Aarhus University

Bioinformatics Research Center

Exploring the Association between Maternal Cardiometabolic Conditions and Autism in Offspring: A Case-Cohort Study

Jinqi Duan

Supervisor

Jakob Grove, Heidi MacLean

June 15, 2023

Abstract

This case-cohort study investigates the potential impact of maternal cardiometabolic conditions(CMCs), particularly preeclampsia and gestational diabetes, on the risk and age at which a child may be diagnosed with Autism Spectrum Disorder (ASD). The research combines polygenic risk scores (PRS) for heart failure and ASD using Generalized Linear Models (GLM) and Cox models in order to reduce genetic confounding, offer individualized risk assessments, and comprehend the interaction between genetic and environmental components. According to the research, preeclampsia and gestational diabetes may each function as a separate risk factor, directly influencing the likelihood of ASD and the age at which it is diagnosed in offspring. These findings highlight the potential link between mother health during pregnancy and the neurodevelopment and behavioral health of the fetus, highlighting the significance of preeclampsia and gestational diabetes management and prevention during pregnancy. This work illustrates the potential of genetic risk scores in improving our understanding of the etiology of ASD and guiding clinical practice. It adds to the expanding body of information relating maternal health, genetic vulnerability, and ASD risk.

Acknowledgement

I would like to express my deepest gratitude to my supervisors, Jakob and Heidi. Their unwavering support and guidance throughout my research journey have been invaluable. Although I am not good with words, my heart appreciates their kindness and patience. Their mentorship extended beyond academics, teaching me invaluable life and work skills. Despite challenges and setbacks, their endless tolerance and understanding have been a source of constant encouragement. I am particularly inspired by their passion for work and life, which has changed me. They embody the qualities and values that I aspire to develop in myself.

In addition, I must express my heartfelt thanks to my family for their constant support and encouragement, especially Jinjie, whose unconditional support has been a pillar of strength for me.

I am also grateful for the opportunity to be a part of Anders Børglum's group, even if it was for a short while. The warm and relaxed atmosphere was truly inspiring and made the experience all the more valuable.

Lastly, I would like to thank the beautiful sunshine in Denmark, which has been a source of comfort and healing, making all worries seem trivial.

Anyway, I am immensely grateful for the mentorship and the opportunity to learn and grow under their guidance. Their influence will continue to resonate in my future endeavors. Thank you, Jakob, Heidi, and everyone who has been a part of this journey.

Contents

1 Introduction **1**

2 Materials and Method **5**

2.1 Data and Data Source 5

2.2 Statistical Analysis 9

2.2.1 Description of the Models 9

2.2.2 The Rationale for Using Models 14

2.2.3 Assumptions and Limitations of Models 17

2.2.4 Explanation of Variables and Covariates 20

2.2.5 Model Selection Process and Criteria 22

3 Results **24**

3.1 Descriptive statistics 24

3.2 Univariable Analysis Results 29

3.3 Model Selection Process 35

3.4 Multivariable Analysis Results 39

3.5 Model Fit and Diagnostics 46

4 Discussion **49**

5 Literature Cited **51**

6 Appendix A: Polygenic Risk Scores **56**

6.1 Selection of Traits and GWAS Summary Data 57

6.2 Computation of Polygenic Risk Scores 58

7 Appendix B: Stepwise selection data **64**

7.1 Preeclampsia 64

7.2 Gestational diabetes 73

List of Tables

1 Selected polygenic risk scores 8

2 Univariate Logistic Regression Summary of Preeclampsia 29

3 Univariate Logistic Regression Summary of Gestational Diabetes 29

4 Univariate Logistic Regression Summary of Hypertension 30

5 Univariate Cox Regression Summary of Preeclampsia 32

6 Univariate Cox Regression Summary of Gestational diabetes 33

7 Model comparison for the first variable selection in preeclampsia's GLM . . 36

8 Model checking after a new variable is added in preeclampsia' GLM 36

9 Model comparison for the PRS variable selection in preeclampsia's GLM . 37

10 Summary of a multiple logistic regression model for Preeclampsia 39

11 Summary of a multiple logistic regression model for Gestational diabetes . 40

12 Summary of a multiple Cox regression model for Preeclampsia 41

13 Summary of a multiple Cox regression model for Gestational diabetes . . . 43

14 VIF for multiple logistic regression of Preeclampsia 46

15 VIF for multiple logistic regression of gestational diabetes 47

16 The Schoenfeld test for the Cox model of preeclampsia 48

17 The Schoenfeld test for the Cox model of gestational diabetes 48