Research Repository UCD

Title	Prediction of Large for Gestational Age Infants in Overweight and Obese Women at Approximately 20 Gestational Weeks
Authors(s)	Du, Yuhan, Mehegan, John, McAuliffe, Fionnuala M., Mooney, Catherine
Publication date	2020-09-24
Publication information	Du, Yuhan, John Mehegan, Fionnuala M. McAuliffe, and Catherine Mooney. "Prediction of Large for Gestational Age Infants in Overweight and Obese Women at Approximately 20 Gestational Weeks." ACM, September 24, 2020. https://doi.org/10.1145/3388440.3414906.
Conference details	The 11th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics (ACM-BCB), Virtual Conference, 21-24 September 2020
Publisher	ACM
Item record/more information	http://hdl.handle.net/10197/12183
Publisher's statement	This is the author's version of the work. It is posted here by permission of ACM for your personal use. Not for redistribution. The definitive version was published in 10.1145/3388440.3414906 http://doi.acm.org/10.1145/3388440.3414906
Publisher's version (DOI)	10.1145/3388440.3414906

Downloaded 2026-05-01 23:45:33

The UCD community has made this article openly available. Please share how this access benefits you. Your story matters! (@ucd_oa)

© Some rights reserved. For more information

Prediction of Large for Gestational Age Infants in Overweight and Obese Women at Approximately 20 Gestational Weeks

Yuhan Du
University College Dublin
Dublin, Ireland
yuhan.du@ucdconnect.ie

John Mehegan
University College Dublin
Dublin, Ireland
john.mehegan@ucd.ie

Fionnuala M McAuliffe
University College Dublin
Dublin, Ireland
fionnuala.mcauliffe@ucd.ie

Catherine Mooney
University College Dublin
Dublin, Ireland
catherine.mooney@ucd.ie

ABSTRACT

Large for gestational age (LGA) births are associated with many maternal and perinatal complications. As overweight and obesity are risk factors for LGA, we aimed to predict LGA in overweight and obese women at approximately 20 gestational weeks, so that we can identify women at risk of LGA early to allow for appropriate interventions. A random forest algorithm was applied to maternal characteristics and blood biomarkers at baseline and 20 gestational weeks' ultrasound scan findings to develop a prediction model. Here we present our preliminary results demonstrating potential for use in clinical decision support for identifying patients early in pregnancy at risk of an LGA birth.

ACM Reference Format:
Yuhan Du, John Mehegan, Fionnuala M McAuliffe, and Catherine Mooney. 2020. Prediction of Large for Gestational Age Infants in Overweight and Obese Women at Approximately 20 Gestational Weeks. In *Proceedings of the 11th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics (BCB '20), September 21–24, 2020, Virtual Event, USA.* ACM, New York, NY, USA, 1 page. https://doi.org/10.1145/3388440.3414906

1 INTRODUCTION

Large for gestational age (LGA), commonly defined as infant's birth weight above 90th percentile for his/her gestational age, is an adverse pregnancy outcome associated with many maternal and perinatal complications [2]. Currently, most published models for the prediction of LGA are focused on the late stages of pregnancy (26 – 37 weeks). Therefore, we aimed to develop a prediction model for LGA at approximately 20 weeks of gestation, in a high risk group of overweight and obese women when the size of the fetus is small and it is early enough to allow interventions to prevent LGA. We developed our model to be easy to use in a clinical setting to help clinicians screening for LGA births in early pregnancy.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
BCB '20, September 21–24, 2020, Virtual Event, USA
© 2020 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-7964-9/20/09.
https://doi.org/10.1145/3388440.3414906

2 METHOD

This research used data collected in the PEARS study (ISRCTN 29316280), a randomized controlled trial of an antenatal behavior change intervention to prevent gestational diabetes mellitus in overweight and obese women conducted at the National Maternity Hospital, Dublin, Ireland [1]. We included 465 PEARS study subjects whose newborn babies' birth weight centile are available. Of them, 11.18% delivered a LGA infant. In order to predict LGA at approximately 20 gestational weeks, we only included maternal characteristics and blood biomarkers at baseline (14.89±1.65 gestational weeks), fetal biometry measured sonographically, biometric percentiles, estimated fetal weight (at the time of the scan) and percentile at approximately 20 gestational weeks, as predictors.

After dropping predictors with missing rate above 20% and imputing missing values in the other predictors with median (numerical) and mode (categorical), the data were randomly split into training set (75%) and test set (25%). A random forest algorithm was applied to the training set, with five-fold cross validation to select the optimal model using highest area under precision-recall curve (AUC-PR) and Synthetic Minority Over-sampling Technique (SMOTE) to balance the data. A minimum subset of predictors was selected based on their variable importance in order to make the model fast and easy to use.

3 RESULT AND DISCUSSION

Evaluated on the test set, the model achieved AUC-PR of 0.27 and area under receiver operating characteristic curve (AUC-ROC) of 0.77. At the false positive rate of 5% and 10%, the model achieved sensitivity of 0.31 and 0.38 respectively.

These preliminary results show the potential of applying machine learning in identifying women at risk of LGA in a clinical setting. Further research will be conducted on the selection of features and model validation in other populations.

REFERENCES

[1] Maria A Kennelly, Kate Ainscough, Karen L Lindsay, Elizabeth O'Sullivan, Eileen R Gibney, Mary McCarthy, Ricardo Segurado, Giuseppe DeVito, Orla Maguire, Thomas Smith, et al. 2018. Pregnancy exercise and nutrition with smartphone application support: a randomized controlled trial. *Obstetrics & Gynecology* 131, 5 (2018), 818–826.

[2] Alina Weissmann-Brenner, Michal J Simchen, Eran Zilberberg, Anat Kalter, Boaz Weisz, Reuven Achiron, and Mordechai Dulitzky. 2012. Maternal and neonatal outcomes of large for gestational age pregnancies. *Acta Obstetricia et Gynecologica Scandinavica* 91, 7 (2012), 844–849.