

EP16.85

Prediction of Spontaneous and Iatrogenic Preterm Birth: Comparison of Statistical and Machine Learning Models

A. McLennan¹, D. L. Rolnik², K. Hjort-Petersen¹, J. Hyett³, F. Da Silva Costa⁴

¹Sydney Ultrasound for Women, Sydney · ²Monash University, Melbourne · ³University of Sydney, Sydney · ⁴Griffith University, Gold Coast · Australia

■ BACKGROUND

Preterm birth (PTB) is the leading cause of neonatal death and comprises two distinct entities, spontaneous and iatrogenic (medically indicated). Whether first-trimester screening markers, and machine learning applied to them, can predict PTB and improve on standard regression models is unclear.

■ OBJECTIVE

- Compare statistical, penalized and machine learning models for first-trimester prediction of PTB using baseline characteristics and biomarkers.
- Model spontaneous and iatrogenic PTB separately, using a competing-risks framework.

■ METHODS

- Retrospective cohort of singleton pregnancies undergoing first-trimester PE screening; PTB defined as delivery before 37 weeks, modelled overall and by subtype.
- Predictors: maternal factors, obstetric history, mean arterial pressure (MAP), Uterine artery pulsatility index (UtA-PI) on ultrasound, pregnancy-associated plasma protein A (PAPP-A), placental growth factor (PlGF) and Fetal Medicine Foundation (FMF) first trimester risk scores for pre-eclampsia and fetal growth restriction.
- Models compared: logistic regression, Cox proportional hazards, Fine-Gray competing risks, penalized regression (LASSO, Elastic Net), Random Forest and gradient boosting (XGBoost).
- Performance assessed by 10-fold cross-validation (AUC, detection rate, Brier score).

■ CONCLUSIONS

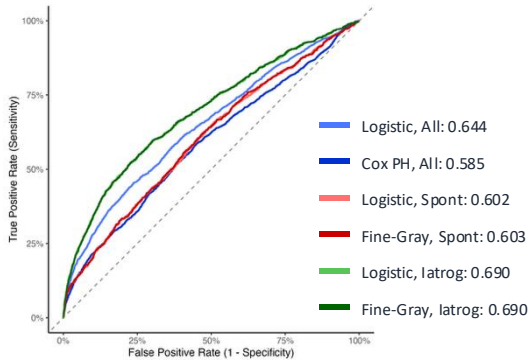
Prediction was modest overall and better for iatrogenic than spontaneous PTB, reflecting the placental basis of the biomarkers. Machine learning offered at most marginal gains over logistic regression. Simpler, well-calibrated models remain preferable for clinical use.

■ RESULTS

Study population: 36,152 singleton pregnancies, with 2,020 preterm births (5.6%): 950 spontaneous (2.6%) and 1,070 iatrogenic (3.0%).

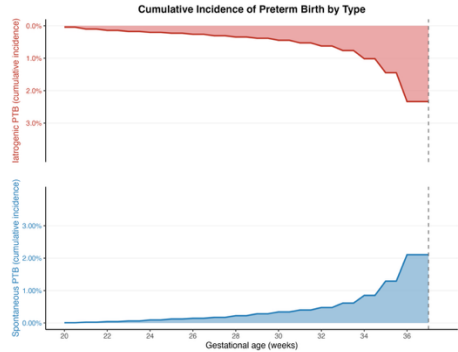
Prediction was modest overall and better for iatrogenic than spontaneous PTB. Machine learning did not outperform logistic regression.

Figure 1. Cross-validated ROC curves



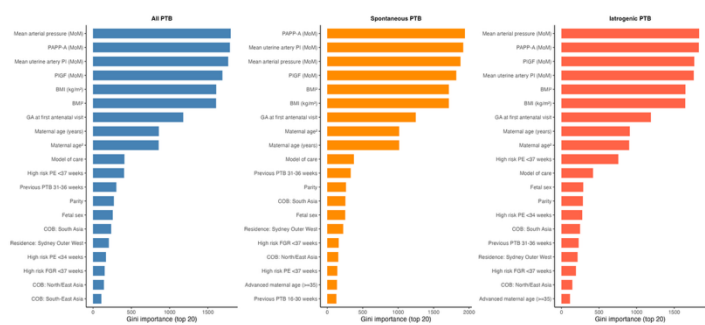
Cross-validated AUCs.

Figure 2. Cumulative incidence of PTB by type



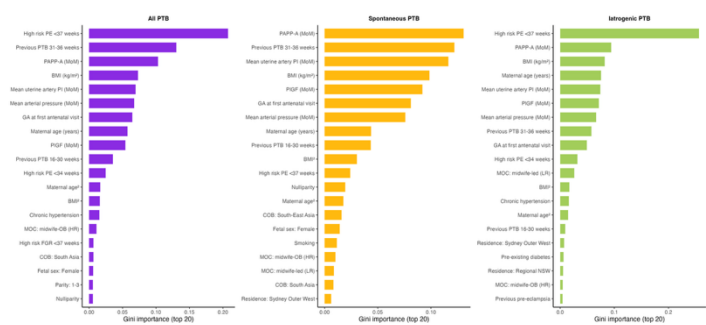
Iatrogenic (top) accumulates faster than spontaneous; final incidence 2.4% vs 2.1%.

Figure 3. Random Forest variable importance (top 20 by Gini index)



Biomarkers (MAP, PAPP-A, UtA-PI, PlGF) and BMI rank highest across all three outcomes; obstetric history contributes to the tree models. Discrimination was near-identical across models.

Figure 4. XGBoost variable importance (top 20 by Gini gain)



■ REFERENCES

1. Ohuma EO, et al. *Lancet* 2023;402:1261-1271. 2. Aughey H, et al. *BJOG* 2023;130:33-41. 3. Austin PC, Lee DS, Fine JP. *Circulation* 2016;133:601-609. 4. Arain Z, et al. *Curr Res Physiol* 2023;6:100099.

Cross-validated AUC by model

AUC	All	Spontaneous	Iatrogenic
Logistic	0.644	0.602	0.690
Cox / FG	0.585	0.603	0.690
LASSO	0.651	0.604	0.699
Elastic Net	0.651	0.604	0.698
Random Forest	0.624	0.559	0.672
XGBoost	0.644	0.586	0.683

Detection at 10% FPR: 24.7 to 34.1% (all), 20.0 to 24.6% (spont.), 37.2 to 40.3% (iatrog.).