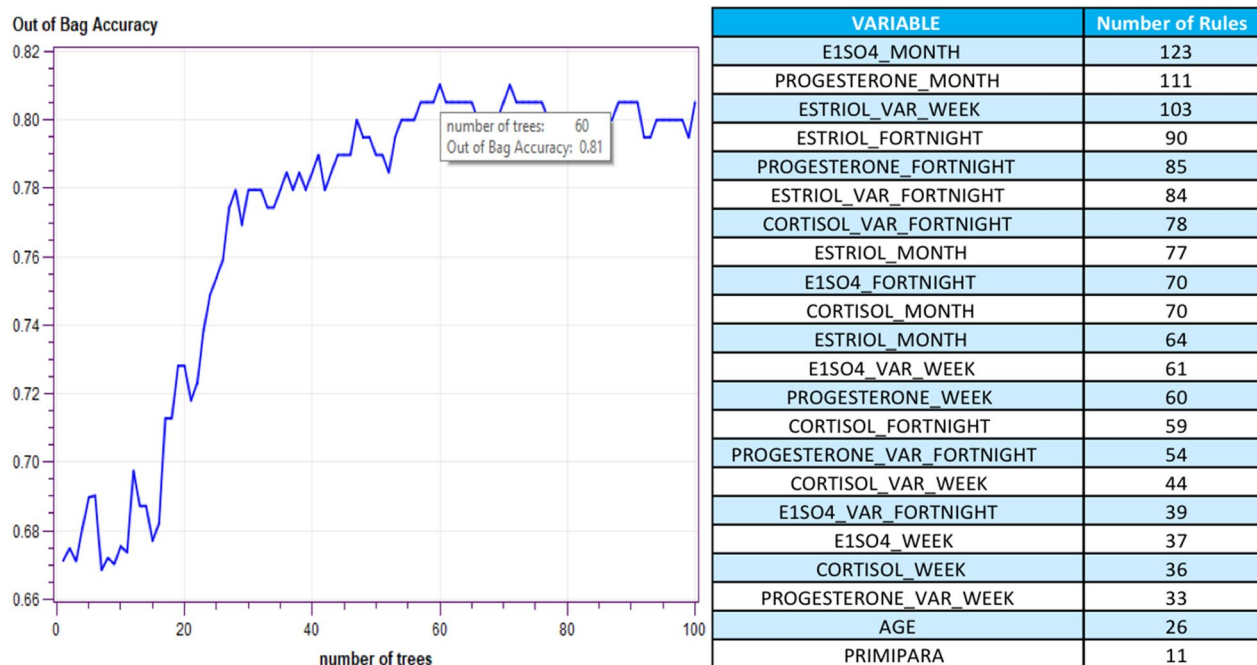
Variable	Estimation	SE	p-value	Odds ratio
INTERCEPT	7.1399	2.8555	0.0124	
ESTRIOL_VAR_WEEK	8.9654	1.9565	<0.0001	999
CORTISOL_FORTNIGHT	0.1987	0.0549	0.0003	1.22
EISO4_VAR_FORTNIGHT	- 0.9205	0.4175	0.0275	0.398
PRIMIPARA (NO)	0.6204	0.2039	0.0023	NO VS YES (3.458)
AGE	- 0.2045	0.0748	0.0062	0.815
PROGESTERONA_MONTH	- 0.916	0.4646	0.0486	0.4

**Table 4.** Estimation, p-values and odds ratios associated with variables of logistic regression model.

In this case, to avoid possible overfitting, we used a modelling strategy used by Enterprise Miner called Out-Of-Bag (OOB) because it is the alternative this software proposed to the well-known cross-validation strategy. In the OOB strategy, the observations associated with each one of the adjusted trees are called the bagged observations, while those excluded are called OOB observations. These last observations were used in order to decide the number of trees to be ensemble to improve the capacity of generalisation of the model, similar to a cross-validation strategy.

The evaluation of the accuracy metric leads us to consider as an optimal model the ones that ensemble a total of 60 or 71 trees (Fig. 2), showing an OOB accuracy of 81.03%. According to the principle of parsimony (Occam's razor) recommended in the context of modelling<sup>26</sup>, we decided to use the simplest model, therefore, the one covering 60 trees was taken. The assessment over the test dataset provided an accuracy of 79.38% (see Table 2).

A list representing the relative importance of the variables in this model is also shown in Fig. 2. This relevance is established based on the degree of participation of these variables in each of the 60 trees that comprise the model. The fact that the same variable can participate in the same tree several times justifies that the participation of each variable can be greater than 60. Therefore, it can be seen that the EISO4 value average over the four weeks of the last month (EISO4\_MONTH variable) is the most important effect to be considered. The same average values of Progesterone-PROGESTERONE\_MONTH and the percentage variation of Estriol- ESTRIOL\_VAR\_WEEK over the last two weeks were, respectively, the second and third most relevant variables, being the



**Figure 2.** Evolution of success rate in training and Out-Of-Bag samples with random forest model.

only ones that participate in more than 100 trees. The first variable associated with cortisol hormone appeared in the seventh position.

In Table 2, the confusion matrices associated with both models are presented. The accuracy, the positive and the negative predictive value (PPV and NPV, respectively) showed that the random forest model provides better results than the logistic model. The area under the Receiver Operating Characteristics (ROC) curve, one of the most important evaluation metrics to verify the performance of any classification model<sup>27</sup>, showed that the random forest model works better also.

Moreover, two criteria were used as a baseline to compare the accuracy of the models. Firstly, the maximum randomness criterion, consisted of assigning the dominant class (birth = 0) to all cases. Although we considered this criterion to be very conservative, it was observed that the reference value it provides is 67/97 (69.07%), which is equal to the accuracy of the logistic regression, but significantly less than that of the random forest model (79.38%). Secondly, the proportional randomness criterion consisted of randomly assigning 0 and 1 values according to the actual proportion of the binary event analysed. Assuming that the predictions are made at random regardless of the real values, the expected probability of coincidences using this criterion is 0.5727 (57.27%).

$$P(\text{actual} = 1 \cap \text{predicted} = 1) + P(\text{actual} = 0 \cap \text{predicted} = 0) = p^2 + (1 - p)^2 = \left(\frac{67}{97}\right)^2 + \left(\frac{30}{97}\right)^2 = 0,5727$$

being  $p = P(\text{actualBirth} = 1)$  and  $1 - p = P(\text{actualBirth} = 0)$ .

The second criterion is considered reasonable to test the true goodness of fit of the model. Then, in the case of the multivariate logistic regression model, the reference precision improves by 11.80 percentage points (69.07–57.27%), while with the random forest model the improvement is 22.11 percentage points (79.38–57.27%) (Table 5).

These results led us to propose the random forest model as appropriate to predict the week in which birth is most likely to occur, although it has less explanatory power than the logistic regression model.

Finally, the confusion matrices, detailed per week, are presented in Table 5. It was observed that the accuracy of the model is higher in the 37th week (85.71%), but this result must be interpreted carefully because in this week, the model tends to predict birth = 0, taking into account that it is the week in which fewer births occur. It should be considered that this prediction was provided by the model in 33 of the 35 cases, and hence in almost 94.30% of the cases. This allowed the highest NPV value (87.88%) but, this should not be considered as an important result because the maximum randomness criterion would provide a value of 85.71% (30/35). On the other hand, the number of positive cases predicted in each week was similar to the actual cases in the weeks numbered 38, 39, 40 (eight vs. 10, six vs. eight, eight vs. seven, respectively) and therefore, the corresponding results could be considered more interesting.

## Discussion

In this study, the physiological variations of P4, C, E3 and E1SO4 during the third trimester are shown. These variations played an important role in pregnancy and in the preparation of labour mechanisms. It is convenient to explain that the activation of labour was not directly related to changes in the levels of estrogen and progesterone; however, this event could be affected by an alteration in the response of the myometrium to the expression of

WEEK	37			38			39			40		
Actual	0	1	OVERALL	0	1	OVERALL	0	1	OVERALL	0	1	OVERALL
<b>Predicted</b>												
0	29	4	33	17	5	22	11	3	14	3	1	4
1	1	1	2	3	5	8	1	5	6	2	6	8
OVERALL	30	5	35	20	10	30	12	8	20	5	7	12
Accuracy	85.71%			73.33%			80.00%			75.00%		
PPV	50.00%			62.50%			83.33%			75.00%		
NPV	87.88%			77.27%			78.57%			75.00%		

**Table 5.** Confusion matrices evaluated per week on test table by random forest model. TN = number of true negative cases, FN = number of false negative cases, FP = number of false positive cases, TP = number of true positive cases, PPV = positive predictive value (precision), NPV = negative predictive value, AUC = area under the ROC.

estrogen receptors (ER) and progesterone receptors (PR). The interaction between PR and ER of the myometrium prevents the action of estrogens during most of the pregnancy, while in preparation for delivery, there is a withdrawal of progesterone and estrogen activation<sup>28</sup>.

We implemented the analysis of these hormones in saliva samples, a useful non-invasive method to predict the week of spontaneous delivery. Our results revealed a positive hormonal correlation between all studied hormones, with C being the strongest correlation to the rest of hormones. Proper functioning and interaction of steroidogenic tissues, in addition to the enzymes involved in their synthesis and transport, are crucial since intrauterine exposure of the fetus to abnormal glucocorticoid levels<sup>29</sup> or sex hormones<sup>30,31</sup> can negatively affect fetal development. Some authors state that there is a relationship between fetal sex with estrone and estradiol levels during pregnancy, however, they did not find statistically significant results<sup>12</sup>.

The role of the different sex hormones has been extensively studied, although there are hardly any articles that interrelate these levels of estrogens, progestogens and glucocorticoids and their variations throughout the third trimester of pregnancy. According to the results obtained in this study, the hormonal analysis of P4, E3, C and E1SO4 at specific times of the third trimester were good indicators of normal values in low-risk pregnancies, and their determination could provide additional information relevant to the control of pregnancy.

In addition, the exponential increase in P4, C, E3 and E1SO4 in third trimester, the period in which the fetus reaches maturity<sup>32</sup>, prepares the fetus and mother for the start of labour, as reported by other authors<sup>8,33–35</sup>. Although its biological importance is ambiguous, E1SO4 seems to serve as an estrogenic reservoir linked to the endometrium transformation of the endometrium necessary for pregnancy maintenance<sup>36</sup>. The increase in E1SO4 at week 35 determines the increases in the other hormones studied. Therefore, it is critical to take this measurement last during the third trimester of pregnancy, in order to obtain a better understanding of the onset of labour processes.

The relevance of P4, C, E3 and E1SO4 in pregnancy and the onset of labour led us to develop a statistical model that could predict the week of birth. In recent years, the need to develop new prognostic methods in the field of obstetrics in order to improve clinical practice regarding the prediction of premature births, delivery methods, and development of pre-eclampsia, among others, has been emphasised<sup>37</sup>.

The knowledge of the week of spontaneous birth one week in advance has multiple benefits for both the pregnant woman and the health personnel, and has been very useful for reducing the number of inductions to labour for post term pregnancies. The development of a mathematical model with a positive predictive value of 70.83%, as well as a negative predictive value of 82.19%, was capable to relate from a woman the variations of the different hormones of only four saliva samples collected from week 34 and to predict childbirth. Therefore, this model is a novel and great advance in the field of obstetrics.

The application of the statistical model developed in low-risk pregnant women, routinely from week 34, contributed to improving the allocation of health resources based on the number of expected deliveries, as an increase in the number of midwives and obstetricians is associated with a decrease in the cesarean section ratio<sup>38</sup>. Also, its routine application contributed to reducing the number of inductions of labour (IOL) for post-term pregnancy, confirming that labour will begin the week following the application of the statistical model. This action could specify the maximum period to maintain a pregnancy, giving the opportunity of a spontaneous delivery and reducing the complications associated with IOL as postpartum haemorrhage, uterine rupture<sup>39</sup>, instrumental deliveries, cesarean sections and used of epidural anaesthesia<sup>40,41</sup>, among others.

For future research, among other aspects, the inclusion of criteria such as pre-pregnancy body mass index, previous preterm birth or history of progesterone medication, would enhance the results of future randomised forest models<sup>45</sup>.

Therefore, due to the development of a statistical model that predicts the due date in low-risk women based on the levels of P4, SO4E1, E3 and C from week 34, fetal survival, the allocation of health resources based on the number of expected delivery, as well as the reduction of the number of inductions to post-term delivery, could be markedly improved.

	All participants recruited
	n = 161
	Participants completed study n = 106 (Dropout rate = 34.2% )
Mean age	34.88 ± 3.29
<b>Prepregnancy BMI (n, %)</b>	
Normal weight	97 (91.5%)
Underweight	9 (8.4%)
<b>Parity (n, %)</b>	
Primiparous	66 (62.2%)
Multiparous	40 (37.7%)
<b>Current pregnancy</b>	
Mean gestational age	38 weeks + 6 days
Birth at 37 weeks	17 (16%)
Birth at 38 weeks	30 (28.3%)
Birth at 39 weeks	24 (22.6%)
Birth at 40 weeks	21 (19.8%)
Birth at 41 weeks	16 (15%)
<b>Mode of delivery</b>	
Spontaneous birth	63 (59.4%)
Induction of labour (post dates + Spontaneous ruptured of membranes* + other reasons)	43 (40.5%)
Induction of labour for post dates (> 41 + 4)	6 (6.3%)
Vaginal birth	83 (78.3%)
Cesarean delivery- without trial of labour**	9 (8.4%)
Cesarean delivery- after trial of labour	14 (13.2%)
<b>Rupture of membranes</b>	
> 24hours	31 (29.2%)
< 24hours	75 (70.8%)
<b>Newborn</b>	
Mean birth weight (g)	3479 ± 356gr
<b>Apgar score &lt; 7</b>	
One minute	8 (7.5%)
Five minutes	4 (3.7%)
10 min	2 (1.8%)

**Table 6.** Data for all recruited women. All indicators were applied according to the SEGO recommendations. \*Spontaneous rupture of membranes was considered spontaneous birth as between 60–80% of births occur within 24 h after SROM and approximately 90% occur by 48 h, however, due to risk of neonatal infection, all study participants were induced 12 h post SROM, as per hospital protocol. \*\*Cesarean delivery without trial of labour was considered as those performed during latent phase of labour without reaching active labour. Elective cesarean section cases were discarded from the total samples.

## Methods

**Sample collection.** This study was carried out at the Nuevo Belén Clinical University Hospital (Madrid, Spain), in collaboration with the Department of Animal Physiology of the Complutense University of Madrid (Spain). The process of patient recruitment and sample collection was carried out for over a year. An informed consent was obtained from all participants recruited on this study.

A total of 161 healthy pregnant women aged between 27 and 44 years (34.88 ± 3.29) were recruited (data from all recruited woman are summarised in Table 6), of which two of the initial participants dropped out before the beginning of the study. We therefore commenced with a total of 159 women, 53 of whom dropped out throughout the study, reaching a total of 106 who completed the study. The exclusion criteria were kidney disease, thyroid disease, autoimmune disease, cancer, pregestational diabetes, pre-gestational hypertension, in vitro fertilisation with heparin treatment, and current steroid treatment.

Eligible women underwent weekly collection of a saliva sample during the third trimester of pregnancy (from week 27 until week of delivery), with a Salivette (Sarstedt, Germany) collection tube. All samples were collected at an established time (10:00 am ± 1 h).

This study was approved by the Clinical Research Ethics Committee of Hospitales de Madrid, with the code CEIm HM hospitals: 16.06.0960-GHM, in accordance with the World Medical Association and the Declaration of Helsinki.

WEEK	37	38	39	40	OVERALL
<b>Birth</b>					
0	90	60	36	16	202
1	16	30	24	20	90
Overall	106	90	60	36	292

**Table 7.** Binary variable for women and week that identifies whether the woman gives birth (1) or not (0).

**Hormone determinations.** The Salivette tubes were centrifuged for 15 min at 2000×g and 4° C. The saliva obtained was stored at a temperature of −20 °C until further hormonal analysis. Polyclonal antisera raised in male rabbits against: 11α-hydroxyprogesterone 11α hemisuccinate BSA (P4), estrone-3-glucuronide: BSA (SO4E1), cortisol 3-CMO: BSA (cortisol) and estriol -6-CMO: BSA (Estriol) were used for assay development (Steraloids. Inc., Newport, RI). P4 (ab: C914), E1SO4 (ab: R522-2), E3 (ab: R4835) and C (ab: R4866) concentrations were assayed by enzyme-immunoassay (EIA).

Briefly, microtiter plates (96-well flat-bottom polystyrene) were coated overnight at 4° C with the appropriate antibody dilutions. Then, plates were washed 3 times and conjugate working solutions (CWS) were prepared by diluting conjugate stocks in assay buffer. Standards and saliva samples were diluted in CWS and analyzed in duplicate. To achieve a competitive reaction, plates were incubated at room temperature for 2 h. Plates were washed 3 times with wash buffer and Enhance K-Blue TMB substrate (Neogen, Lexington, KY) was added to each well. To stop colorimetric reaction, stop solution was added to each well. Absorbance was read at 450 nm in an automatic microplate reader. Hormone concentrations were calculated by means of software developed for this technique (ELISA AID, Eurogenetics, Belgium). A standard dose–response curve was constructed by plotting the binding percent-age ( $B/B_0 \times 100$ ) against steroid hormone standard concentrations. P4, E3, SO4E1 and C concentrations were expressed in ng/ml<sup>42</sup>.

The validation technique parameters' recovery rates, sensitivity, intra- and inter-assay coefficients of variation and parallelism were assayed as previously reported by Illera et al.<sup>42</sup>. The EIA techniques and antibodies used were developed and validated in the Endocrinology Laboratory of the Department of Animal Physiology (Faculty of Veterinary Medicine, University Complutense of Madrid, Spain).

**Statistical analysis.** In order to achieve the desired precision for the estimates of the population mean hormonal values in each week of observation, the sample size was estimated with a confidence level of 95%, and an error of the estimate of ±0.2 times the standard deviation, according to this consideration we obtained  $n \geq 96.04$ .

$$\text{Error of the estimate} = z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

$$z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq 0.2\sigma$$

$$\alpha = 0.05$$

$$n \geq \left( \frac{1.96}{0.2} \right)^2 = 96.04$$

To achieve the required precision and based on previous experience with other studies, an approximate loss between 30 and 40% of women has been estimated throughout the sample collection process (in the worst case, the dropout rate was 40%, leading us to specify  $n \geq 161$ ). Finally, after concluding the study,  $n = 106$ , which is above the initial consideration resulting in a dropout rate 34.2%.

Descriptive statistical analysis for mean and standard deviation of each hormone based on the week of delivery was performed. Repeated measures correlation coefficient (rmcorr) between hormonal values during the third trimester were estimated using PROC MIXED procedures by SAS<sup>24</sup>.

Validation technique parameters (recovery rates, sensitivity, intra- and inter-assay coefficients of variation) were calculated according to Andreasson et al.<sup>43</sup>. Parallelism was calculated using ANCOVA analysis<sup>44</sup>. Data were expressed as mean ± standard error. In all statistical comparisons, type I error set to maximum of 0.05.

**Methodology for the construction of a childbirth prediction model.** The methodology employed for predicting the probability of spontaneous delivery in the week following from the latest hormonal levels observed is presented in this section. This methodology is based on two phases: processing data and adjustment of mathematical model.

**Data processing.** A binary variable has been defined for each pair (Woman, Week), identifying whether a subject gives birth (1) or not (0) each week. This is the variable whose value is to be predicted, and its distribution by weeks is shown in Table 7. This table presents a total of 292 cases, distributed into 202 cases with a target equal to 0 and 90 cases with a target equal to 1. This table is divided into two tables (Table 8), one named “training dataset used to adjust the model” and another one named “test dataset used to check it”. The proportions associated with one and another dataset are 2/3 and 1/3, respectively. The cases assigned to each dataset were selected at random, taking into account that the patients registered in one and the other table must be exclusive.



WEEK	Training dataset					Test dataset				
	37	38	39	40	OVERALL	37	38	39	40	OVERALL
<b>Birth</b>										
0	60	40	24	11	135	30	20	12	5	67
1	11	20	16	13	60	5	10	8	7	30
Overall	71	60	40	24	195	35	30	20	12	97

**Table 8.** Learning dataset and test dataset.

INDICATOR_WEEK	Reflects the level of the indicator in the week before WE <i>For example: CORTISOL_WEEK measures the value of C in the week before WE</i>
INDICATOR_FORTNIGHT	Reflects the average level of the indicator in the fortnight before WE (the average of the indicator in the last two weeks) <i>For example: ESTRIOL_FORTNIGHT measures the average of the two E3 values recorded in the two weeks prior to WE</i>
INDICATOR_MONTH	Reflects the average level of the indicator in the month before WE (the average of the indicator in the last four weeks) <i>For example: PROGESTERONE_MONTH measures the average of the four P4 values recorded in the four weeks prior to WE</i>
INDICATOR_VAR_WEEK	Reflects the percentage variation of the indicator between the two weeks prior to WE $\text{INDICATOR\_VAR\_WEEK} = \frac{\text{INDICATOR\_WEEK\_1} - \text{INDICATOR\_WEEK\_2}}{\text{INDICATOR\_WEEK\_2}}$
INDICATOR_VAR_FORTNIGHT	Reflects the percentage variation of the indicator between the last two fortnights before WE $\text{INDICATOR\_VAR\_FORTNIGHT} = \frac{\text{INDICATOR\_FORTNIGHT\_1} - \text{INDICATOR\_FORTNIGHT\_2}}{\text{INDICATOR\_FORTNIGHT\_2}}$
AGE	Continuous variable that refers to the age of the woman
PRIMIPARA	Binary variable that identifies whether it is the first time (value 1) or not (value 0) that the woman give birth

**Table 9.** Dictionary of variables. WE = week in which the target is evaluated, whether or not the woman gave birth.

Once the target (variable to be predicted) was defined, it was necessary to define the explanatory inputs used to make the prediction. Basically, these variables measured mean values and percentage variations of the hormonal indicator during different time periods. The detail of the definition of these variables is shown in Table 9. Considering that the data was collected on Mondays of each week, the explanatory variables were measured in the four weeks prior to which the probability of delivery was estimated, including the forecast week itself. Therefore, the hormonal data collected at weeks 34, 35, 36, and 37 were used to predict whether the birth would occur at week 37, as well as weeks 38 and 39, and finally the data collected at weeks 37, 38, 39 and 40 were used to predict whether the birth would occur at week 40 or week 41. The resulting table, made up of the pair (Woman, Week), the binary target variable (birth or non-birth) and the 22 explanatory variables, served as input for different statistical models.

Received: 17 June 2020; Accepted: 23 February 2021

Published online: 10 March 2021

## References

- Usandizaga, J. A. & Escalante, J. M. Fisiología del embarazo. In *Obstetricia y ginecología* (eds Usandizaga, J. A. et al.) 96–129 (Marban, 2011).
- Galal, M., Symonds, I., Murray, H., Petraglia, F. & Smith, R. Postterm pregnancy. *FVV on ObGyn* **4**(3), 175–187 (2012).
- López Araque, A. B., López, M. D. & Linares, M. Emotional state of primigravid women with pregnancy susceptible to prolongation. *Invest. Educ. Enferm.* **33**(1), 92–101 (2015).
- National Institute for Health and Care Excellence (NICE), *Inducing labour: NICE guideline [CG70]*. (2018)
- Mongelli, M., Wilcox, M. & Gardosi, J. Estimating the date of confinement: ultrasonographic biometry versus certain menstrual dates. *Am. J. Obstet. Gynecol.* **174**, 278–281 (1996).
- Keelan, J. Intrauterine inflammatory activation, functional progesterone withdrawal, and the timing of term and preterm birth. *J. Reprod. Immunol.* **125**, 89–99 (2018).
- Menon, R. Human fetal membranes at term: dead tissue or signalers of parturition?. *Placenta* **44**, 1–5 (2016).
- Morel, Y. et al. Evolution of steroids during pregnancy: maternal, placental and fetal synthesis. *Annales d'Endocrinologie* **77**, 82–89 (2016).
- Pasqualini, J. R. Enzymes involved in the formation and transformation of steroid hormones in the fetal and placental compartments. *J. Steroid. Biochem. Mol.* **97**, 401–415 (2005).
- Kim, S. Y., Kim, S. K., Lee, J. S., Kim, I. K. & Lee, K. The prediction of adverse pregnancy outcome using low unconjugated estriol in the second trimester of pregnancy without risk of Down's syndrome. *Yonsei Med. J.* **41**, 226–229 (2000).
- Falah, N., Torday, J., Quinney, S. K. & Hass, D. M. Estriol review clinical application and potential biomedical importance. *Clin. Res. Trials* **1**(2), 29–33 (2015).

12. Kuijper, E. A. M., Ket, J. C. F., Caanen, M. R. & Lambalk, C. B. Reproductive hormone concentrations in pregnancy and neonates: a systematic review. *Reprod. Biomed. Online* **27**, 33–63 (2013).
13. Escobar, J. C., Patel, S. S., Vesahy, V. E., Suzuki, T. & Carr, B. B. The human placenta expresses CYP17 and generates androgens de novo. *J. Clin. Endocrinol. Metab.* **96**, 1385–1392 (2011).
14. Pařízek, A., Koucký, M. & Dušková, M. Progesterone, inflammation and preterm labor. *J. Steroid Biochem. Mol. Biol.* **139**, 159–165 (2014).
15. Vrachnis, N., Malamas, F.M., Sifakis, S., Tsikouras, P., Iliodromiti Z., Immune aspects and myometrial actions of progesterone and CRH in labour. *Clin. Dev. Immunol.* 937618 (2012)
16. Li, X. Q., Zhu, P., Myatt, L. & Sun, K. Roles of glucocorticoids in human parturition: a controversial fact?. *Placenta* **35**, 291–296 (2014).
17. Mendelson, C. R., Montalbano, A. P. & Gao, L. Fetal- to maternal signaling in the timing of birth. *J. Steroid. Biochem. Mol. Biol.* **179**, 19–27 (2017).
18. Zabala, E. *et al.* Mathematical modelling of endocrine systems. *Trends Endocrinol. Metab.* **30**(4), 244–257 (2019).
19. Leng, G. & Macgregor, D. J. Mathematical modelling in neuroendocrinology. *J Neuroendocrinol* **20**, 713–718 (2008).
20. Meertens, L. J. E. *et al.* Prediction models for the risk of spontaneous preterm birth based on maternal characteristics: a systematic review and independent external validation. *Acta Obstet. Gynecol. Scand.* **97**, 907–992 (2018).
21. Baçtanlar, Y. & Ozuysal, M. Introduction to machine learning. *Methods Mol Biol* **1107**, 105–128 (2014).
22. Goldstein, B. A., Navar, A. M. & Carter, R. E. Moving beyond regression techniques in cardiovascular risk prediction: applying machine learning to address analytic challenges. *Eur Heart J.* **38**, 1805–1814 (2017).
23. Peissig, P. L. *et al.* Relational machine learning for electronic health record-driven phenotyping. *J. Biomed. Inform.* **52**, 260–270 (2014).
24. Hamlett, A., Hamlett, R., Russ, L., Russ, W., On the use of PROC MIXED to estimate correlation in the presence of repeated measures, SAS Users Group International, Proceedings of the Statistics and Data Analysis Section, 1–7 (2004).
25. Malley, J. D., Malley, K. G. & Pajevic, S. *Statistical learning for biomedical data* (Cambridge University Press, 2011).
26. Sorrel, M. A., Abad, F. J., Olea, J., de la Torre, J. & Barrada, J. R. Inferential item-fit evaluation in cognitive diagnosis modeling. *Appl. Psychol. Meas.* **41**(8), 614–631. <https://doi.org/10.1177/0146621617707510> (2017).
27. Zou, K., O'Malley, J. & Mauri, L. Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation* **115**, 654–657. <https://doi.org/10.1161/CIRCULATIONAHA.105.594929> (2007).
28. Mesiano, S. *et al.* Progesterone withdrawal and estrogen activation in human parturition are coordinated by progesterone receptor A expression in the myometrium. *J. Clin. Endocrinol. Metab.* **87**(6), 2924–2930 (2002).
29. Miranda, A. & Sousa, N. Maternal hormonal milieu influence on fetal brain development. *Brain Behav.* **8**, e00920 (2018).
30. Cardoso, R. C., Puttabyatappa, M. & Padmanabhan, V. Steroidogenic versus metabolic programming of reproductive neuroendocrine, ovarian and metabolic dysfunctions. *Neuroendocrinology* **102**, 226–237 (2015).
31. Pluchino, N., Russo, M. & Genazzani, A. R. The fetal brain: role of progesterone and allopregnanolone. *Horm. Mol. Biol. Clin. Investig* **27**, 29–34 (2016).
32. Cunningham, F. *et al.* *Williams obstetricia* 23rd edn. (Mexico D.F, 2011).
33. Kota, S. K. *et al.* Endocrinology of parturition. *Indian J. Endocrinol. Metab* **17**(1), 50–59 (2013).
34. Soma-Pillay, P., Nelson-Piercy, C., Tolppanen, H. & Mebazza, A. Physiological changes in pregnancy. *Cardiovasc. J. Afr.* **27**(2), 89–94 (2016).
35. Chatuphonpraster, W., Jarukamjorn, K. & Ellinger, I. Physiology and pathophysiology of steroid biosynthesis, transport and metabolism in the human placenta. *Front Pharmacol.* **9**, 1027 (2018).
36. Gibson, D. A. *et al.* Sulfation pathways: a role for steroid sulphatase in intracrine regulation of endometrial decidualisation. *J. Mol. Endocrinol.* **61**(2), M57–M65. <https://doi.org/10.1530/JME-18-0037> (2018).
37. Kleinrouweler, C. E. *et al.* Prognostic models in obstetrics: available, but far from applicable. *Am. J. Obstet. Gynecol.* **214**(1), 79–90. e36. <https://doi.org/10.1016/j.ajog.2015.06.013> (2016).
38. Zbir, S., Rozenberg, P., Goffinet, F. & Milcent, C. Cesarean delivery rate and staffing levels of the maternity unit. *PLoS ONE* **13**(11), e0207379 (2018).
39. Ashwal, E. *et al.* The incidence and risk factors for retained placenta after vaginal delivery - a single center experience. *J. Matern. Fetal Neonatal Med.* **27**(18), 1897–1900 (2014).
40. Jay, A., Thomas, H. & Brooks, F. In labor or in limbo? The experiences of women undergoing induction of labor in hospital: Findings of a qualitative study. *Birth* **45**, 64–70 (2018).
41. Rydahl, E., Eriksen, L. & Juhl, M. Effects of induction of labour prior to post-term in low-risk pregnancies: a systematic review. *JBI Database Syst. Rev. Implement Rep* **17**(2), 170–208 (2019).
42. Illera, J. C. *et al.* Assessment of ovarian cycles in the African elephant (*Loxodonta africana*) by measurement of salivary progesterone metabolites. *Zoo Biol.* **33**(3), 245–249. <https://doi.org/10.1002/zoo.21124> (2014).
43. Andreasson, U. *et al.* A practical guide to immunoassay method validation. *Front. Neurol.* **6**, 179 (2015).
44. Carlsson, M. O., Zou, K. H., Yu, C. R., Liu, K. & Sun, F. W. A comparison of nonparametric and parametric methods to adjust for baseline measures. *Contemp. Clin. Trials.* **37**, 225–233 (2014).
45. Lee, K. S. & Ahn, K. H. Application of artificial intelligence in early diagnosis of spontaneous preterm labor and birth. *Diagnostics* **10**, 733 (2020).

## Acknowledgements

This research was supported by the Spanish Ministry of Science, Innovation and Universities, under the I+D+i Programme, grant number PID2019-106254RB-I00.

## Author contributions

S.A.: design of study, acquisition, analysis and interpretation of data, article drafting and revising; S.C.: design of study, acquisition, analysis and interpretation of data, article drafting and revising; D.V.: analysis and interpretation of data, development of mathematical model, article drafting and revising; L.S.: analysis and interpretation of data, development of mathematical model, article drafting and revising; G.S.: design of study, acquisition, analysis and interpretation of data, article drafting and revising; M.J.I.: design of study, acquisition, analysis and interpretation of data, article drafting and revising; J.C.I.: design of study, acquisition, analysis and interpretation of data, article drafting and revising.

## Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-84924-0>.

**Correspondence** and requests for materials should be addressed to S.C.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021