Early Prediction of Small-for-Gestational-Age (SGA) Babies at Second Trimester using Machine Learning Models

Shier Nee Saw[1],Citra Nurfarah Zaini Mattar[2], Arijit Biwas[2], Hwee Kuan Lee[3,4,5], Choon Hwai Yap[1]

1. Department of Biomedical Engineering, National University of Singapore, Singapore; 2. Department of Obstetrics and Gynecology, Yong Loo Lin School of Medicine, National University of Singapore, National University Health Systems, Singapore; 3. Bioinformatics Institute, Agency for Science, Technology and Research (A*STAR). 4. CNRS, IPAL, Singapore, Singapore. 5. Department of Computer Science, School of Computing, National University of Singapore, Singapore

Introduction

- **Small-for-gestational-age (SGA**) refers to babies with birthweight (BW) below the 10th centile.
- Current diagnosis of SGA screening includes (i) fetal size measurements and (ii) Doppler ultrasound.

Figure 1. Two term neonates at the same gestational age Image: https://nursekey.com/15-the-newborn-at-risk-conditions-associated-with-gestational-age-and-development

- **Problem Statements:**
  - 46% of SGA is undetected before delivery by conventional ultrasound methods (Uterine RI, Uterine PI) [Lindqvist and Molin, 2005].
  - Undetected SGA neonates have higher adverse perinatal outcomes (perinatal death, cerebral palsy, intracranial hemorrhage) [Lindqvist and Molin, 2005].

- Machine Learning model is a type of data analysis that uses algorithm to learn patterns from dataset and makes predictions.
  - Detect myocardial infarction [Dohare et.al, 2018]
  - Predict cardiovascular disease [Yekkala and Dixit, 2018, Poplin et al., 2018]

Methods and Materials

Data Collection

- Retrospective study. BW < 10th centile was defined as SGA while BW > 10th centile was defined as control.
- We collected a total of 202 SGA, 145 Control.
- 16 measurements at second trimester were collected, such as maternal demographics, fetal brain and size measurements, uterine Doppler and birthweight (BW).

Machine Learning Model

- Support Vector Machine (SVM), Random Forest (RF) and Neural Network (NN) were explored.
- Data were split into training (80% of data) and testing (20% of data) datasets.
- Training data were used to train the model while testing data were used to evaluate the performance of the models.
- Performance of models were compared with clinical diagnosis where diagnosis was made based on RCOG Green Top Guidelines no. 31.

Objectives

- **To use machine learning models in investigating the prediction of SGA at birth** (BW below the 10th centile) using second trimester ultrasound measurement and maternal demographics.
- To **determine the most important feature in predicting SGA at birth.**

Results

(A) Performance Comparison between Machine Learning models and Clinician Diagnosis

- Using clinical guideline to predict SGA achieved only 32% in sensitivity.
- Machine learning models sensitivity in predicting SGA were 2 times higher than the clinical diagnosis with a sensitivity of 65% in predicting SGA at birth,
- However, ML model had lower specificity (56-70%) than clinical diagnosis.

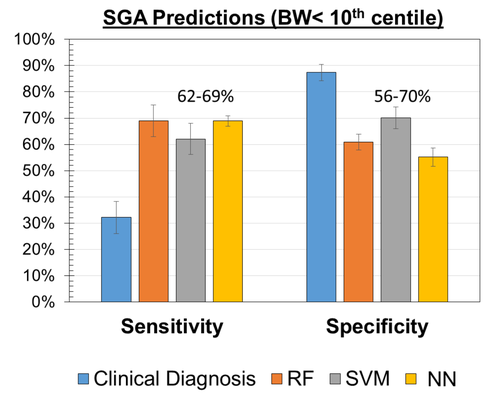
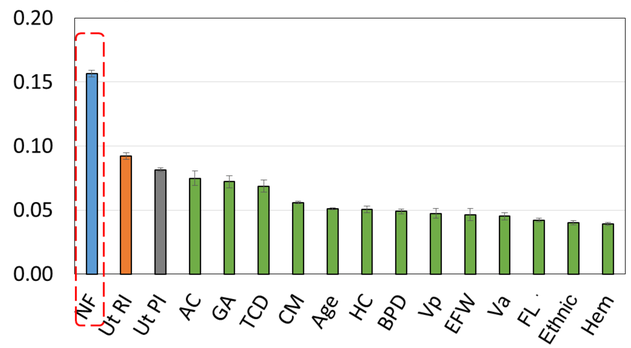


Figure 2. Comparison between clinical diagnosis and machine learning models' capabilities in predicting SGA at birth using the second trimester measurements on the testing datasets. RF: Random Forest, SVM: Support Vector Machine, NN: Neural Network

(B) Features Importance

- Nuchal fold thickness had highest importance score in predicting SGA, instead of Uterine RI and PI which are the clinical predictors for SGA.



Abbreviation:
GA: Gestational Age
NF: Nuchal Fold Thickness
Ut RI: Uterine Resistance Index
Ut PI: Uterine Pulsatility Index
AC: Abdominal Circumference
FL: Femur Length
EFW: Estimated Fetal Weight
BPD: Biparietal Diameter
TCD: Transverse Cerebellar Diameter
Va: Anterior Horn of Lateral Ventricle
Vp: Posterior Horn of Lateral Ventricle
Hem: Hemisphere of fetal cerebral

Figure 2. Parameters importance ranking in SGA prediction, ranking in the ascending order. Nuchal fold was shown to be the most important parameter in predicting SGA at birth using second trimester ultrasound measurements.

(C) Statistical Analysis

- SGA neonates have significantly lower nuchal fold thickness at second trimester.
- Possible reason: Could be a due to reduced growth rates, resulting in reduced nuchal fold adipose tissue at neck area.

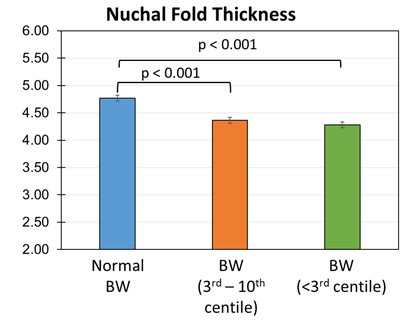


Figure 3. ANOVA test results comparing nuchal fold thickness for fetuses with normal birthweight (BW), 3rd-10th centile BW and BW below the 10th centile

Conclusions

- Machine learning allows **early prediction of SGA at birth** and achieved **higher sensitivity than** the clinical diagnosis.
- Machine learning revealed that **nuchal fold thickness** at second trimester was **important in predicting SGA at birth**.
- Machine learning will **help to identify mothers and babies** who would **benefit from nutritional intervention** earlier than is possible.

Future Work

- Increase the robustness and generalizability of the model using more data from other health institutions.
- Identify applications for the model to predict when an adjustment to a mother's diet is needed to prevent SGA.
- Incorporate the model into digital health platforms to provide real-time information to obstetricians about SGA risk when evaluating pregnant mothers via ultrasound.

Contact

Shier Nee, SAW, PhD
Bioinformatics Institute, Agency of Science, Technology and Research
30 Biopolis St, Singapore 138671
Email: sawsn@bii.a-star.edu.sg

Scan and connect with Shier Nee

Martin Lim, MS, MBA
Co-Founder and CEO, ONWARD Health Research Singapore
Email: martin.lim@onwardhealthresearch.com

Scan and connect with Martin

Acknowledgement

The National University of Singapore Young Investigator Award 2015 (PI: Yap) for funding.