

Maternal Associations with Body Mass Index and Gestational Weight Gain

Hypothesis

- Risk of postpartum obesity can be predicted based on health and social-economic factors

Background

- 54% of women in industrialized nations are being classified as either overweight or obese.
- Recent systematic reviews have indicated that excessive GWG increases the risk of the child becoming overweight/obese by 30-40%.
- Overweight pre-pregnancy BMI shows an increased risk in gestational diabetes, large-for-gestational age babies, and pre-eclampsia.

Study Population – All Our Families Cohort

- <http://allourfamiliesstudy.com/background-2/>
- Eligibility:
 - ≥18 years of age, being less than 25 weeks of gestation at the time of recruitment, receiving prenatal care in Calgary, Canada, and be willing to complete written questionnaires in English.
- Approved by the Child Health Research Office and the Conjoint Health Research Ethics Board of the Faculties of Medicine, Nursing, and Kinesiology at the University of Calgary, and written informed consent was obtained prior to blood collection.
- Women responded to questionnaires that assessed participant demographics, socioeconomic status, psychosocial, health, and lifestyle data at multiple timepoints (**Q1 = <25 weeks gestation, Q2 = 34-36 weeks gestation, Q3 = 4 months postpartum**) throughout and after their pregnancy.

Methodology

- **Notebook:** [Project 1215.ipynb](#)
- **Summary**
 - Questionnaire results from 470 Women were used to estimate obesity risk based on 40 variables. ([AOB Sum.csv](#))
 - Data was split into training (75%) and testing sets (25%).
 - Principle component analysis was used to reduce the number of variables while retaining 90% variance. 19 principle components were identified.
 - Four classification models were trained and tested, and their results were compared based on the confusion matrix:
 - Logistic Regression (accuracy: 92.4%)
 - Naïve Bayes (accuracy: 82.2%)
 - K nearest neighbor (accuracy: 88.1%)
 - Random forest (accuracy: 91.5%)

Results and Discussion

- Principle component analysis can greatly reduce the number of variables needed for making prediction.
- All models tested appear to give good performance. Logistic Regression model and random forest model seem to have slightly better performance.
- This approach serves as a good foundation for incorporating additional variables in the future.
 - Serum metabolite concentrations using nuclear magnetic resonance
 - Serum metal ion concentrations using inductively coupled plasma mass spectrometry
- The objective is to use the model to determine obesity risk and provide timely health advises during pregnancy.