Doppler-based fetal heart rate analysis markers for the detection of early intrauterine growth restriction

Lisa Stroux, University of Oxford
Christopher W. Redman, University of Oxford
Antoniya Georgieva, University of Oxford
Stephen J. Payne, University of Oxford
Gari Clifford, Emory University

Journal Title: Acta Obstetricia et Gynecologica Scandinavica
Volume: Volume 96, Number 11
Publisher: Wiley: 12 months | 2017-11-01, Pages 1322-1329
Type of Work: Article | Post-print: After Peer Review
Publisher DOI: 10.1111/aogs.13228
Permanent URL: https://pid.emory.edu/ark:/25593/tjw9v

Final published version: http://dx.doi.org/10.1111/aogs.13228

Copyright information:
© 2017 Nordic Federation of Societies of Obstetrics and Gynecology

Accessed August 4, 2019 6:58 AM EDT
Doppler-based Fetal Heart Rate Analysis Markers for the Detection of Early Intrauterine Growth Restriction

Dr Lisa Stroux, PhD¹, Professor Christopher W. Redman, MB, BCHir², Dr Antoniya Georgieva, PhD², Professor Stephen J. Payne, PhD¹, and Professor Gari D. Clifford, PhD³

¹Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, UK
²Nuffield Department of Obstetrics & Gynaecology, University of Oxford, UK
³Departments of Biomedical Informatics and Engineering, Emory University and Georgia Institute of Technology, USA

Abstract

Introduction—One indicator for fetal risk of mortality is intra-uterine growth restriction (IUGR). Whether markers reflecting the impact of growth restriction on the cardiovascular system, computed from a Doppler-derived heart rate signal, would be suitable in its detection antenatally were studied.

Material and Methods—We used a cardiotocography archive of 1163 IUGR cases and 1163 healthy controls, matched for gestation and gender. We assessed the discriminative power of short-term variability (STV) and long-term variability (LTV) of the fetal heart rate, computed over episodes of high and low variation aiming to separate growth-restricted fetuses from controls. Metrics characterizing the sleep state distribution within a trace were also considered for inclusion into an IUGR detection model.

Results—Significant differences in the risk markers comparing growth-restricted with healthy fetuses were found. When used in a logistic regression classifier, their performance for identifying IUGR was considerably superior before 34 weeks gestation. LTV in active sleep was superior to STV (AUROC of 72% compared to 71%). Most predictive was the number of minutes in high variation per hour (AUROC of 75%). A multivariate IUGR prediction model improved the AUROC to 76%.

Conclusion—We suggest that heart rate variability markers together with surrogate information on sleep states can contribute to the detection of early-onset IUGR.

Keywords
Cardiotocography; Doppler; heart rate variability; intrauterine growth restriction; low-cost automated fetal monitoring
**Introduction**

IUGR is a pathological decrease in fetal growth rate whereby the fetus cannot reach its full genetic growth potential, and is associated with significantly increased mortality and morbidity (1, 2).

Developing countries carry the highest burden of perinatal mortality (98%) (3) of which a significant proportion can be attributed to IUGR-based complications (IUGR prevalence estimated at over 10%) (4). Key reasons for this include lack of systematic screening for IUGR and robust referral.

In the context of a well-equipped hospital several measures are typically available to the clinician when screening for, diagnosing and managing IUGR. In addition to maternal history, available tests include symphysis fundal height measurement, the biophysical profile score and fetal ultrasound biometry and multivessel Doppler studies.

In the absence of sophisticated equipment in low-resource settings, IUGR detection is limited to the identification of maternal risk factors and the measurement of fundal height over time. The latter has high inter-observer and intra-observer variability (5) and its diagnostic value has been reported with highly varied results, even for relatively well trained clinical staff (6). Given the scarcity of information available to healthcare workers in low resource settings, additional insight into the developmental status of the fetus may provide valuable decision support. The fetal cardiac signal is the most accessible source of physiological information on the well-being of the fetus. A low-cost method to record this signal is through the use of hand-held Doppler devices.

Growth restriction remains an area of active research due to the complexity surrounding accurate identification and adequate management. Several researchers have investigated either the detection of growth restriction (7–14) or its management (15–19) by assessing its impact on the fetal cardiovascular system. One extensively researched measure of fetal well-being is heart rate variability (HRV).

Studies in literature assessing the difference in HRV marker values between IUGR and healthy fetuses have however been largely underpowered, rendering any generalization of results beyond the subjects studied unfeasible (7–14).

We therefore set out to investigate whether analysis of the Doppler-based fetal heart rate signal (on a convenience sample of CTG traces) can contribute to the detection of IUGR risk.
Material and methods

Database

To confirm the utility of selected markers for the detection of IUGR fetuses antepartum, we utilized the extensive computerized cardiotocography (cCTG) archive of fetal heart rate variables collected between 1990 and 2011 at the John Radcliffe hospital in Oxford, UK. The cardiotocograms were recorded and analyzed using the Sonicaid FetalCare system (Huntleigh Healthcare Ltd.). Ethics approval to use this database was given by NRES Committee South Central – Oxford A (REC Reference 13/SC/0153).

All fetuses with a birthweight under the third percentile [adjusted Yudkin birthweight centile which is gender and gestational age specific (20)] were considered growth-restricted and included in the dataset for analysis. All cCTG traces selected were the last recording of each woman prior to delivery. The control group of normal fetuses was matched for both gender and gestational age at the time of recording. The resultant study dataset of 1163 cases and 1163 control cases is the largest of its kind reported on by more than a factor of 10.

To confirm an initial assumption that IUGR fetuses can be better distinguished from healthy fetuses earlier in gestation the study database was divided into two subsets. The first subset included fetuses with their last trace recorded at or before 34 weeks, the second subset all fetuses with their last trace recorded after 34 weeks of gestation. The characteristics of the two populations are summarized in Table 1.

cCTG markers

Fetal STV and LTV were readily available in the database, computed by the cCTG system according to Dawes and Redman (21).

The fetal heart rate shows change corresponding to the respective biological sleep state. A healthy fetus cycles through episodes of active and quiet sleep from 28 weeks of gestation, which the Dawes/Redman system screens for as part of its automated analysis (21). Periods of active sleep are associated with fetal movement and high heart rate variability. Quiet sleep in contrary is characterized by the absence of movement and low heart rate variability (22).

An episode of active sleep is an indicator of fetal wellbeing and is one of the key criteria used in the Dawes/Redman system to advise on normality.

If criteria are met, the system can be switched-off, resulting in recordings of dissimilar length, typically between 10 to 60 minutes, where short traces contain a high percentage of active sleep and longer traces can be composed of, for example, 54 minutes of quiet sleep with low variability followed by only a short period of high variation. Overall measures of short- and long-term variability calculated for each trace therefore contain a bias due to a potentially large sampling error dependent on the recording length. STV and LTV were therefore computed as average over periods of active and quiet sleep in each trace.
Metrics characterizing the trace composition, such as the number of episodes of quiet, active
and indeterminate sleep and their average duration, have been included as features in the
discrimination model (All features are listed in Table 2).

Statistical analyses

As the distribution of the selected risk markers STV and LTV is non-Gaussian, their
correlation was assessed using Spearman’s rho. It was investigated whether the differences
in risk marker values were statistically significant in the growth restricted and normal
population using the non-parametric Mann-Whitney U-test, visualized with box plots. A p-
value of <0.05 was selected as the level of significance.

The data were quasi randomly divided into a training and test set (70/30 split). Binary
logistic regression classifiers were trained for all individual single risk markers, as well as
for every possible combination including gestational age (in weeks) at time of recording
using five-fold cross-validation on the training set. The associated average area under the
receiver operator curve (AUROC or simply AUC) was computed on the held out folds for
comparison of the predictive power of the risk markers and their combinations. The AUC
statistic of the overall best performing feature combination was compared for statistical
significance with the AUC of the remaining classifiers achieving the best result provided the
selection of input variables. Additionally, the AUC statistic of the best performing single
marker was compared for statistical significance with the AUC of the remaining.
Significance values were adjusted with the Holm-Bonferroni method to correct for multiple
testing and falsely rejecting the null hypothesis (Type 1 errors) (23). The best performing
single- and multiple-feature classifier was tested on a held-out test set and the AUC reported
for comparison with the training results.

Data pre-processing

The dataset contained a total of 387 traces (314 IUGR and 73 healthy controls) with no
episodes of high variability and 1410 traces (580 IUGR and 820 healthy controls) with no
episodes of low variability. Traces with no high variability indicate that the monitored fetus
did not enter a period of high heart rate variation within 60 minutes of recording. Fetuses
that do not enter a phase of active sleep within one hour may be suspected of being
compromised; they are therefore of particular interest in this investigation and should be
included in the analysis. The majority of traces with no low variation are traces that have
met criteria after an episode of high variability and have been terminated before a period of
low activity could occur. The absence of low variation in a trace does therefore not allow for
assumptions concerning fetal wellbeing to the same degree. However, short- and long-term
variability during episodes of low heart rate variation may contain information on the
severity of compromise. Substitute long-term variability and short-term variability values
were therefore estimated for both ‘missing’ high and low periods.

Substitution estimates were established based on the data available. For long-term variability
in active sleep episodes (LTVinHi) the lowest value of existing LTVinHi measures was
identified in each gestational age group with ten or more traces available. A regression line
was fitted and LTVinHi substitute values computed for each week of gestation.
This approach was based on the assumption that the minute ranges in traces without an identified episode of high variability would have always been equal to or lower than this constructed threshold, as they would have otherwise been classified as of high variation.

The same approach was used to estimate substitutes for LTVinLo and STVinHi. No clear threshold could be identified in short-term variability in quiet sleep episodes; substitute values were inferred from a regression line fitted to the mean of existing STVinLo measures instead.

Compared to the thresholds used in the cCTG analysis to classify periods of varying variability, as described by Pardey (21), the substitute values appear to be a reasonable choice. LTVinHi substitute values lie above the 32ms and LTVinLo below the 30ms reported.

**Results**

The average distribution of episodes of high, low and indeterminate variability per gestational age group is shown in Figure 1. In accordance with previous studies, an increase of active periods with maturation of the fetus, concomitant with its sleep organization is clearly visible. This is true for both IUGR and normal traces. IUGR fetuses show a lower percentage of high variability (active sleep) compared to the normal population, in particular before 35 weeks of gestation.

Variability features appear to discriminate better earlier on in gestation, in both male and female traces. This can be observed in the example boxplots showing short-term variability averaged over episodes of active sleep (Figure 2). As expected the quality of risk markers declined with increasing gestational age.

The performance of individual features for 23–34 weeks of gestation is listed in Table 3. The number of minutes in high variation per hour (HiMinph) was most predictive with an AUC of 75% on the training set and 73% on the test set. HiMinph was statistically not significantly superior to the variability risk markers in active sleep, LTVinHi (74% training, 72% test) and STVinHi (71% training, 71% test). Both LTVinHi and STVinHi were found to be significantly different in IUGR traces compared to normal controls (p<0.01). The AUC of multivariate models also declined with increasing gestational age. The model including the features gestational age, long-term and short-term variability in high variation episodes, the average duration in high variation and the number of high episodes in the trace, could improve the discriminative performance to 77% on the training set and 76% on the test set. The results are summarized in Table 4, including those for 35 to 42 weeks of gestation for comparison.

The correlation between HRV markers was assessed in those traces containing periods of both low and high variability. The correlation in IUGR and normal traces was investigated separately to uncover whether relationships depend on the developmental state of the fetus. LTVinHi and STVinHi were found to be correlated in both IUGR traces and normal traces, however to a greater extent in the control group (+0.627, p<0.001 and +0.765, p<0.001
respectively). LTVinLo and STVinLo also showed a correlation in both traces of growth restricted fetuses and healthy controls (+0.690, p<0.001 and +0.585, p<0.001 respectively).

**Discussion**

Our aim was to evaluate the utility of a novel combination of Doppler-based fetal HRV markers to identify fetal growth restriction. Analysis of our extensive dataset confirmed that the distribution of marker values differs significantly in growth-restricted fetuses compared to normal controls, also when averaging over periods of active and quiet sleep respectively.

The Dawes/Redman system applies an automated approach to the detection of fetal activity states based on the presence of accelerations and long-term variability measures tested against age-specific thresholds (21). The alternative method of sleep state assessment described by Nijhuis (24) relies on experts visually scoring FHR traces according to defined patterns. Multi-expert annotations were not feasible in the context of this investigation and with a view of translating this research to a low-resource environment an automated approach is favored.

Sleep organization is part of the fetal maturation process. It has been described in literature that sleep states can only be reliably identified in the second half of the third trimester by pattern assessment (at around 36 weeks of gestation) (24); the documentation of the Dawes/Redman system reports cycles of active and quiet sleep (high and low variation) to be present from 28 weeks onwards. Whether episode of high and low variation in younger and compromised fetuses in this analysis truly represent respective sleep states is unknown. However, a clear difference between IUGR and healthy fetuses in trace composition could be observed in our study and the low percentage of high variability in IUGR fetuses may well be interpreted as delayed or compromised sleep state organization as observed in literature (25, 26).

The importance of activity states in the analysis of variability markers was suspected and could be confirmed in the investigation. The univariate model outcomes of markers computed for individual episodes of high and low variability have revealed the superior discriminative power of markers associated with active compared to quiet sleep episodes.

This indicates that the fetal behavioral state may confound heart rate variability analysis in general. This has also been observed by Sriram et al. (27). Two investigations could be identified in literature that took sleep state information into account when analyzing the difference in HRV metrics in IUGR and normal fetuses (13, 27). Both found their respective metrics under investigation to be lower in growth restricted compared to normal fetuses and both included fetuses of less than 36 weeks of gestation. A direct comparison of results was not possible as previous studies were based on magnetocardiogram data and sleep state identification was based on expert review.

This investigation could confirm the dependence of heart rate variability markers on gestational age, which has been reported in the literature before (7, 12). Risk marker values increased with gestational age in both growth-restricted and normal cases. The rate of
change in compromised fetuses was, however, much higher with marker values being considerably lower earlier in gestation and approximating to control values towards term.

The results suggest that early-onset IUGR (≤34 weeks of gestation) and late-onset IUGR (>34 weeks of gestation) do not impact the fetal cardiovascular system in the same way.

The study outcomes indicate that HRV markers together with information on sleep states might contribute to the detection of early-onset IUGR. The multivariate classifier is of acceptable clinical utility [with an AUC of greater than 0.75 (28)]. Identifying early-onset IUGR is of particular importance in settings where prevalence is high due to potentially addressable factors such as malnutrition or substance abuse. Positive perinatal outcomes have recently been reported for this high-risk group (29) and attributed to advances in monitoring and neonatal care. However, even relatively simple interventions, such as targeted nutritional programs, may lead to significant impacts in IUGR rates. Appropriate timely referral is therefore of utmost importance.

The difficulty of detecting late-onset IUGR has been previously noted in relation to other monitoring modalities, such as umbilical artery Doppler (12, 30). This evidence, together with the lower performance of our approach in the late stages of pregnancy, indicates that alternative Doppler-based identifiers should be a priority for future research.

We note several limitations in our study. First, we note that this is a retrospective study based on a large convenience sample; subjects were not recruited specifically for the purposes of identifying IUGR. As such, a prospective study will be needed to confirm our approach. Second, the work includes population-specific growth charts used for the identification of IUGR babies; consideration as for the appropriateness of growth curves is required when translating presented findings to other study populations. The patients were drawn from a population of mostly white UK residents, and therefore the metrics may not apply to any demographic which is significantly different. It is important to note that IUGR is a complex phenomenon, its etiology is wide ranging and challenges surrounding differentiation from SGA babies persist. However, regardless of the etiology of IUGR, we expect to see FHR changes (a lack of maturation of the central nervous system relative to gestational age) indicative of growth restriction that suggests onward referral to identify the etiology and hence the appropriate action. The key difference will be that there may be a rate of change of maturation related to the effect that may require a recalibration of the features and coefficients of the model. Third, fundal height measurements were not available in the dataset and therefore could not be included in our study for comparative effectiveness. However, as noted earlier, evidence around this method is highly varied. Nevertheless, there may indeed be independent predictive power in fundal height measures, and future work should, if practical, include fundal height as an additional parameter in our model to investigate if the measure provides additional predictive power for a group of midwives with a given training level. Fourth, we also note that fetal growth restriction was defined by the newborn’s birth weight percentile (<3rd percentile) only. Antenatal measures to assess IUGR, such as Doppler velocimetry of fetal vessels and maternal uterine artery were not accounted for, as they were not available.
Finally, while we acknowledge that 0.76 is a modest AUC, we believe the results indicate that there is good evidence to suggest that FHR-based assessment may be a constituent part of any low cost IUGR screening method. New studies, currently under way, are aimed at adding additional non-HRV metrics to improve this AUC.

The low temporal resolution of the cardiotocograph’s fetal heart rate signal is suspected to be a major limitation for IUGR detection. Ongoing work therefore concerns the development of raw Doppler-based, pseudo beat-to-beat, fetal heart cycle segmentation. We hypothesize that risk markers derived from such a signal can improve IUGR identification and we are in the process of collecting (hand-held) Doppler data to test this hypothesis. The translation of this research to the use of hand-held Doppler in low-resource environments is a work in progress (31).

Acknowledgments

LS acknowledges the support of the RCUK Digital Economy Programme grant number EP/G036861/1 (Oxford Centre for Doctoral Training in Healthcare Innovation). AG was supported by the Action Medical Research and the Henry Smith Charity. GC acknowledges the support of the National Institutes of Health, the Fogarty International Center and the Eunice Kennedy Shriver National Institute of Child Health and Human Development, grant number 1R21HD084114-01.

Abbreviations

- **AvgDurHi**: Average duration of episodes of high variation
- **AvgDurLo**: Average duration of episodes of low variation
- **AUC**: average area under the receiver operator curve
- **(c)CTG**: (computerized) Cardiotocography
- **LoMin**: Number of minutes in low variation normalized by the length of the recording
- **LTV**: long-term variability
- **LTVinHi**: Long-term variability averaged over periods of high heart rate variation
- **LTVinLo**: Long-term variability averaged over periods of low heart rate variation
- **HiEpiStrt**: Onset of the first high variation episode
- **HiMin**: Number of minutes in high variation normalized by the length of the recording
- **HRV**: heart rate variability
- **IUGR**: intra-uterine growth restriction
- **MCG**: Magnetocardiography
- **NoHiEpi**: Number of high variation episodes normalized by the length of the recording
- **NoLoEpi**: Number of low variation episodes normalized by the length of the recording
**RecGest**  Gestational age as estimated at time of recording

**SGA**  small for gestational age

**STV**  short-term variability

**STVinHi**  Short-term variability averaged over periods of high heart rate variation

**STVinLo**  Short-term variability averaged over periods of low heart rate variation

**References**


Doppler-based heart rate variability markers together with information on sleep states can contribute to the detection of early-onset IUGR. This supports the notion of developing screening algorithms for resource-poor settings with high IUGR prevalence based on low-cost technology.
Figure 1.
The average percentage of minutes of low, indeterminate and high HRV in IUGR (a) and normal (b) traces. Fetal traces from the entire database (8783 traces of 2326 patients) over 10 minutes in length were grouped into extremely premature (21–26 weeks), moderately premature (27–30 weeks), mildly premature (31–34 and 35–37 weeks), term (39–40) and post-term (41–42 weeks) (Range of recording length: 11–60 min).
Figure 2.
Top graphs: The difference of STVinHi in the male (a) and female (b) IUGR population (red circles) compared to controls (white boxes). Bottom graphs: The number of male (a) and female (b) IUGR cases and controls in each gestational age group. Of all age groups with N \(\geq 15\) for both cases and controls the two male study populations differ significantly at 29 and 31 to 36 weeks \((p < 0.05)\), the two female study populations differ significantly at 28, 30, 31, 32, 36 and 37 weeks \((p < 0.05)\).
Table 1
Characteristics of the study subgroups of 23 to 34 weeks and 35 to 42 weeks gestation. Data are presented as quantity or the median with the range provided in brackets.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>IUGR</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>23–34 weeks gestation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of fetuses</td>
<td>463</td>
<td>463</td>
</tr>
<tr>
<td>Gestational age at CTG (weeks)</td>
<td>32 [23–34]</td>
<td>32 [23–34]</td>
</tr>
<tr>
<td>CTG recording length (minutes)</td>
<td>60 [10–60]</td>
<td>20 [10–60]</td>
</tr>
<tr>
<td>Time to delivery (days)</td>
<td>1.29 [0.1–86]</td>
<td>46.04 [0.1–126.6]</td>
</tr>
<tr>
<td><strong>35–42 weeks gestation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of fetuses</td>
<td>700</td>
<td>700</td>
</tr>
<tr>
<td>Gestational age at CTG (weeks)</td>
<td>37 [35–42]</td>
<td>37 [35–42]</td>
</tr>
<tr>
<td>CTG recording length (minutes)</td>
<td>20 [10–60]</td>
<td>16 [10–60]</td>
</tr>
<tr>
<td>Time to delivery (days)</td>
<td>2.06 [0.1–41.3]</td>
<td>5.57 [0.1–50.1]</td>
</tr>
</tbody>
</table>
Table 2

Selected features from cardiotocographic risk markers, trace characteristics and patient information for IUGR classification.

<table>
<thead>
<tr>
<th>Risk marker (feature) [unit]</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>LTVinHi [ms]</td>
<td>The long-term variability averaged over periods of high heart rate variation.</td>
</tr>
<tr>
<td>STVinHi [ms]</td>
<td>The short-term variability averaged over periods of high heart rate variation.</td>
</tr>
<tr>
<td>LTVinLo [ms]</td>
<td>The long-term variability averaged over periods of low heart rate variation.</td>
</tr>
<tr>
<td>STVinLo [ms]</td>
<td>The short-term variability averaged over periods of low heart rate variation.</td>
</tr>
<tr>
<td>NoHiEpi [ph]</td>
<td>The number of high variation episodes normalized by the length of the recording.</td>
</tr>
<tr>
<td>NoLoEpi [ph]</td>
<td>The number of low variation episodes normalized by the length of the recording.</td>
</tr>
<tr>
<td>HiMin [ph]</td>
<td>The number of minutes in high variation normalized by the length of the recording.</td>
</tr>
<tr>
<td>LoMin [ph]</td>
<td>The number of minutes in low variation normalized by the length of the recording.</td>
</tr>
<tr>
<td>AvgDurHi [min]</td>
<td>The average duration of episodes of high variation.</td>
</tr>
<tr>
<td>AvgDurLo [min]</td>
<td>The average duration of episodes of low variation.</td>
</tr>
<tr>
<td>HiEpiStrt [min]</td>
<td>The onset of the first high variation episode.</td>
</tr>
<tr>
<td>RecGest [weeks]</td>
<td>The gestational age as estimated at time of recording.</td>
</tr>
</tbody>
</table>
Table 3

The AUCs of individual risk markers for fetuses of 23–34 weeks. Results are reported on a held out test set after training using five-fold cross-validation on training data (70/30 split). The best performing feature, the minutes per hour in high variation (HiMinph), was significantly superior than gestational age and markers associated with low variation episodes (p<0.01), with the exception of (LoMinph). No difference was found when comparing to features also derived from high variability episodes. Gender has been excluded from the analysis as it will not be available in anticipated use context.

<table>
<thead>
<tr>
<th>Features*</th>
<th>23–34 weeks gestation</th>
<th>Training data AUC (% [95% CI])</th>
</tr>
</thead>
<tbody>
<tr>
<td>LTVinHi</td>
<td>74 [70–78]</td>
<td></td>
</tr>
<tr>
<td>LTVinLo</td>
<td>69 [65–73]</td>
<td></td>
</tr>
<tr>
<td>STVinHi</td>
<td>71 [67–75]</td>
<td></td>
</tr>
<tr>
<td>STVinLo</td>
<td>46 [41–50]</td>
<td></td>
</tr>
<tr>
<td>AvgDurHi</td>
<td>73 [69–76]</td>
<td></td>
</tr>
<tr>
<td>AvgDurLo</td>
<td>71 [66–75]</td>
<td></td>
</tr>
<tr>
<td>NoHiEpiph</td>
<td>75 [71–79]</td>
<td></td>
</tr>
<tr>
<td>NoLoEpiph</td>
<td>69 [65–73]</td>
<td></td>
</tr>
<tr>
<td>HiMinph</td>
<td>75 [71–79]</td>
<td></td>
</tr>
<tr>
<td>LoMinph</td>
<td>71 [67–75]</td>
<td></td>
</tr>
<tr>
<td>HiEpiStrtMin</td>
<td>63 [59–67]</td>
<td></td>
</tr>
<tr>
<td>RecGest</td>
<td>48 [43–52]</td>
<td></td>
</tr>
</tbody>
</table>

* All features are defined in Table 2.
Table 4

The best performing multivariate logistic regression classifier for 23–34 (top) and 35–42 weeks of gestation (bottom). Results are reported on a held out test set after training using five-fold cross-validation on training data (70/30 split).

<table>
<thead>
<tr>
<th>23–34 weeks gestation</th>
<th>Training data</th>
<th>Test data</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC [%]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Features *: RecGest, LTVinHi, STVinHi, AvgDurHi, NoHiEpiph)</td>
<td>77</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td>(95% CI = 73–80)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>35–42 weeks gestation</th>
<th>Training data</th>
<th>Test data</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC [%]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Features *: LTVinHi, NoHiEpiph)</td>
<td>57</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td>(95% CI = 54–61)</td>
<td></td>
</tr>
</tbody>
</table>

* All features are defined in Table 2.