Polygenic scores of body mass index, mediation through eating behaviours, and their interaction with dietary intakes in childhood.

Danick Goulet

Thesis submitted to the University of Ottawa
in partial fulfillment of the requirements for the
Ph.D. in Epidemiology

School of Epidemiology and Public Health
Faculty of Medicine
University of Ottawa

PREFACE

Ethical disclosures

The research conducted as part of this thesis included data gathered from human subjects from the Quebec Longitudinal Study of Child Development. Until participants reached 10 years of age, informed consent was obtained from the parents before each wave of data collection. After that point, the children gave informed consent directly at each wave of data collection. The initial protocol for the Quebec Longitudinal Study of Child Development was approved by the Ethics commitees of the Institut de la Statistique du Québec and the Centre Hospitalier Universitaire Sainte-Justine Mère-Enfant. The secondary analysis of the data was approved by the University of Ottawa Research Ethics Board (ethics file number: H-01-23-8018, approval date: 10/02/2023). The certificate of approval is available in Appendix 1.

Author contributions

Danick Goulet (Ph.D. candidate) is the author accountable for the content of this thesis and is responsible for leading the conduction of the presented research project with the supervision of Dr. Lise Dubois (supervisor) and Dr. Michel Boivin (external co-supervisor, Université Laval, Québec, Canada). More specifically, Danick Goulet was responsible for designing the study methods, completing statistical analyses, designing the figures and tables, writing the three individual manuscripts (first author in all cases) and the thesis, and communicating the results at conferences when applicable. Dr. Dubois and Dr. Boivin were considered senior authors for all three manuscript and co-corresponding authors.

Dr. Dubois and Dr. Boivin were instrumental in the development of the overall thesis topic following several conversations held in the first year of the candidate's Ph.D., with follow-up conversations

shaping the overall directions of the research project. Dr. Dubois was the main driver for the choice of members of the thesis advisory committee, composed of Dr. Beth Potter, Dr. Julian Little, and Dr. Christopher Gravel. The committee was formed to include expertise in various areas relating to the thesis subject, including nutrition, human genetics, children's health and development, epidemiology, and causal inference. All members of the committee participate in meetings at least once a year to follow the overall progress and advise on the next steps when necessary. Through those meetings, members of the committee held discussions concerning the development of the studies and were involved in revisions of the manuscripts. Dr. Gravel further provided insights into the statistical analyses presented by meeting with the Candidate when necessary, or using email communications, to ensure the appropriateness of the methods used. All members of the committee appear as authors of all three manuscripts included in the thesis with the exception of Dr. Potter for the first manuscript. Although she participated in discussions relating to the manuscript, Dr. Potter felt she did not reach the minimum requirements for authorship and voluntarily withdrew her consideration.

Dr. Boivin invited the candidate to Université Laval on two occasions to present the results from the research project to an audience of colleagues with relevant expertise and receive guidance, specifically regarding the genetic aspects of the research. I also want to acknowledge the contributions of Dr. Till Andlauer, Dr. Stéphane Paquin, Dr. Geneviève Morneau-Vaillancourt, and Dr. Isabelle Ouellet-Morin, who were involved in the quality control of the genetic data of the Quebec Longitudinal Study of Child Development participants that are used in the research. Dr. Geneviève Morneau-Vaillancourt also authored documentation on polygenic score construction using PRS-CS with the Quebec Longitudinal Study of Child Development genetic data.

ABSTRACT

Previous studies found that the effect of polygenic scores (PGS) on body mass index (BMI) is at least partly mediated through eating behaviours, and that PGS can mitigate or elevate the association of dietary intakes on BMI. However, this evidence has several limitations. For instance, there are few investigations in child populations or using longitudinal data. Also, the majority of PGS for BMI in studies of children are constructed from adult genome-wide association studies (GWAS) summary statistics, even though the genetic architecture of obesity may change throughout life.

To address those limitations, the thesis has three specific objectives: 1) to measure the genetic susceptibility to obesity in children using two PGS derived from adults and children GWAS summary statistics, respectively, and compare their association with BMI and discrimination of obesity, 2) to determine the extent to which the association between adult- and child-derived PGS and BMI is mediated through eating behavioural traits (over-eating, fussy eating) assessed in early childhood, and 3) to estimate the association between preschool dietary (food and macronutrient) intakes and BMI (mean and change with age) displayed in childhood and identify the ability of adult- and child-derived PGS to modify those associations.

The objectives are developed in three manuscripts. In the first manuscript, we showed that a PGS derived from adult summary statistics has a stronger association with BMI z-scores from 4 to 13 years, compared to a PGS derived from child summary statistics. Using longitudinal growth curve mediation analysis in the second manuscript, we observed that over-eating expressed from 2 to 6 years of age mediated the association between both PGS and BMI z-scores from 6 to 13 years of age. Notably, the proportion mediated by over-eating decreased as children grew older. Leveraging linear mixed effects models in the third manuscript, we showed that the association between specific dietary intakes (protein, lipid, and total energy) and BMI increased with the value of the adult-derived PGS, but not the child-derived PGS. Overall, this thesis sheds light on the underlying role of dietary habits in genetic susceptibility to obesity across childhood and teenage years.

ABBREVIATIONS

BMI – Body mass index.

CEBQ – Child Eating behaviour Questionnaire.

GIANT – Genetic Investigation of Anthropometric Traits consortium.

GREML – Genomic relatedness restricted maximum likelihood.

GWAS – Genome-Wide Association Study.

LD – Linkage disequilibrium.

LDSC – Linkage disequilibrium score regression.

PGS – Polygenic score.

SD – Standard deviation.

TFEQ – Three-Factor Eating Questionnaire.

ACKNOWLEDGEMENTS

First, I want to thank my supervisor, Dre Lise Dubois. J'ai été très chanceux de pouvoir compter sur ton expertise et ton mentorat tout a long de la complétion de mon Doctorat. J'apprécie particulièrement d'avoir été en mesure d'effectuer l'entièreté de mon Doctorat de la maison, et l'opportunité de travailler sur d'autres projets afin de développer mes aptitudes pour l'analyse de données. Je suis honoré d'avoir été ton dernier étudiant et te souhaite une excellente retraite.

I also want to highlight my co-supervisor, Dr. Michel Boivin, and the three members of my thesis advisory committee, Dre Beth Potter, Dr. Julian Little, and Dr. Christopher Gravel. Michel, merci beaucoup pour ton mentorat tout au long de mon Doctorat. Je suis particulièrement reconnaissant pour l'opportunité d'aller visiter ton équipe à l'Université Laval à quelques reprises. Beth, Julian, and Chris, thank you all for your continued support. I am certain the advice given throughout the writing of three articles will help me for a long time in the future.

I also want to thank all the participants of the QLSCD. The importance of participants in the research process cannot be understated, without you this thesis would not have been possible. The same goes for all the staff and researchers that worked on the QLSCD at one time or another for more than 20 years. While the work accomplished is often done in the background, you have my sincere gratitude.

Thank you to all the professors at the School of Epidemiology and Public Health of the University of Ottawa. I have no doubts the various skills I have acquired attending courses will serve me for years to come.

Finally, a big thank you to my family and my close friends throughout these last four years. J'ai eu la chance d'avoir le support de très bon amis (certains de longue date, et d'autres plus récent). Ces dernières années ont été remplies de changements et je suis fier de pouvoir compter sur vous dans les temps les plus difficiles. Je ne peux également pas oublier le support de ma famille. Caroline et Paul-André, merci de votre support inconditionnel pendent près de 30 ans. Audrey-Anne, je suis chanceux de t'avoir comme soeur.

FUNDING

Throughout the completion of this thesis, I was financially supported by two sources of funding:

2021/01 to 2024/12: Ph.D. admission scholarship, University of Ottawa.
2022/04 to 2024/03: Doctoral training award - Ontario Graduate Scholarship (OGS), Ontario government.

LIST OF TABLES

Tables:

Table 3-1. Description of eating behaviour scores and individual questionnaire items. 44
Table 3-2. Description of food intake variables and individual questionnaire items. 46
Table 4-1. Characteristics of participants (n=717) 72
Table 5-1. Characteristics of study participants. 104
Table 5-2. Association of the PGS with eating behaviors and BMI in children. 106
Table 5-3. Proportion of the effect of PGS on BMI z-score mediated by eating behaviours. 107
Table 6-1. Characteristics of study participants with food intake and macronutrient and energy intake data available. 136
Table 6-2. Age-specific association between food, macronutrient and energy intake, and BMI at 4, 8, and 13 years, and association with change in BMI with age.....138
Table 6-3. Coefficient for the interaction of food, macronutrient and energy intake with PGSs.....139
Table 6-4. Age-specific association of food, macronutrient and energy intake with BMI by PGS value. 140

Supplementary Tables:

Supplementary Table 4-1. Characteristics of participants, included vs. excluded. 77
Supplementary Table 4-2. PGS association with BMI z-score by age. 78
Supplementary Table 4-3. Obesity prediction for PGS, by age. 79
Supplementary Table 4-4. PGS association with BMI z-score by age, stratified by sex..... 80
Supplementary Table 4-5. Obesity odds ratio by age, stratified by sex..... 81
Supplementary Table 5-1. Proportion of the effect of PGS on BMI z-score mediated by over-eating, IPW for attrition.....118
Supplementary Table 6-1. Comparisons of characteristics of study participants with food intake and genetic data and excluded QLSCD participants. 145
Supplementary Table 6-2. Comparisons of characteristics of study participants with macronutrient and energy intake and genetic data and excluded QLSCD participants. 146

Supplementary Table 6-3. Age-specific association between juice and fruit drinks intake and BMI at 4, 8, and 13 years, and association with change in BMI with age, by sex. 147
Supplementary Table 6-4. Age-specific association between macronutrient intake and BMI at 4, 8, and 13 years, and association with change in BMI with age, adjusted for total energy intake. 148
Supplementary Table 6-5. Age-specific association of juice and fruit drinks intake with BMI by PGS value, by sex. 149
Supplementary Table 6-6. Age-specific association of protein and lipid intake with BMI by PGS value, adjusted for total energy intake. 150
Supplementary Table 6-7. Age-specific association between food, macronutrient and energy intake and BMI at 4, 8, and 13 years, and association with change in BMI with age, with IPW. 151
Supplementary Table 6-8. Coefficient for the interaction of food, macronutrient and energy intake with PGSs, with IPW. 152
Supplementary Table 6-9. Age-specific association of food, macronutrient and energy intake with BMI by PGS value, with IPW 153

LIST OF FIGURES

Figures:

Figure 3-1. Flowchart of the number of participants with specific data available from the QLSCD..... 47
Figure 4-1. Linear association between PGS and BMI z-scores, 4 to 13 years old. Beta ± 95% CI..... 73
Figure 4-2. Mean BMI z-scores per PGS quintile at (A) 4 years, (B) 6 years, (C) 7 years, (D) 8 years, (E) 10 years, (F) 12 years, and (G) 13 years. 74
Figure 4-3. Proportion of obesity per PGS quintile at (A) 4 years, (B) 6 years, (C) 7 years, (D) 8 years, (E) 10 years, (F) 12 years, and (G) 13 years. 75
Figure 4-4. Obesity odds ratio for adult and child PGS, 4 to 13 years old. 76
Figure 5-1. Latent growth curve mediation model. 108
Figure 5-2. Latent growth curve mediation model, for fussy eating. 109
Figure 5-3. Latent growth curve mediation model, for over-eating. 110
Figure 6-1. Predicted trajectories of BMI throughout childhood by (A) the child-derived PGS, and (B) the adult-derived PGS. 141

Supplementary Figures:

Supplementary Figure 4-1. AUC plot, logistic regression, QLSCD 4-13 years. 82
Supplementary Figure 4-2. Calibration plot, logistic regression, QLSCD 4-13 years. 83
Supplementary Figure 4-3. Linear association between PGS and BMI z-scores adjusted for covariates, 4 to 13 years old..... 84
Supplementary Figure 4-4. Obesity odds ratio for adult and child PGS adjusted for covariates, 4 to 13 years old. 85
Supplementary Figure 5-1. Flow chart of study participants from the Québec Longitudinal Study of Child Development (QLSCD) participants. 117
Supplementary Figure 6-1. Flow chart depicting QLSCD study sample.....144

LIST OF APPENDICES

Appendix 1: Copy of the ethics certificate of approval.

Appendix 2: Characteristics of QLSCD participants with BMI data available at seven waves of data collection.

Appendix 3: Proof of submission email for manuscript 1, Canadian Journal of Physiology and Pharmacology.

Appendix 4: PRS-RS reporting checklist for Manuscript 1.

Appendix 5: Published manuscript 2

Appendix 6: Proof of authorization to reuse manuscript 2 in the thesis.

Appendix 7: STROBE reporting checklist for Manuscript 2.

Appendix 8: Proof of submission email for manuscript 3, Journal of Nutritional Science.

Appendix 9: STROBE reporting checklist for Manuscript 3.

TABLES OF CONTENTS

PREFACE ii

Ethical disclosures ii

Author contributions ii

ABSTRACT iv

ABBREVIATIONS v

ACKNOWLEDGEMENTS vi

FUNDING vii

LIST OF TABLES viii

LIST OF FIGURES x

LIST OF APPENDICES xi

TABLES OF CONTENTS xii

CHAPTER 1 - INTRODUCTION 1

1.1 Problem statement 1

1.2 Thesis objectives 3

CHAPTER 2 - LITERATURE REVIEW 5

2.1 Obesity, a health problem with public health ramifications 5

2.1.1 Defining and measuring obesity 5

2.1.2 Health consequences of obesity 7

2.2 The genetic contribution to obesity 9

2.2.1 Heritability of obesity 9

2.2.2 Candidate gene approach to obesity genetics 11

2.2.3 Molecular genetic approach to obesity genetics 12

2.3 Using polygenic scores to measure genetic susceptibility to obesity 14

2.3.1 Defining polygenic scores 14

2.3.2 Polygenic scores of body mass index in adulthood 15
2.3.3 Polygenic scores of body mass index in childhood 16
2.4 Interplay between dietary habits and genetic susceptibility to obesity 19
2.4.1 Conceptual framework: the behavioural susceptibility theory 19
2.4.2 Evidence supporting the behavioural susceptibility theory 20
2.5 Eating behaviours and their role in translating genetic susceptibility to obesity 22
2.5.1 Measuring eating behaviours 23
2.5.2 Empirical evidence: eating behaviours mediating the effect of a PGS on BMI 24
2.6 Interaction between dietary intakes and genetic susceptibility to obesity 27
2.6.1 Empirical evidence: Interaction of PGS and dietary scores in relation to BMI 28
2.6.2 Empirical evidence: Interaction of PGS and food intake in relation to BMI 29
2.6.3 Empirical evidence: Interaction of PGS and macronutrient intake in relation to BMI 31
CHAPTER 3 - METHODS 33
3.1 The Quebec Longitudinal Study of Child Development (QLSCD) 34
3.1.1 General description 34
3.1.2 Data collection 35
3.2 Outcome variables: anthropometric measures 35
3.3 Genetic variables: polygenic score construction 36
3.3.1 General construction process of a polygenic score 37
3.3.2 Target data and genetic quality control 37
3.3.3 The importance of the choice of genome-wide association study summary statistics for polygenic score construction 38
3.3.4 Description of the genome-wide association study summary statistics chosen 40
3.3.5 PRS-CS procedure 40
3.4 Dietary habit variables 42
3.4.1 Eating behaviours 42

3.4.2 Dietary intakes 45
3.5 Data analyses 48
3.5.1 Objective 1 48
3.5.2 Objective 2 49
3.5.3 Objective 3 50
CHAPTER 4 - MANUSCRIPT 1 51
Article preface 51
Title page 52
Abstract 53
Introduction 54
Materials and methods 56
Results 59
Discussion 62
References 68
Tables and figures 72
Supplementary materials 77
CHAPTER 5 - MANUSCRIPT 2 87
Article preface 87
Title page 88
Abstract 90
Introduction 91
Methods 93
Results 98
Discussion 100
Conclusion 103
Tables and Figures 104

References 112
Supplementary materials 117
CHAPTER 6 - MANUSCRIPT 3 120
Article preface....120
Title page....121
Abstract....122
Introduction....123
Methods 125
Results 129
Discussion....131
Supplementary materials 144
References 154
CHAPTER 7 - DISCUSSION 159
7.1 Summary of main findings 159
7.1.1 Manuscript 1....159
7.1.2 Manuscript 2....160
7.1.3 Manuscript 3....161
7.2 Contributions to public health and clinical care....162
7.2.1 Identify targets for obesity prevention and management....162
7.2.2 Reinforcing the conceptual framework of obesity 164
7.2.2 Usefulness of PGS for clinical care....166
7.3 Thesis strengths and limitations....167
7.3.1 Strengths....167
7.3.2 Limitations 168
7.4 Suggestions for future research....171
7.5 Conclusions....171

BIBLIOGRAPHY 173

APPENDICES 187

Appendix 1. Copy of the ethics certificate of approval. 187

Appendix 2. Characteristics of QLSCD participants with BMI data available at seven waves of data collection. 189

Appendix 3. Proof of submission email for manuscript 1, Canadian Journal of Physiology and Pharmacology. 190

Appendix 4: PRS-RS reporting checklist for Manuscript 1. 192

Appendix 5: Published manuscript 2 197

Appendix 6: Proof of authorization to reuse manuscript 2 in the thesis. 209

Appendix 7: STROBE reporting checklist for Manuscript 2. 215

Appendix 8: Proof of submission email for manuscript 3, Journal of Nutritional Science. 218

Appendix 9: STROBE reporting checklist for Manuscript 3. 219

CHAPTER 1 - INTRODUCTION

1.1 Problem statement

In recent years, few diseases drew the attention of epidemiologists more than obesity. The World Health Organization declared a global epidemic of obesity in 1997, due to its dual burden of mortality and morbidity, combined with a prevalence that is increasing at an alarming rate [1]. Obesity can be described as an excess accumulation of fat in the body that results in negative health outcomes. Body mass index (BMI) is the most used method to measure excess adiposity in observational studies due to its ease of implementation despite limitations regarding the location of body fat and the differentiation of lean and fat mass [1]. According to the most recent data available, high BMI was responsible for 3.7 million deaths, and the loss of 129 million disability-adjusted life years worldwide in 2021 [2]. The high health burden of obesity is related to the consequences of excess fat on cardiometabolic health and the increased risk of several health problems such as Type-2 diabetes, hypertension, cancer, coronary heart disease, sleep apnea, or mental health issues [1].

The prevalence of obesity and the mean BMI level escalated dangerously worldwide in the last three decades in both adults and children [3, 4]. The roots of obesity can be tracked down to early development [5]. Even though recent data suggest that the rise in overweight and obesity in childhood has tapered off in recent years [6, 7], levels of excess weight are still at distressing heights. For instance, more than a quarter of children were affected by overweight or obesity in Canada in 2019 [8]. This is concerning since childhood comorbidities are common and can lead to serious long-term health complications [9]. Obesity prevention efforts should ideally start at a young age considering most cases of childhood obesity persist into adulthood [10], and difficulties in reversing pediatric obesity in the long-term [9].

Uncovering the factors related to the occurrence of diseases is a crucial endeavour of epidemiologists, especially in the context of a major public health issue like obesity, so that prevention measures can be aimed at relevant targets and tailored to the appropriate individuals at risk. Accordingly, major efforts have been made to uncover the determinants related to the rapid increase in the prevalence of obesity. Exposure to food environments encouraging unhealthy dietary habits was identified as a major

contributor driving the global obesity increase [11]. More specifically, the global food system has evolved in a way that favorize the availability of energy-dense and highly processed foods due to improved distribution systems and widespread marketing strategies [11]. This so-called obesogenic environment leads to obesity by abetting individual behaviours ushering in an unbalance between energy intake and energy expenditure.

However, not everyone is similarly affected by exposure to an obesogenic environment. The complex nature of obesity, characterized by the interplay between several environmental and genetic factors, contributes to this observation. Obesity is strongly heritable, with genetic factors accounting for up to 80% of the variation in BMI [12, 13]. Obesity is polygenic, which means that it consists of hundreds of individual variants each contributing small increments to heritability [14]. Polygenic scores (PGS) are often used to assess the genetic susceptibility to obesity in individuals by aggregating the effects of those variants into a single index. Although there has been a surge of PGS for BMI applied to adult samples [15], only a few studies examined their use in children. Moreover, the few studies with child samples employed PGS derived from genome-wide association studies (GWAS) summary statistics obtained from adults due to lower sample sizes available for children. However, the strength of the associations between variants and BMI may change over time [16], as the effect of some genetic variants may emerge throughout development [17]. This highlights the need for investigations comparing the use of PGS developed using GWAS summary statistics obtained from adult and child populations.

PGS can be leveraged to broaden the understanding of the mechanisms underpinning genetic susceptibility to obesity. A common theory describes appetite regulation as a key component translating the influence of genetic factors on obesity [18]. Within this framework, exposure to an obesogenic environment facilitating access to energy-dense foods leads to variations in body weight due to genetic differences in how individuals respond to such an environment [18]. There is empirical evidence to support this theoretical framework. For instance, the expression of a large proportion of genetic variants associated with BMI variation is enriched in brain regions related to addiction, reward, and appetite regulation in adults [19, 20] and children [21]. Additionally, genetic variation for BMI is higher in individuals living in countries characterized by more obesogenic environments [12], and in children living in homes characterized by higher obesogenic risk [22]. Those observations are consistent with the view that genetic factors interact with the environment to create an obesity risk, and that appetite-related traits play a central role in that process.

However, evidence is lacking regarding which appetite-related behaviours translate genetic susceptibility to obesity. Furthermore, it is not clear if specific dietary habits interact with genetic factors in relation to individual BMI. Mediation and interaction analyses can be leveraged to study those specific aspects of gene-environmental interplay. Mediation refers to how the effect of an exposure on an outcome can be mediated through a third variable [23]. In contrast, interaction happens when the effect of an exposure on the outcome differs depending on the value of a third variable [23].

In recent years there has been the emergence of studies using PGS to examine mediation and interaction in the context of genetic susceptibility and dietary habits. Indeed, some studies observed that appetite-related behaviours (e.g., disinhibition, uncontrolled eating, emotional eating and susceptibility to hunger) mediate a significant portion of the association between PGS and BMI [24, 25]. Other studies found interactions between PGS composed of variants associated with BMI and dietary intake (e.g., sugar-sweetened beverages, fried foods) in relation to BMI [25, 26]. However, most of the evidence was gathered using cross-sectional designs with adult study samples. Those features hinder causal inference and limit generalizability of child populations. This highlights the need for studies investigating these aspects of gene-environmental interplay using longitudinal designs in samples of children.

1.2 Thesis objectives

The main **research goal** of the proposed thesis is to document the interplay between genetic factors and dietary habits in relation to BMI by examining the extent to which eating behavioural traits translate (i.e., mediate) genetic susceptibility to obesity and to identify putative dietary intakes that interact with genetic susceptibility to predict obesity in children and young teens. To answer this question, we propose a secondary analysis of data originating from a cohort followed since birth in 1997-1998 (Quebec, Canada). It is unclear how BMI-PGS constructed using adult and children's genetic data compare across childhood and teenage years. Therefore, two BMI-PGS have been developed to estimate genetic exposition, one from adult summary statistics and the other from child summary statistics, before tackling eating behaviours and dietary intake traits. The main research goal can be further detailed into three specific objectives:

1) To measure the genetic susceptibility to obesity in children using two PGS derived from summary statistics from adults and children GWAS, respectively, and compare their association with BMI and discrimination of obesity.
2) To determine the extent to which the association between adult- and child-derived PGS and BMI is mediated through eating behavioural traits (over-eating, fussy eating) assessed in early childhood (i.e., preschool age).
3) To estimate the association between preschool dietary (food and macronutrient) intakes and BMI (mean and change with age) displayed in school-aged children and identify the ability of adult- and child-derived PGS to modify those associations.

The thesis is organized in seven chapters, including the present introduction. Chapter 2 presents the literature review, which includes key concepts relating to obesity, the theoretical framework, and a summary of existing evidence relating to the thesis objectives. Chapter 3 contains the methods used throughout the thesis in more detail than discussed in the three scientific articles included in the thesis. Chapters 4, 5, and 6 present the three manuscripts serving as the results section. Finally, chapter 7 presents a discussion based on the integration of the results with their specific contributions and a look ahead with suggestions for future research.

CHAPTER 2 - LITERATURE REVIEW

The following literature review first describes concepts related to obesity, the trait of interest of the thesis. We offer an overview of the conceptualization of obesity and how it can be measured in practice, followed by its health consequences. The contribution of genetic factors on variations in body weight is then discussed through the lens of twin studies and molecular genetic analyses. We also introduce PGS and the current state of the literature for scores predicting BMI in both adults and children. We then present an overview of the interplay between dietary factors and genetic susceptibility to obesity by highlighting the theoretical constructs underpinning this process. Finally, we look over the current state of the evidence and highlight limitations regarding the role of eating behaviours in mediating the effect of PGS on BMI and the interaction between dietary intakes and PGS in relation to BMI.

2.1 Obesity, a health problem with public health ramifications

2.1.1 Defining and measuring obesity

Obesity can be defined as the accumulation of body fat to a level that is detrimental to one's health [1]. At the individual level, the accumulation of fat is caused by a disbalance between energy intakes and energy outputs. Energy within the body is carried by the molecule adenosine triphosphate, which is synthesized using the macronutrients (carbohydrates, proteins, lipids) ingested as food. Variations in energy intake between individuals arise from differences in the quality and quantity of foods consumed, which changes depending on factors including age, sex, energy expenditure, genetics, eating behaviours, and accessibility [27].

Energy outputs correspond to the resting metabolism, dietary thermogenesis, and the energy expended due to physical exertion [28]. For the majority of individuals, the basal metabolism is responsible for the most energy output, but can vary between individuals due to age, sex, genetics, and body

composition with lean mass leading to higher energy consumption [28, 29]. The most modifiable aspect of energy output is physical activity, which explains why the proportion of total energy expenditure due to physical activity can vary from 15% to 50% [28]. Overall, when the energy consumed in the form of macronutrients exceeds the energy necessary for body functions, the excess energy is primarily stored in the form of triglycerides in adipose tissue [27].

Obesity is a complex chronic disease characterized by the interplay between multiple demographic (e.g., age, sex), environmental (e.g., food accessibility, transportation systems), behavioural, and socioeconomic determinants (e.g., income, education), in addition to genetics. The genetic contribution to obesity can be classified as either monogenic or polygenic. The monogenic form of obesity is usually more severe, develops early in life, and is characterized by a low environmental contribution offset by a large genetic contribution [15]. Genes affecting monogenic obesity are typically rare (minor allele frequency ≤ 1%) with large effect sizes and high penetrance (e.g., LEP, POMC, MC4R, and ADCY3) [15]. In comparison, hundreds of common genetic variants (minor allele frequency > 1%) with small individual effects are involved in the polygenic form of obesity [15]. There is a strong interplay between the genetic contribution of polygenic obesity and environmental factors. Although the two forms of genetic risk for obesity have distinct characteristics, both monogenic and polygenic obesity may be underpinned by similar biological mechanisms. The central nervous system is thought to have an overlapping role in the development of monogenic and polygenic forms of obesity through impairments in appetite and reward-related pathways [15, 30]. Monogenic obesity is much rarer than polygenic obesity, which is the form the most encountered. Considering the interest in the interplay between genetic and dietary factors in this thesis, the following sections relate to polygenic obesity.

BMI (calculated as weight in kilograms divided by the square of height in metres) is the most frequently used measure of adiposity in observational studies because of its ease of use and low cost. However, BMI has its limitations, including the lack of distinction between lean and fat mass, and being blind as to the location of adipose tissue [31]. Fat mass accumulation is more strongly associated with nefarious health consequences compared to an excess of lean mass [32]. The distinction between visceral and subcutaneous fat is also of importance considering the accumulation of visceral fat and subcutaneous fat present different cardiometabolic risk profiles [33, 34]. This could explain the higher association with cardiovascular disease prognosis [35] and the higher risk of overall cardiovascular events [36] found for visceral fat accumulation compared to subcutaneous fat. Incidentally, other assessment techniques better at measuring abdominal fat, such as waist-to-hip ratio, waist

circumference, skinfold thickness, and costlier or less accessible techniques (computed tomography, bioelectric impedance, hydrostatic weighting) are also used to assess the excess of adiposity [25, 32, 37]. Nevertheless, because it has a high correlation with abdominal fat indicators, and has a strong association with cardiovascular risk in both adults and children [38, 39], BMI remains firmly established as the most widely investigated measure of obesity despite these constraints [25].

In Canada, the classification of obesity according to BMI in adulthood is in line with the World Health Organization (WHO). According to their BMI, individuals can be classified as being in a situation of underweight (<18.5 kg/m^2), normal weight (18.5-24.9 kg/m^2), overweight (25.0-29.9 kg/m^2), or obesity (≥30 kg/m^2) [40]. While subcategories are sometimes used for definition with higher distinction [41], these categories are the most commonly used in observational studies internationally. The classification of obesity in childhood is slightly more complex considering BMI varies according to age and sex. Therefore, the classification of obesity in childhood hinges on BMI z-scores calculated according to reference growth charts for children under 5 years [42] and children and adolescents between the ages 5 and 19 [43]. The WHO recommends cut-offs in BMI z-scores of >2 for overweight and >3 for obesity in children under 5 years [42], compared to >1 for overweight and >2 for obesity in children of 5 to 19 years [43].

2.1.2 Health consequences of obesity

The WHO declared a global epidemic of obesity in 1997, highlighting the disease as a major public health problem over the world [1]. The prevalence of obesity has increased drastically in the last few decades. According to the latest data from the NCD Risk Factor Collaboration, the prevalence of obesity increased from 8.8% and 4.8% to 18.5% and 14.0% for women and men, respectively, between 1990 and 2022 [3]. Alarmingly, almost half of the adult population of women (45.2%) and men (44.3%) is now considered at least overweight [3], an increase from 39% of the world's adult population in 2016 [4]. In Canada, 61.3% and 58.7% of women and men were affected by overweight or obesity and 36.0% and 30.5% were affected by obesity in 2022 [3]. There is robust evidence that obesity is related to increased mortality rates and a high morbidity burden among affected individuals. In 2021, obesity was among the deadliest health problems, accounting for approximately 3.7 million deaths and the loss of 129 million disability-adjusted life years, ranking sixth and seventh, among

attributable risk factors [2]. Compared to individuals with healthy weight, adults with overweight and obesity have an 11% and 44% increased risk of mortality of all causes [44]. There is additional evidence that suggests the increased risk of mortality is dose dependent on the increase in body weight, with the risk of mortality increasing from 29% to 39% for each 5 kg/m^2 of BMI across North America, Europe, and East Asia in those with a BMI of 25 kg/m^2 and over [44]. This is supported by similar results from a large-scale cohort study where among individuals 40 years and older with BMI of 25 kg/m^2 and over, a rise in BMI of 5 kg/m^2 was associated with an increase of 21% in the risk of mortality from all causes [45].

Although the association between obesity and increased mortality in the general population is well documented, an "obesity paradox" has been observed in which obesity appears to be protective among people with chronic illnesses. This has been observed in cardiovascular, diabetes, cancer, and respiratory patients [46]. Although there is substantial evidence that this paradox is a biological plausibility rather than a statistical artifact, many possible methodological explanations have been considered, such as reverse causation, the limits of BMI as a measure of obesity, the role of confounding factors, and the presence of selection bias [46, 47].

The high burden of mortality associated with obesity is mostly related to increased morbidity, especially cardiovascular diseases, Type II diabetes and kidney diseases, responsible for most of the mortality due to high BMI [48]. In addition to being more at risk of developing concomitant diseases, individuals afflicted with obesity present a higher risk of developing multimorbidity [49]. Individuals affected by obesity have more than 12 times the risk of developing more than two comorbidities compared to those of normal weight in a large European population of adults [49]. Indeed, higher BMI is related to a higher risk of several other health issues, including Type-2 diabetes [11, 50-52], various cancers (e.g., liver, ovaries, prostate, leukemia) [53-55], hypertension [11, 56], and cardiovascular diseases [57-61]. In addition to non-communicable diseases, obesity is also related to a higher risk of contracting hospital and community-acquired infections [55], such as COVID-19, where individuals contracting the disease have a higher risk of hospitalization and severe illness if they are affected by obesity [15, 62, 63].

Obesity in Canada is less prevalent in childhood and adolescence than in adulthood. Still, more than a quarter of children are in a situation of overweight or obesity [8]. Contrary to data from adults, the prevalence of overweight and obesity in children and adolescents seems to have reached a plateau in recent years [6, 7]. According to the developmental origin of health and diseases hypothesis, the source

of obesity goes back as early as prenatal development [5]. Children classified as having obesity are around five times more likely than children of a healthy weight to have obesity in adulthood, and as much as 70% of adolescents with obesity remain affected by obesity by 30 years of age [10], suggesting that obesity tracks from childhood to adulthood. This translates into an increased risk of mortality and morbidity reaching adulthood. Indeed, previous systematic review and large-scale observational studies found that children and adolescents with overweight or obesity were at higher risk of premature mortality [64] and cardiovascular-related mortality [65]. There is also substantial evidence that supports the association between childhood and adolescent obesity and an increased risk of disorders including type-2 diabetes, hypertension, dyslipidemia, ischemic heart disease, coronary heart disease, and cancer [64, 66, 67], although some highlight the uncertainty in whether the associations are dependent on adult obesity or not [66].

In addition to weight accumulation over time and the associated risks in adulthood, childhood obesity may also negatively affect health in the short term. Accumulating evidence points to an association between obesity and a rise in cardiovascular risk factors (e.g., hypertension, type II diabetes, dyslipidemia , or high cholesterol levels), obstructive sleep apnea, puberty irregularities (e.g., prepubertal acceleration, menstrual irregularities, or polycystic ovary syndrome), and psychological problems (e.g., low self-esteem, depression, or social isolation) [37]. The high burden related to childhood obesity highlights the need for early prevention centred around early life risk factors of obesity and tailored to mitigate the effects of those factors.

2.2 The genetic contribution to obesity

2.2.1 Heritability of obesity

Heritability refers to the proportion of inter-individual variation of a specific phenotype that is explained by genetic factors [68]. Narrow-sense heritability describes the additive genetic effects (i.e., the sum of the effects of the genes affecting a trait), while broad-sense heritability describes all sources of genetic variations, including dominance and gene by gene interaction effects [69]. Methods used to estimate the heritability of phenotypes can be divided into two groups: family-based designs and genomic methods. The first methods used to derive heritability estimates were family-based. Twin

studies estimate the narrow-sense heritability (additive genetic effects) by assessing and comparing phenotype variations between monozygotic twins (who share 100% of their genes) and dizygotic twins (who share on average 50% of their genes) [69]. Twin studies typically follow a model which estimates the narrow-sense heritability [69]. The model decomposes the phenotype variation into three components: additive genetic refers to variability explained by genetic effects, common familial environment describes variability due to environmental determinants shared by both twins, and unique environmental factors identify distinct effects on each twin [25].

Although there have been debates about the possible overestimation of heritability in twin studies due to the interaction between genes and environment [70], those studies are still commonly used to quantify the relative role of genetic factors in phenotype variation in a population. Two recent large-scale pooled analysis of twin studies from infancy to early adulthood [13], and throughout adulthood [12] estimated BMI heritability to vary from 40% to 80% throughout the life course. The lowest estimates of heritability were observed at around 4 or 5 years of age at 40%, which then rapidly increased until early adulthood (19 years of age) at 80%. It was hypothesized that the parallel rise in additive genetic effects and decrease in shared environmental factors could reflect the increased role of children's autonomy regarding their food selection and preferences [25]. There also seems to be sex-specific genetic factors affecting BMI, which become more prominent starting at puberty [25]. From a life course approach, results from longitudinal twin studies suggest that genetic susceptibility to obesity tracks well across life [25], but also that new variants affecting BMI may emerge after early childhood [17].

Compared to twin studies, genomic methods to estimate heritability generally rely on unrelated individuals that were genotyped [69]. Those methods include, but are not limited to, linkage disequilibrium (LD) score regression (LDSC) [71], and genomic relatedness restricted maximum likelihood (GREML) [72]. LDSC and GREML estimate the proportion of variance explained in a trait by common variants assessed in a GWAS [69]. An analysis using LDSC in GWAS summary statistics from over 1 million participants of European ancestry estimated a BMI heritability from common variants of 18.8% [73]. In comparison, a study using GREML-LDMS, a variant of GREML correcting for LD and allele frequency bias, estimated the BMI heritability from common variants to be 27% [74]. This result is not surprising considering heritability estimates for common variants obtained from GREML are consistently higher compared to LDSC [75]. The gap between the heritability estimated through twin studies and the heritability related to common variants is commonly referred to as missing

heritability, which could be explained by a variety of factors, such as the influence of rare variants, the overestimation of heritability in twin designs, gene-environment interactions, or epigenetic processes [74, 76, 77]. Notably, there is a strong relationship between the GWAS sample size and the proportion of variance explained by variants [78].

Overall, combined evidence from twin studies and genomic analyses supports the presence of a large genetic influence on the development of obesity. Despite highlighting the substantial role of genetic factors, heritability estimates, on their own, cannot identify the specific genetic variants associated with traits. To that end, genetic association studies were conducted to identify specific genetic variants affecting variation in BMI. Genetic association studies can be broadly defined as either hypothesis-driven or hypothesis-free. Those two approaches will be presented in the next two sections in the context of BMI. Considering the hypothesis-free approach is at the centre of this thesis and because it has far outpaced the hypothesis-driven approach in terms of discovery, it will be discussed in more detail.

2.2.2 Candidate gene approach to obesity genetics

One of the first approaches used to identify specific common genetic variants associated with a change in body weight was the candidate gene association approach. This type of study is hypothesis-driven in the sense that variants located in genes relevant to specific biological pathways are tested for their association with the trait of interest. Initiated in the mid-1990s for obesity traits, candidate gene association studies tested variants located in hundreds of genes thought to have a role in obesity risk based on previous biological studies in humans or animal models [15]. However, akin to the mythological figure Sisyphus rolling a massive boulder up a mountain, just for the boulder to roll down every time repeatedly, despite the years of efforts invested in candidate gene studies, only a handful of variants were ever validated for their association with obesity traits. Indeed, due to small sample sizes and their minuscule contributions, most genetic variants tested failed replication, limiting the overall impact of this type of study in the field [15]. Nevertheless, candidate gene methods can still be helpful in the present era of molecular genetics, where thousands of genetic variants are associated with specific traits to study specific pathways and more complex mechanisms governing the association between variants and traits.

2.2.3 Molecular genetic approach to obesity genetics

Considering the genetic component of obesity is mainly characterized by the interplay of thousands of variants across the genome, each with minor effects, association studies need to be properly powered to identify those small individual effects. The introduction of efficient genotyping technologies has facilitated the study of associations between genetic variants and various traits. As a result, GWAS are now common in investigating the genetic architecture of an extensive array of traits, such as differences in BMI. The first variant identified via GWAS for its association with BMI was a common variant on the *FTO* gene (rs9939609) [79], a region now well recognized for its role in obesity etiology. The variant showed relatively high effect size with homozygous carriers of the effect allele having 67% higher odds of obesity compared to homozygous non-carriers in the study [79]. This first GWAS had only ~5,000 participants, which explains why other loci with small effects on BMI were not identified. Additional variants located on genes such as TMEM18, MTCH2, MAP2K5, ADCY9 were identified in the subsequent years using gradually larger GWAS meta-analysis [80-82].

Since then, large databases and consortia have been leveraged to provide enough statistical power to appropriately identify small effects attributed to variants across millions of association tests run in parallel. For example, the latest update of the Genetic Investigation of Anthropometric Traits consortium (GIANT) (n ~ 340 000) has been analyzed on its own [19], as well as in conjunction with the UK biobank (n ~ 700 000) [14] to identify 97 and 941 variants, respectively, independently associated with BMI, collectively accounting for ~2.4% and ~6.0% of the variance of BMI when those variants are aggregated. Genetic association studies of BMI are now common – as of 2024, the GWAS Catalog includes 357 GWAS of BMI [83].

We also note that other than BMI, other obesity-related traits can be used to represent adiposity, even though these GWAS tend to have lower sample size considering the higher resources necessary for their implementation. Notably, GWAS of body fat percentage [84-86] and adipose tissue [87] have been completed and their results can hint at possible mechanisms overlooked by GWAS centred on BMI. For instance, variants identified in those GWAS, but not previous analyses of BMI have links to insulin resistance and adipose tissue differentiation, which suggest a unique contribution to the accumulation of subcutaneous body fat specifically [84, 87]. However, an important caveat of the current corpus of GWAS is that most studies published to date have been performed in samples of European ancestry [88],

a tendency that has increased in recent years with the availability of large-scale biobanks of participants of European ancestry. Results obtained from those studies may not be generalizable to other populations.

Additionally, most of GWAS published have been conducted in adult populations, with only a few association studies looking at children and adolescents. Since some variants affecting BMI may emerge throughout the life course [17], investigating genetic variants related to BMI in childhood should be prioritized. The first efforts at replicating the associations found in adults showed that most, but not all, BMI-related variants identified in childhood and adolescents were also associated with BMI in adulthood [89]. Since most variants associated with BMI were identified in adult populations and because information on genetic susceptibility to obesity in childhood could offer hints at possible early prevention, GWAS have also been conducted directly among children. Although they are few and have smaller sample sizes than their adult counterparts, these studies also identified the variants with the largest effect sizes on BMI, and new associations showcasing possible differences in genetic susceptibility across the lifespan [21, 90-97].

The largest GWAS meta-analysis in children to date included 41 studies for a total of 61,111 children aged between 2 and 10 years old [21]. The study identified 25 variants significantly associated with BMI; only two of these were not previously known as BMI-related variants, which suggested that the biological underpinnings of childhood genetic obesity are largely similar to that in adulthood, and that genetic factors may contribute in large part to the tracking of obesity from childhood to adulthood. This aligns with the observation that the 25 variants associated with BMI were differentially expressed in the brain [21], similar to a consistent observation in adult GWAS [19]. The study also observed a strong genetic correlation between childhood BMI and other anthropometric traits (BMI in adulthood, waist-to-hip ratio, and body fat percentage), and to a lesser extent with cardiometabolic traits (type-2 diabetes, diastolic blood pressure, and coronary artery disease), which may contribute to the association observed between childhood BMI and nefarious cardiometabolic outcomes later in life [21]. Individually, the hundreds of BMI-related variants identified throughout the years do not have substantial effects on BMI. However, taken together, they could substantially predict the BMI of individuals, which led to the representation of genetic susceptibility as the amalgam of effects of many variants. For example, a score including all 25 variants was found to explain 3.6% of the variance in BMI in a cohort of 1169 children aged seven years old [21].

2.3 Using polygenic scores to measure genetic susceptibility to obesity

2.3.1 Defining polygenic scores

In the past decade, thousands of genetics variants have been identified to be significantly associated with a wide range of complex diseases by genome-wide association studies [98]. Since the individual effects of those common variants are small in magnitude, PGS have been introduced to represent genetic susceptibility to a trait by aggregating the effect of multiple genetic variants. Those scores provide more power than assessing exposition to genetic variants individually and for some diseases the risk conferred by being in the upper tail of its distribution can be compared to established clinical risk factors [98]. Contrary to studies relying on genetic indices within a population, such as twin designs, PGS can infer genetic risk of individuals. PGS are now a common tool to study genetic susceptibility with almost 700 publications reported in the PGS Catalog [99] as of December 2024.

PGS can be leveraged for various purposes that can be sorted into two categories: 1) knowledge-driven utility, which refers to the acquisition of a better understanding of the biological underpinnings of the development of a disease, and 2) clinical utility, which describes how genetic risk can be used to achieve better clinical outcomes. Examples of the application of PGS related to acquiring better knowledge of a specific disease of trait include the study of gene-environment interactions, the investigation of mechanisms connecting genetic susceptibility to variation in the trait of interest, or Mendelian randomization studies [100]. Most complex diseases are characterized by the interplay between genetic and environmental factors [101]. For example, the increase in BMI associated with higher genetic risk is mitigated in individuals reporting higher physical activity and lower television watching [102]. Additionally, PGS can be used to represent genetic susceptibility in individuals and help distinguish between different forms of pleiotropy (i.e., genetic factors associated simultaneously with more than one trait). Biological pleiotropy describes the phenomenon where genetic factors are associated with multiple traits independently, whereas mediated pleiotropy refers to mechanisms in which a trait is an intermediary between the association of genetic variants and another trait [103]. Although other types of investigations, such as twin designs, can share evidence that traits are genetically correlated, PGS are unique in that they can ascertain the specific type of pleiotropy.

There is potential for PGS to be successfully used in the course of disease and in multiple aspects of clinical care, such as disease risk prediction, diagnosis, treatment decision-making, or prognosis [101]. The clinical implementation of PGS starts with the validation in large samples using statistical genetics and epidemiological methods [98]. As highlighted recently [98], this phase has grown more robust in recent years. First with significant methodological advances (PGS construction methods) [104] and the implementation of large-scale data banks [105], and second with the development of the PGS Catalog [99], an open resource of published PGS, and the Polygenic Risk Score Reporting Standards [106] providing a more standardized approach to reporting results from PGS analyses. However, the translation of PGS to the clinical utility of PGS has several limitations regarding the designing of valid pipelines to calculate and synthesize the PGS information for individuals, and the physician-patient management decisions [98]. For instance, PGS for complex diseases, including obesity, do not currently have the predictive ability to have a meaningful clinical role in identifying individuals at higher disease risk individually [101]. There are also concerns that the over-representation of PGS for European ancestry may exacerbate health disparities [88], and that using genetic information to incite behaviour changes in patients may not be effective [107].

2.3.2 Polygenic scores of body mass index in adulthood

There is already an extensive repertoire of PGS predicting either continuous body weight variables or the risk of obesity in adults. However, the literature surrounding such scores in children and adolescents is less prominent. Therefore, the present section will offer a brief overview of PGS of obesity in adults before offering a more exhaustive look at the existing scores in children and adolescents. The first attempts to construct genetic scores aggregated the individual effects of a few variants identified through GWAS by simply summing the risk alleles into a score. Those scores often counted only a dozen variants associated with BMI and explained less than 1% and the variation in BMI [80, 108]. This is considerably low, considering the proportion of variation explained by common variants was recently estimated at 23.4% [71]. In 2017, a review of the literature relating to genetic scores composed of BMI variants identified through GWAS identified seven such studies [109]. Although the largest PGS counted 97 individual variants, the score only explained 2.7% of the variance of BMI [19], reflecting the limited success of PGS predicting obesity-related traits.

Since then, developments in molecular genetics and the creation of large biobanks have allowed the completion of several large-scale GWAS, such as the GIANT meta-analysis [19] which was then amplified with UK biobank data [14], and the identification of more than 1000 independent variants associated with BMI-related traits [15]. Already 91 scores have been registered in the PGS Catalog [99] in December 2024. Those PGS have been reported or evaluated in a total of 29 different studies published from 2017 to 2024. The GIANT meta-analysis of 238 944 adults [19], data from the UK biobank of 453 169 adults [110], and the combination of GIANT and UK biobank data of 681 275 adults [14] were used to construct most of the PGS. Noticeably, the more recent scores [111-116] tended to not focus exclusively on the prediction of BMI, but were rather more methodological studies with various objectives, such as improving prediction in admixed populations, or comparing different PGS construction methods with various traits.

Due to different procedures used to construct each PGS, the number of variants retained in each score varies from 14 to over 7 million. Most PGS were developed using a variation of pruning and thresholding methods, which selected genome-wide significant variants, resulting in scores with hundreds of variants. A few other scores opted to use more recent statistical shrinkage techniques, such as LDpred [117], lassosum [118], PRS-CS [119], or SBayesR [120], resulting in scores with a high number of variants. When reported, the proportion of variance explained by the PGS was generally higher in PGS with a higher number of variants using more recent construction methods. For example, a study [121] built a repository of PGS for 27 exposures (including BMI) using four different PGS construction methods. The scores achieving the highest proportion of variance explained (9.7% to 10.3%) were those which included just over 1 million variants using PRS-CS and lassosum, while other scores achieved proportions below 8.1% [99]. This aligns with two other PGS developed using LDpred including approximately 2 million variants each, which showed a proportion of variance explained of 8.5% [16] and 13.1% [122].

2.3.3 Polygenic scores of body mass index in childhood

Contrary to the effervescence of PGS of BMI developed for adult samples, scores involving children and adolescents are scarce. Studies involving a PGS predicting BMI in the context of children and adolescent populations range from a few thousand participants [123-128] to more than 300 000

individuals [16]. Some studies used prospective designs, which allowed for the investigation of the effect of PGS on changes in BMI, or the difference in the strength of the associations throughout childhood. Two studies used both summary statistics from adults and children to compare their association with BMI [128, 129]. The other studies used exclusively GWAS summary statistics from adult samples. More specifically, data from the GIANT meta-analysis [19] and a meta-analysis of data from the GIANT consortium and the UK biobank [14] are commonly used for this purpose.

Studies differed in one major way: whether they used pruning and thresholding, or statistical shrinkage techniques to develop the PGS. Most of the studies constructed a PGS based on pruning and thresholding to include only variants with independent associations with BMI above a specific threshold of statistical significance. A prospective study [126] where 1037 individuals were followed from birth to 38 years old constructed a PGS including 29 variants associated with BMI from the GIANT consortium [81]. The increase in one standard deviation (SD) in the PGS was significantly associated with both higher mean BMI and higher BMI growth throughout childhood (3 to 13 years). Notably, weight gain between birth and the age of 3 mediated the association between the PGS and obesity risk in adulthood [126]. The scores built in Coleman et al. [123] and Sanz-de-Galdeano et al. [127] used GWAS summary statistics from the GIANT meta-analysis [19], albeit with different statistical significance thresholds of 0.0032 and 1×10^{-8} resulting in scores of 2321 and 97 individual variants, respectively. The PGS developed in Coleman et al. explained 4.7% of the variance in BMI in 3414 children of 11 years old and the score was significantly associated with an increase of 0.21 log(BMI) ($p = 1.59 \times 10^{-37}$) and an increase of BMI from 11 to 16 years old ($\beta = 0.09$, $p = 4.96 \times 10^{-6}$). In comparison, the proportion of variance of BMI explained by the PGS developed by Sanz-de-Galdeano et al. increased slightly from 4.2% at 15.4 years to 4.9% at 21.7 years in 2730 individuals. One other prospective study [124] constructed a PGS of 941 variants based on a threshold of 1×10^{-8} from the GWAS summary statistics of the meta-analysis of the GIANT consortium and the UK biobank [14].The PGS explained 6.5% of the variance of BMI in 1289 individuals at 16 years and, like previous studies, it remained relatively stable between 5.8% and 6.6% from 11 to 22 years old. Individuals in the highest decile of genetic risk had more than a six-fold increase in the odds of having overweight or obesity compared to those in the lowest decile at 16 years (OR = 6.41, 95% CI: 2.95–15.56, $p < 0.001$).

In contrast to the pruning and thresholding methods used in the studies presented above, Khera et al. [16] and Hüls et al. [125] both used the same PGS opting for the statistical shrinkage technique LDpred

[117]. The score, referred to here as PGS-Khera, included over 2.1 million variants weighted with the effects from the GIANT meta-analysis [19]. In 7861 children followed from birth to 18 years old in Khera et al., the weight difference between the bottom and top decile increased with age. At birth that difference was only 0.06 kg ($p = 0.02$), while it grew to 3.5 kg ($p < 0.0001$) at 8 years old and 12.3 kg at 18 years ($p < 0.0001$). Overall, PGS-Khera explained 11% of the variance of BMI in 3098 children aged 2 to 16 years in a second study [125]. The proportion explained increased throughout childhood from 2% at 2 years to 18% at 14 years. Furthermore, an increase in one SD of the PGS-Khera was associated with an increase of 0.33 kg/m^2 (95% CI: 0.30, 0.37), and individuals with PGS-Khera within the top decile had 3.63 times higher odds (95% CI: 2.57, 5.14) of being classified with obesity compared to those in other deciles. The results from these studies suggest that PGS composed of a high number of variants can explain a larger proportion of variance compared to PGS including a limited number of variants based on specific thresholds.

Supporting those results, a study [128] constructed four different PGS derived from GWAS summary statistics from a large meta-analysis [14]. The different scores included 144,588 variants, 13,137 variants, and 1,610 variants (significance threshold = none, $p<0.001$, and $p<1\times10^{-8}$), in addition to the same PGS-Khera composed of over 2 million variants used in previously discussed studies [16]. The proportion of variance in BMI z-scores explained by each score gradually increased from 2 to 12 years old in a sample of 9254 children. The highest estimates were observed at 11-12 years old with 9%, 7%, 9%, and 12% of variance explained for the $p<0.001$, $p<1\times10^{-8}$, no threshold, and PGS-Khera, respectively [128]. This aligns with the studies presented above showing that PGS developed using statistical shrinkage methods tend to have a better ability to predict BMI in childhood, compared to PGS using pruning and thresholding.

Finally, studies have used PGS derived from summary statistics from both adult and child data to investigate if those scores predicted BMI in a similar fashion when the target sample is composed of children. A secondary analysis from a recent study [129] observed that PGS derived from GWAS summary statistics in adult [14] and child populations [97] behaved differently in 28,681 Norwegian children from birth to 8 years old at 12 time points. Both scores showed the same overall pattern where the initial BMI z-score discrepancy between infants at high (tenth decile) vs. low genetic risk (first decile) at birth was small and then slowly increased with age. The main difference is that while there is a rise in the discrepancy in BMI z-score between the tenth and the first decile of both PGS starting at 3 years, the growth in discrepancy was more pronounced for the adult-derived PGS [129]. For

comparison, the difference in the BMI z-score between the tenth and the first decile of the adult-derived PGS increased from 0.36 at 3 years to 0.80 at 8 years, while the difference increased from 0.31 to only 0.41 for the child-derived PGS [129]. These results suggest that genetic factors affecting adult adiposity already characterize the genetic architecture of obesity as early as 3 years old. It is important to note that the GWAS used for the adult-derived PGS was composed of more than 10 times the sample size of the GWAS used for the child-derived PGS, which could lead to higher precision of effect size detected in adult vs. child studies. This aligns with similar results reported by a study [128] described earlier where BMI z-scores of 9254 children were examined from 2 to 12 years in relation to PGS derived using different construction methods. In addition to the four scores described in the previous section obtained from adult GWAS summary statistics, the study also examined a score based on summary statistics from children [97]. However, the child-derived PGS only explained 1-2% of variance in BMI z-scores from 2 to 12 years, a lower estimate compared to its adult-derived counterparts, which could achieve up to 7-12% explained BMI variance [128].

Overall, we highlight three main takeaways from studies predicting BMI in children target samples using PGS: 1) that the association between PGS and BMI increases throughout childhood , 2) that PGS construction methods using statistical shrinkage techniques allowing for the inclusion of a high number of variants achieve better predictions of BMI in childhood, compared to pruning and thresholding methods, and 3) that PGS derived from adult summary statistics tend to predict BMI with higher precision than PGS derived from child summary statistics.

2.4 Interplay between dietary habits and genetic susceptibility to obesity

2.4.1 Conceptual framework: the behavioural susceptibility theory

Although the predominance of an obesogenic environment illustrated by increased access to energy-dense food and the promotion of sedentary lifestyles is thought to be a major determinant in the increase in obesity rates in recent years [11], it does not account for variations in obesity outcomes in individuals exposed to similar environments [24]. This observation contributed to a shift in the conceptualization of obesity etiology. The central nervous system is now recognized as a centre for

energy balance regulation, with neurodevelopmental dysfunctions leading to obesity [30]. This aligns with the behavioural susceptibility theory, attributed to Professor Jane Wardle [130]. The theory posits that genetic susceptibility to obesity is translated through neurological functions regulating appetite, observed as variations in eating behaviours [131]. Thus, variations in body weight between individuals exposed to similar obesogenic environments are driven by different genetically determined eating behaviours.

The assumptions that 1) appetite-related traits mediate the association between genetic susceptibility and body weight, and that 2) there is an interaction between genetic factors and the environment in relation to body weight early in development are important premises underlying the behavioural susceptibility theory [18]. Conceptually, mediation and interaction investigations answer different but complementary questions regarding the association between exposure and outcome. Mediation is described as the "[…] phenomenon whereby a cause affects an intermediate and the change in the intermediate goes on to affect the outcome […]" (p.8) [23]. Mediation investigations aim to explain "how" a cause (an exposure in an observational setting) affects an outcome [23]. In the context of the interplay between genetic susceptibility to obesity and the environment, the cause refers to genes associated with obesity, and the intermediate, often called mediator, refers to eating behaviours. In comparison, interaction is described as the "[…] phenomenon whereby one exposure, characteristic, or a state somehow alters the effect of a different exposure, characteristic, state, or cause […]" (p.9) [23]. Interaction analyses examine questions related to "for whom" or "in what context" an exposure affects an outcome [23]. Contextualizing the interaction process around the behavioural susceptibility theory, the phenomenon refers to genetic factors altering the effect of the environment on body weight. Mediation and interaction analyses can be leveraged to add evidence supporting the behavioural susceptibility theory and gain a better understanding of the biological underpinnings of the development of obesity in early life.

2.4.2 Evidence supporting the behavioural susceptibility theory

Results from twin studies support the premise that BMI and eating behavioural traits share a common genetic architecture. Recently reviewed [18], twin studies in adults found in all [132, 133] but one [134] occasion a significant genetic correlation between BMI and appetite-related traits. Those results

were also observed in children in a study [135] of 2402 twin pairs observing genetic correlation between body weight and appetite at 3 months.

Supporting those observations, gene expression investigations have provided additional evidence that a large portion of the influence of genetic factors on body weight relies on neurological processes related to addiction and reward. Notably, the expression of genetic variants identified through large-scale GWAS has been shown to be enriched in regions of the central nervous system (hypothalamus, pituitary gland, hippocampus, and limbic system) known to be responsible for appetite and emotion regulation [19]. More specifically, a recent expression investigation of various variants identified through previous GWAS identified the substantia nigra and insula as key regions of the central nervous system where gene expression was the most enriched [20]. Furthermore, genetic colocalization revealed 60 brain proteins whose concentration could be linked to genetic variants associated with BMI [136]. Those proteins were in brain regions related to reward sensation associated with dietary intake and the regulation of appetite and satiety. As the authors of the investigation pointed out, a few of those proteins could be linked directly to cortex gene expression levels [136], suggesting that investigating protein concentrations could reveal mechanisms beyond those identified through gene expression. Overall, the current evidence clearly supports the presence of a shared genetic influence between BMI and appetite.

The presence of gene-environment interplay in relation to obesity was initially supported by twin designs examining differences in responses to interventions between sets of monozygotic twin pairs performed in the nineties. For example, following investigators found 3.4 and 6.8 times more variance in weight between twin pairs than between brothers and sisters following overfeeding and energy deficit protocols, respectively [137]. This strongly suggested that genetic factors could affect individual responses to environmental stimuli. Twin studies can further provide evidence supporting the presence of an interaction between genetic susceptibility and an obesogenic environment in relation to body weight at the population level by examining genetic variation in different contexts. Socioeconomic factors can be an indicator of exposure to obesogenic environments since it can affect access to a healthier choice of foods or lead away from or toward behavioural choices influencing the development of obesity. For example, Canadians living in neighbourhoods characterized by higher unemployment, lower income, and with a higher percentage of renters are more likely to be exposed to fast-food establishments [138, 139], which can lead to higher observed consumption of fast food in these groups

[140]. Additionally, Canadian children attending schools considered to be of low socioeconomic status have access to almost double the density in fast-food restaurants [141].

LLewellyn and colleagues [18] recently reviewed evidence showing that adults and children living in situations of lower socioeconomic quality generally have higher genetic variation in BMI. Notably, the genetic variation of BMI is higher in adults living in North America and Australia vs. East Asia [12], living in countries with higher gross domestic product [142], and characterized by lower educational attainment [143, 144]. Similarly, the genetic variation of BMI is higher in children living in North America and Australia vs. East Asia [13], with parents with lower educational attainment [145], and living in high-risk obesogenic homes [22]. This evidence aligns with observations from large-scale twin studies showing that variation in BMI explained by genetic factors changes drastically across the life course, especially in childhood [13]. Socially, children usually gain independence as they age, which increases their vulnerability to the environment and enhances their susceptibility to their genetic profile [131]. Overall, the results from twin studies presented above support the presence of gene-environment interaction in relation to BMI at the population level. Nonetheless, analyses at the individual level would strengthen the evidence. For instance, the identification of a significant interaction between a PGS and other risk factors (socioeconomic status [102], and consumption of western-styled diet [146]) in relation to BMI solidify the corresponding results from twin studies.

2.5 Eating behaviours and their role in translating genetic susceptibility to obesity

The previous section established the conceptual framework surrounding the interplay between genetic factors and the environment in relation to obesity and laid out existing evidence supporting the behavioural susceptibility theory. While we have also presented evidence supporting the importance of the central nervous system, and more specifically appetite, in translating genetic susceptibility to obesity, a next step would be to examine if this can be observed at the individual level, and to determine what observable eating behavioural traits can mediate the effects of genetic factors on body weight. As demonstrated in section 2.3, PGS can be used to quantify genetic susceptibility directly in individuals. Therefore, this section describes the most common measurement tools for eating behaviours and then presents the existing evidence regarding eating behaviours as putative mediators of the association between PGS and BMI.

2.5.1 Measuring eating behaviours

The Three-Factor Eating Questionnaire (TFEQ), devised by Albert Stunkard and Samuel Messick [147], is the most widely used questionnaire and evaluates three dimensions of eating behaviours (cognitive restraint, disinhibition, and susceptibility to hunger) displayed in adulthood [147]. The final questionnaire included cognitive restraint (21 items), disinhibition (16 items), and susceptibility to hunger (14 items) [147]. The cognitive restraint behaviour referred to the restriction of food intake by conscious mechanisms, while disinhibition behaviour related to the action of over-eating in response to external incentives and susceptibility to hunger describes the perceived sensation of hunger [147].

Different versions of the questionnaire were created by other investigators with validations in other populations. For example, the TFEQ-R18 is a reduced version of the initial TFEQ that was refined in a population of Swedish participants with obesity [148]. The resulting questionnaire was reduced to 18 total items that related to three factors. Cognitive restraint was composed of six items that were originally also part of the cognitive restraint factor of the TFEQ. Uncontrolled eating combined nine items from the disinhibition and susceptibility factors of the original TFEQ. Emotional eating was composed of three items that were originally part of the emotional subscale of disinhibition [148]. The TFEQ-R18 was later validated in normal weight populations [149] and is commonly used to assess eating behaviours in adulthood.

The Child Eating behaviour Questionnaire (CEBQ) [150], developed by Jane Wardle, is the first extensive measure characterizing the appetite of children and is meant to be administered to parents and effectively measures eight behaviours (scales) related to appetite. The questionnaire was developed using an approach where the scales of interest were chosen to represent theoretical constructs supported by the literature, compared to a purely empirical approach. Eating style constructs of interest were initially selected by examining the existing literature, including other questionnaires (mostly for adults), such as the TFEQ, which allowed the selection of six constructs of interest: satiety responsiveness, responsiveness to food cues, emotion eating, general interest in eating, speed eating, and food fussiness [150]. The constructs responsiveness to social factors, distractibility, emotional undereating, and desire for drinks were then added for a total of 10 constructs following parental interviews [150].

The questionnaire was then refined across multiple samples using principal component analysis, which resulted in a final version of the CEBQ including eight scales composed of a total of 35 individual items. The questionnaire is composed of four "positive eating response scales", namely food responsiveness, enjoyment of food, emotional over-eating, and desire to drink [150]. Similarly, satiety responsiveness, food fussiness, slowness in eating, and emotional undereating represent "negative eating response scales" [150]. A subsequent validation analysis reported that the appetite traits measures using the questionnaire persist from early to late childhood [151], and that the scales are associated with objectively measured eating behaviours [130].

2.5.2 Empirical evidence: eating behaviours mediating the effect of a PGS on BMI

Most of the evidence examining the association between a PGS and BMI in adults used the TFEQ or some of its variations (e.g., R18, or R21) to measure eating behaviours (uncontrolled eating, emotional eating, cognitive restraint, disinhibition, or susceptibility to hunger) and their role as an intermediate variable. The sample size for those studies varied from 768 [152] to 5863 [153] adult participants. Analyses of three different PGS in two different studies showed that high uncontrolled eating and high emotional eating measured using the TFEQ at least partly mediated the association between the PGS and BMI [153, 154]. Uncontrolled eating mediated respectively 6% and 12% of the effect of PGS in a French cohort (n=2154) and a United Kingdom cohort (n=3515) [154]. In comparison, the meta-analysis of two Finnish cohorts (n = 4632 and 1231) estimated that uncontrolled eating mediated 16% of the association between a PGS and BMI [153].

Similarly, both studies observed that associations between PGS and BMI were mediated by emotional eating in the magnitude of 11% and 10% in the French and United Kingdom cohorts [154], and 13% in the Finnish meta-analysis [153]. Another study [155] did not find significant mediation of emotional eating between a PGS comprising close to one million variants and BMI in a Finnish cohort (n = 1055). However, the percentage mediated (10.3%) was similar to that observed in the studies conducted on emotional eating in the cohorts from France and the United Kingdom [153, 154]. This could be explained by how the emotional eating variable was constructed via principal component analysis rather than a specific questionnaire.

Another study performed with a cohort of 768 Canadians [152] obtained similar results for disinhibition and susceptibility to hunger. In the study, disinhibition and susceptibility to hunger mediated respectively 9% and 4% of the association between a PGS and BMI. The concordant results with emotional and uncontrolled eating is not surprising considering the authors of the original article note that different versions of the TFEQ use similar items to measure uncontrolled eating, disinhibition and susceptibility to hunger, and that the subscale emotional susceptibility to disinhibition is identical to emotional eating [152]. Interestingly, the authors also noted that the subscale emotional susceptibility to disinhibition was not a significant mediator on his own [152], contrary to the results observed in the studies mentioned above. This could reflect the smaller proportion of mediation observed for emotional eating compared to uncontrolled eating, albeit this time with a lower sample size.

The only study using a prospective design alongside eating behaviours measured using the TFEQ identified that disinhibition and susceptibility to hunger mediated the association between a PGS and BMI, but that disinhibition was the main driver behind the mediation process [156]. The study observed that 34% of the association between the PGS and BMI was mediated by disinhibition over a follow-up of 20 years in 2464 adults from the United Kingdom [156]. Comparatively, susceptibility to hunger was only responsible for 10% of the association between the PGS and BMI. Those proportions are notably higher in magnitude than those observed in the study performed in Canadians. This could be explained by the higher sample size and the usage of a prospective design allowing for more precise estimations.

Cognitive restraint can also be investigated using the TFEQ. However, the association between PGS and BMI was not found to be mediated by cognitive restraint in the French and United Kingdom cohorts [154], and the Canadian cohort [152] mentioned earlier. Finally, a study of two UK cohorts of adults (n = 1,219 and 1,468) using the TFEQ-51 observed that the indirect effect of a PGS on BMI was higher through disinhibition and its subscales compared to external or internal hunger [157]. Overall, those results suggest that uncontrolled eating plays a substantial role in mediating the association between genetic susceptibility to obesity and higher body weight in adults, and that changes in appetite level in response to external cues (disinhibition) play a larger role, compared to internal cues (susceptibility to hunger). Although there is evidence that emotional eating plays a role in this mediation process, the evidence is less solid.

Few studies examined the potential role of eating behavioural traits as mediators of genetic susceptibility to obesity measured with PGS in children [158-160]. The CEBQ was used in two of those studies to specifically investigate food prone behaviours (enjoyment of food, food responsiveness, emotional over-eating), and food avoidant behaviours (slowness in eating, and satiety responsiveness). Satiety responsiveness mediated the association between a PGS and BMI z-scores in 10-year-old children (n = 2258) from the United Kingdom, with the addition of the behaviour reducing the association between the PGS and BMI z-score from 0.177 to 0.167 [158]. On the other hand, a prospective study of 995 Norwegian children observed that the five behaviours of enjoyment of food, food responsiveness, emotional over-eating, slowness in eating, and satiety responsiveness collectively did not mediate the effect of a PGS on BMI from 6 to 8 years [159]. Slowness in eating, the only behaviour found to be significantly correlated with the PGS, was not a mediator when considered individually [159]. Examining the role of appetitive traits in genetic susceptibility to higher BMI in French children followed from birth to 5 years old, another study found that those with higher PGS were more likely to have a high appetite at 2 years, as measured by a single question [160]. The association between the PGS and BMI z-score from 2 to 5 years was mediated by higher appetite at two years. Notably, the proportion mediated continuously decreased from 47% at two years, to 35% at three years, 28% at four years, and 24% at five years [160].

As presented above, there is evidence supporting the hypothesis that eating behaviours are partly mediating genetic susceptibility to obesity, but this literature still has crucial limitations. For instance, most of that evidence was observed in adult samples. Furthermore, the studies related to children observed mixed findings with appetite behaviours identified as mediators in some [158, 160], but not all instances [159]. Although there is supporting evidence suggesting that children's eating behaviours can track into adulthood [161], more studies following eating behaviours throughout life are needed to confirm this statement. In the meantime, investigators should not assume that results pertaining to eating behaviours in adulthood also apply to children. Additionally, a few studies used a prospective design. This is especially important in the context of mediation analyses where the temporality of events is crucial. Indeed, at its simplest, a mediation analysis calls for a defined sequence of events where an exposure affects a mediator, which in turn affects the outcome of interest. A prospective design ensures that these events happen in order where cross-sectional designs can fail to account for reverse causation. Finally, aside from one PGS including ~1,000,000, all other PGS that were used in the mediation analyses presented in the previous section were constructed using fewer than 100 variants. Considering obesity is a polygenic disease where hundreds of variants are known to contribute

to variations in body weight, a wide array of biological mechanisms are involved in the development of obesity. In addition to explaining less overall variation in body weight, only including a few dozen variants also risks limiting the range of possible mechanisms captured by those variants. Thus, there is an advantage to using PGS construction methods leveraging the whole genome to include more variants associated with obesity.

2.6 Interaction between dietary intakes and genetic susceptibility to obesity

As was presented in section 2.4, twin studies on the heritability of BMI support the presence of gene-environment interactions in relation to body weight on a population level. PGS can be helpful in identifying those interactions at the individual level. Several studies using PGS to measure genetic susceptibility to obesity have already observed that environmental factors (e.g., TV watching, physical activity and sleeping habits) can interact with PGS to influence BMI [24, 102]. Considering that the appetitive control pathway is a major part of the biological underpinnings of obesity [15], dietary habits are good candidates to represent a proximal dimension of the environment. Therefore, this section relates to prior investigations pertaining to the role of dietary intakes as moderators of the association between a PGS (exposure) and BMI (outcome). The term moderation is often used interchangeably with the phenomenon of interaction. In this thesis, the term moderation is used specifically to describe the process whereby the presence of an exposure modifies (or moderates) the effect of another exposure on the outcome of interest. Importantly, the presence of an interaction between two variables implies that the moderation of effect works both ways. For instance, although the moderation process of interest in this thesis relates to the PGS as a moderator of the effect of dietary intakes on BMI, some studies discussed below viewed the interaction with dietary intake as the moderator of the effect of a PGS on BMI. For this section, we categorized studies based on dietary intake variables. Dietary scores refer to any composite score that is not specific to the consumption of a specific dietary component. Food intakes relate to the consumption of specific types of food or drinks (e.g., fruits, or sugar-sweetened beverages). Macronutrient intakes refer to the total consumption of energy or macronutrients (lipids, proteins, and carbohydrates) and their subtypes.

2.6.1 Empirical evidence: Interaction of PGS and dietary scores in relation to BMI

Although dietary scores do not differentiate between specific aspects of the diet, they can be informative in aggregating the effects of multiple dietary factors into a single score and characterizing the diet as a whole. In general, there is evidence suggesting that a high genetic risk could strengthen the effect of dietary intakes on BMI, but as was recently pointed out [26], some dietary scores may be better suited to identifying interactions. Three dietary quality scores (the Alternative Healthy Eating Index, the Dietary Approach to Stop Hypertension, and the Alternate Mediterranean Diet) composed of individual items from a semiquantitative food frequency questionnaire, with higher scores defining a healthier diet, were used in two different prospective investigations with a total of 14,046 adults [162], and 31,058 adults [163]. The first study found that an increase in a diet quality score (Alternative Healthy Eating Index-2010 and Dietary Approach to Stop Hypertension, but not Alternate Mediterranean Diet) was associated with a decrease in BMI over four years [162]. The decrease in BMI was stronger in participants with higher values of the PGS, suggesting that the benefits of a healthier diet are more prominent in individuals at higher genetic risk of obesity. The reduction in BMI associated with a one SD increase in the Alternative Healthy Eating Index-2010 score was lower (-0.12 kg/m^2) in those of low genetic risk compared to those of high genetic risk (-0.18 kg/m^2). Similarly, the association between the Dietary Approach to Stop Hypertension score and a reduction in BMI was lower in those of low genetic risk (-0.14 kg/m^2) compared to those of high genetic risk (-0.19 kg/m^2) [162]. This indicates that the benefits (lower BMI) of a healthier diet are increased in individuals of high genetic risk. The other study found that the association between a PGS and change in BMI (measured every four years) was lower in participants with healthier diets from all three scores, suggesting that a healthier diet may mitigate the effect of genes on obesity [163]. However, we note that two additional studies of Chinese (n = 7817) [164] and Swiss (n = 3033) [165] adults did not find a significant interaction between the Alternative Healthy Eating Index-2010 and a PGS in relation to BMI. The mixed results for the Alternate Mediterranean Diet align with other components of the literature where some studies observed an interaction between a PGS and a Mediterranean diet in relation to adiposity measures [166], while others did not [165].

Another dietary score composed of individual items obtained from a food frequency questionnaire found evidence that the protective effect of a healthier diet is stronger in genetically susceptible individuals. The decrease in BMI associated with higher scores of a healthy plant-based diet index,

based on 17 individual food items, almost doubled from the lowest quartile of genetic risk to the highest quartile in more than 100,000 adults from the UK biobank [167]. This aligns with another study, where the inverse association between the Healthy Eating Index and BMI was found to be strengthened in individuals with higher values of a PGS in a Canadian adult study (n = 6,087) [168]. Notably, the Healthy Eating Index was found to be one of the most reliable diet scores interacting with genetic susceptibility to obesity in a recent systematic review [26]. On the other hand, the analysis of 68,317 adults of European ancestry from 18 cohorts did not find that a diet composite score could modify the association between a PGS and BMI, with the modification going in either direction [169]. Using factor analysis, another group developed four dietary scores describing different styles of diet (Western-style diet, Korean-style diet, rice-based diet, and plant-based diet) and compared the effect of a PGS on BMI according to those diet scores in a Korean hospital-based cohort of over 50,000 individuals [146]. The odds ratio of obesity for high PGS compared to low PGS significantly decreased from 1.46 to 1.39 in those of low and high plant-based diet, respectively. Similarly, the genetic association significantly increased from 1.47 to 1.62 in those of low and high fried food diet, respectively [146]. Finally, a significant interaction was observed between a PGS and eating a Mediterranean diet for an association with both BMI and waist circumference in a European sample of 605 adolescents [170], adding evidence for younger individuals that different dieting styles can influence the association between genetic susceptibility and obesity. Overall, the current body of evidence seems to largely confirm that having a healthier diet can diminish the effects of genetic susceptibility on obesity, and, conversely, that the beneficial effects on BMI of a healthier diet are accentuated in individuals at higher genetic risk of obesity. However, it remains to be determined if specific types of food drive this interaction.

2.6.2 Empirical evidence: Interaction of PGS and food intake in relation to BMI

A recent review and meta-analysis [26] of the studies examining the interaction between PGS and dietary intakes in relation to anthropometric outcomes revealed that sugar-sweetened beverage and fruit and vegetable intakes were among the dietary factors with the most evidence supporting moderating effects. A study of three large prospective cohorts from the United States (n = 33,097) found that the consumption of sugar-sweetened beverages modified the association between a PGS and BMI and the risk of obesity four years later [171]. The association between an increment of 10 alleles and BMI and

the risk of obesity increased from 1.00 and 1.19 for an intake of less than one serving per month to 1.78 and 3.16 for one or more serving per day, respectively [171]. Those results were adjusted for lifestyle factors, including total energy intake, suggesting the observed interaction is independent of energy intake. Those results aligned with a cross-sectional study of 26,726 adults of European ancestry [172]. The association between an increment of ten alleles and BMI increased from 0.83 kg/m^2 in participants with seldom to low sugar-sweetened beverage intake, compared to 1.31 kg/m^2 in participants with medium to high sugar-sweetened beverage intake [172]. These results were also adjusted for energy intake. Complementary to those results, the consumption of soft drinks was found to interact with a PGS to influence yearly changes in waist circumference in 4765 adults [173]. The increase in one allele per serving of soft drink per day was associated with an increase of 0.05 cm per year, adjusted for lifestyle factors, including total energy intake [173].

However, other studies contradicted those findings. A prospective study of 14,046 health professionals in the United States did not find a significant interaction between a PGS and servings of sugar-sweetened beverages and fruit juice per day for an association with BMI change over 4-year intervals [162]. Similarly, a cross-sectional study of 46,526 adults from the United Kingdom found that the association between a PGS and BMI was not different between participants with the consumption of 0 fizzy drinks daily and those with at least one drink daily [102]. Additionally, the quality of beverages was not found to moderate the association between a PGS and BMI in an Iranian study (n = 202) [174]. Finally, a study was performed in Finnish children (n = 1142), where the intake of various foods, including sugary juice drinks, did not interact with a PGS in relation to BMI [175].

Other types of foods have also been investigated, including fruits and vegetables, fried foods, and meats. The prospective study of health professionals presented earlier [162] also investigated fruits, vegetables, and meats. The study observed that an increase in 10 alleles per one SD in servings of fruits and vegetables per day was associated with a decrease of 0.05 and 0.04 kg/m^2 in BMI every 4 years [162] adjusting for total energy intake. On the other hand, no interaction was found for the consumption of fish and processed meats [162]. The interaction between PGS and fruit and vegetable intake was replicated by the same group of researchers a year later. The association between a 10-allele increment and BMI change every 4 years was -0.02 kg/m^2 in participants with a high increase in fruit and vegetable consumption, compared to 0.09 kg/m^2 in participants with a high decrease in consumption, adjusting for lifestyle factors, including total energy intake [176]. Finally, conflicting results have been observed for a putative interaction between fried foods and genetic susceptibility. A

prospective investigation of three cohorts (n = 37,423) observed that the association between an increment of 10 alleles and BMI and risk of obesity 3-4 years later increased from 1.10 kg/m^2 and 1.61 for an intake of less than one serving of fried food per week to 2.20 kg/m^2 and 2.72 for four or more servings per week, respectively [177]. However, a cross-sectional study of 46,526 adults found that the association between a PGS and BMI was not different between participants with a consumption of 0 meals of fried food daily and those with at least one daily meal [102].

2.6.3 Empirical evidence: Interaction of PGS and macronutrient intake in relation to BMI

Few studies have examined potential interactions between genetic susceptibility to obesity measured by PGS and macronutrient intakes. Those studies show mixed results and mostly use cross-sectional designs. Notably, a cross-sectional study of 711 adults from Spain [178] used body fat mass percentage measured using bioelectric impedance to measure body weight and to classify obesity instead of BMI. The study found a significant interaction between a PGS and proteins (total, animal, and vegetable), lipids (total, saturated fatty acids, and polyunsaturated fatty acids), carbohydrates (total and complex), and total energy intake in association with body fat percentage and obesity risk [178]. All macronutrient analyses were adjusted for total energy intake, suggesting that the interactions observed were not due to an increase in energy consumption.

However, although a higher consumption of energy, proteins, lipids, and carbohydrates were significantly associated with higher BMI and risk of obesity (not adjusted for total energy intake) in three Korean cohorts (n = 35,094), no interaction was observed with two different BMI-PGS [179]. The lack of significant interaction between genetic susceptibility measured using a PGS and protein and lipid consumption was also noted in the cross-sectional study of 46,526 adults [102] mentioned previously, although this time macronutrient intake was measured as a percentage of energy consumed. Similarly, energy intake and the proportion of proteins, carbohydrates, and lipids ingested were not found to interact with a PGS in a Chinese study (n = 7817) [164].

Finally, a higher consumption of fat (total, saturated, and monounsaturated) in relation to the total diet was observed to interact with a PGS to affect BMI, adjusting for total energy intake in a cross-sectional study combining 2,817 adults [180]. Overall, the mixed nature of the results may be explained by the low number of studies and individual differences between them. Specifically, differences in how

energy intake is taken into consideration, or demographic differences between study samples could explain how results between studies varied.

The evidence presented in the present section supports the hypothesis that the effect of dietary intakes on BMI can be moderated by individual differences in genetic susceptibility (and vice versa), but there are still important limitations. First, much of the evidence was gathered from studies using adult samples. Considering the dietary choices made in adulthood may differ from those in childhood based on differences in individual independence between those periods [25], there need to be additional studies performed specifically in childhood. Second, more studies using prospective designs are needed, as cross-sectional designs are unable to account for the possible bidirectionality [181] of the association between dietary intakes and BMI. Third, most of the PGS used in the studies presented above were constructed using pruning and thresholding methods restricting the number of variants included. Using statistical shrinkage techniques to incorporate considerably more variants may allow the observation of stronger associations of PGS with BMI and diminish the chance of missing important biological mechanisms related to genetic susceptibility to obesity.

CHAPTER 3 - METHODS

All the data used throughout this thesis came from longitudinal assessments gathered from the Quebec Longitudinal Study of Child Development (QLSCD). This section will introduce the methodological aspects necessary to examine all three of the thesis objectives. The finer points of the methods are expanded upon in the three manuscripts presented in the Results section (Chapters, 4, 5, and 6). The current section covers additional details about the recruitment and the data collection process of the QLSCD, details how the main outcome of the thesis (BMI) was devised, presents a bigger picture of the PGS construction process, introduces the eating behaviours and dietary intake variables under study, and discusses the data analysis techniques employed.

Box 1. Restatement of the three thesis objectives

Objective 1: To measure the genetic susceptibility to obesity in children using two PGS derived from adults and children GWAS summary statistics, respectively, and compare their association with BMI and discrimination of obesity.

Objective 2: To determine the extent to which the association between adult- and child-derived PGS and BMI is mediated through eating behavioural traits (over-eating, fussy eating) assessed in early childhood (i.e., preschool age).

Objective 3: To estimate the association between preschool dietary (food and nutrient) intakes and BMI (mean and change with age) displayed in school-aged children and identify the ability of adult- and child-derived PGS to modify those associations.

3.1 The Quebec Longitudinal Study of Child Development (QLSCD)

3.1.1 General description

The QLSCD is a prospective cohort study of 2120 children born in 1997-1998 in the province of Quebec, Canada. At its conception, the aim of the study was initially to document childhood development and adaptation difficulties in the first five years of life [182]. The study then continued to further investigate the relation between developmental characteristics of children or their environment and biopsychosocial outcomes from childhood to early adulthood. The study united a large number of researchers and experts from various interdisciplinary fields related to childhood development [183]. Since its creation the study has contributed to a growing library of journal articles detailing the relationship between early life characteristics (e.g., internalizing and externalizing behaviours, parental environment, or peer victimization) and later developmental outcomes (e.g., educational achievement, substance use, suicidal ideation, cognitive development or psychosocial development) [183].

Details about the QLSCD recruitment process and follow-up until age 21 years are available in a cohort profile [183], with important information summarized in the following lines. The children part of the QLSCD were born between October 1997 and July 1998. The children were randomly sampled based on living areas and birth rates based on the Master birth register to be representative of the Canadian Province of Quebec born in those years. An initial sample of 2940 families was selected based on inclusion criteria (mother's pregnancy lasted for 24 to 42 weeks and mothers could speak either French or English). From the initial sample of 2940 families, 2120 were part of the final longitudinal sample. Families from the initial sample were lost due to not being found (n = 172), being excluded (n = 93), being unreachable (n = 14), refusing to participate (n = 438), or being part of an oversampled population from Montérégie (n = 103).

3.1.2 Data collection

The participants were followed annually or biannually from ages 5 months to 25 years at the latest data collection wave in 2023. Data collection waves at 17 months, 29 months, and 41 months will henceforth be referred to as 1 year, 2 years and 3 years for simplicity. Another data collection wave happened between the ages of 45 months and 56 months and will be referred to as 4 years. The data was collected through various questionnaires. From 5 months to 17 years, parents (usually mothers) filled questionnaires by themselves or by interviewers (paper or informatic). From 7 years onward, participants also directly filled out paper or informatic questionnaires, with online questionnaires being the norm starting at age 15 years. Teachers, school administrators, brothers and sisters also answered questionnaires on occasions. A complete list of the questionnaires used in the QLSCD, and technical documentation are publicly available on the study website (www.jesuisjeserai.stat.gouv.qc.ca). The data collected varied in themes, from developmental characteristics in early childhood to academic achievement and overall health outcomes later in childhood, and transition into adulthood at the latest waves of data collection. The three objectives of this thesis relate to BMI or obesity as the outcomes of interest and involve the use of genomic data. Therefore, the following sections relate to anthropometric and genetic data collections and variable construction. Then, we detail eating behaviour and dietary intake measures, since these are the focus of the second and third thesis objectives, respectively.

3.2 Outcome variables: anthropometric measures

Height and weight were measured through different means throughout the QLSCD. From ages 5 months to 3 years height and weight were reported by the parents. From ages 4 years to 13 years trained technicians directly measured the height and weight of the participants using a standardized protocol. From 15 years onward, the participants reported their height and weight themselves. Given that self-reported measures of height and weight tend to underestimate the actual BMI [184], and that parents of children with overweight have a tendency to underestimate their children's weight [185], using BMI derived from reported height and weight measurements may lead to biased association estimates. Therefore, BMI values for this thesis were derived from direct measurements available at

ages of 4, 6, 7, 8, 10, 12, and 13 years old. BMI z-scores were calculated using the World Health Organization Growth reference data [42, 43] in order to account for developmental changes based on age and sex. Weight categories were calculated based on the World Health Organization classification for children under 5 years [186] and for children and teens between 5 and 19 years [43]. Overweight was defined as having a BMI above 2 SD for children under 5 years of age, and above 1 SD for children aged between 5 and 19 years of age. Similarly, obesity was defined as having a BMI above 3 SD for children under 5 years of age, and above 2 SD for children between 5 and 19 years of age.

Due to the longitudinal nature of the study, a portion of participants naturally drop out of the study through time. The description of participants with BMI data available at the specified collection waves is presented in Appendix 2. The number of participants with BMI data decreased from 1529 (72% of the initial QLSCD sample) on the four-year-old collection wave, to 1229 (58% of the initial QLSCD sample) at the 13-year-old collection wave. Overall, compared to all QLSCD participants, participants with BMI at later time points had higher representation of girls, had mothers with higher education, lived in households with higher income, and had mothers with a higher probability of being born in Canada. In terms of body weight, participants with BMI at later collection waves were characterized by higher BMI and higher proportions of overweight and obesity.

3.3 Genetic variables: polygenic score construction

A major part of the thesis is the construction of two PGS to represent the individual genetic susceptibility to increased BMI. Those two scores are then used to address all three of the thesis objectives and are an essential part of the three presented manuscripts. A general description of the two PGS was given in the manuscripts, but this section aims to provide a more complete picture of the construction process. More specifically, the section describes the overall process used to construct those scores, details about the genetic quality control process applied to the target sample (QLSCD participants), a description of the GWAS summary statistics chosen, a discussion about the choice of summary statistics, and an overview of the specific construction procedure (PRS-CS).

3.3.1 General construction process of a polygenic score

Construction of a PGS usually follows a basic structure, requiring two primary sources of information: GWAS summary statistics that are generally publicly available and include individual associations between variants and the trait of interest, and target data that are usually not publicly available and include individual genotype information [104]. The PGS is first constructed using the summary statistics (variants and their effect size) to identify and weight variants associated with the trait of interest to be included in the score, a process often referred to as PGS "training". Until recently, most PGS included variants pruned from LD for independence and selected those that passed a p-value threshold, generally referred to as pruning and thresholding. That method is still widely used to construct PGS across common diseases, but other procedures have recently emerged. For instance, statistical shrinkage methods adjust for the uncertainty in the estimation of variant effect size in analyses with a large number of variables to shrink small effect coefficients toward zero [187]. Statistical shrinkage often relies on Bayesian statistical approaches and includes methods like LDpred [117], SBayesR [120] or PRS-CS [119]. These methods allow the inclusion of variants from the entire genome while allowing for a wider range of genetic architectures. Those methods have been shown to outperform other procedures and work better with large sample sizes [119]. The development of the PGS then needs to consider LD by either including only independent variants, or by accounting for the correlation between variants using LD reference data for the population of interest [104]. The score is then calculated in the study (training) sample with available phenotypic and genotypic data, which is used in conjunction with demographic and other non-genetic variables to form the risk model to be fitted appropriately.

3.3.2 Target data and genetic quality control

Biological material was collected for a subsample of 992 participants when they were 10 years of age. DNA was extracted from blood or saliva samples and the genotype of 978 participants was derived using the Illumina Infinium PsychArray-24 Beadchip. There were three rounds of genetic quality control. The first round included exclusion criteria for variants (minor allele frequency below 0.01, and

genotyping rate below 0.98), and exclusion criteria for individuals (call rate below 0.95, gender mismatch, genetic duplicates, cryptic relatives, genetic outliers, and heterozygosity deviations). The second round was performed in remaining individuals only to exclude additional variants (non-autosomal variants, call rates below 98%, minor allele frequency below 5% and Hardy-Weinberg Equilibrium p-values below 1×10^{-3}). The final round of exclusion for variants occurred after the imputation process (minor allele frequency below 1%, Hardy-Weinberg Equilibrium p-values below 1×10^{-6}, and INFO metric below 0.8). Imputation aims to multiply the number of variants available by leveraging the correlation between variants. The imputation process was completed using Phase 3 of the 1000 genomes project reference data [188] and the programs SHAPEIT v2 (r837) [189] and IMPUTE2 [190]. A total of 816 participants and 8,465,216 variants, most of them imputed, passed all three steps of quality control and imputation.

3.3.3 The importance of the choice of genome-wide association study summary statistics for polygenic score construction

Choosing the most appropriate input GWAS summary statistics is crucial, as it can have substantive effects on how it affects the trait of interest. In addition to the statistical power of the discovery GWAS, the predictive power of a PGS may be reduced when there are differences between the PGS target population and the discovery GWAS population [191]. Those differences can be related to the genetic background or encompass non-genetic determinants of trait variability (e.g., sociodemographic characteristics, or phenotypic assessment). For instance, PGS are not easily transferable across ancestries due to stratifications in allele frequencies, LD, and environmental exposures [88]. For example, pairs of PGS for height have a stronger correlation when the discovery GWAS are both of European ancestry ($r \geq 0.73$), compared with pairs of European and African ancestries ($r \leq 0.31$) [192]. Generally, PGS have lower predictive ability when discovery and target samples come from different ancestries [88, 192, 193]. For example, the same PGS constructed using discovery GWAS of European ancestry explains a higher variance of BMI in a target population of European descent (5.8%), compared to African American descent (1.5%) [88].

Careful consideration must also be given to non-genetic factors that could affect phenotypic variability between the discovery and target samples. For instance, phenotypic assessment, be it the definition of

the phenotype measured, or how the phenotypes are empirically measured, should ideally be as similar as possible [191]. For example, most of the GWAS for obesity-related traits use BMI to represent the accumulation of adipose tissue. However, those PGS may not be best suited to measure other similar obesity-related traits. A PGS constructed using BMI-associated variants explained more variance of BMI (10.8%) compared to waist circumference (8.8%) in children aged 2-16 years [125]. Indeed, the specific variants associated with these adiposity measures and their individual effects may differ substantially. For instance, a PGS constructed with a GWAS for BMI had only a moderate correlation with a PGS for waist circumference (0.36 to 0.44), and a null correlation with a PGS for waist to hip ratio (0.00 to -0.04) in three independent cohorts [173].

Considering genetic factors are known to interact with environmental factors in complex diseases, differences in sociodemographic characteristics between discovery and target samples may also limit the prediction of a PGS. Most samples used in recent GWAS originate from large biobanks. Considering participants in those biobanks are volunteers, they may be generally healthier compared to some PGS target populations, such as hospital-based cohorts or samples from socioeconomically deprived communities. This could result in differences in baseline characteristics between discovery and target samples, or the presence in the target sample may be dependent on specific factors, which could hamper the resulting PGS prediction [88]. Looking at a more concrete example of specific demographic factors that can influence phenotype variability, although there is overlap in genetic variants associated with BMI in childhood and adulthood, the effect of some variants may change with age. Interestingly, two different studies showed that PGS derived from adult discovery samples explained a higher proportion of BMI variance in early childhood (birth to 8 years, and 2 to 12 years), compared to PGS derived from child discovery samples [128, 129]. While this observation could be explained by more precise effect estimates in adult discovery samples due to a larger number of participants, this could also reflect that genetic variants related to adult BMI may already emerge after the adiposity rebound around 5 or 6 years of age [129]. This last example highlights that the choice of GWAS summary statistics to use in the development of a PGS is complex and requires careful considerations in light of the target sample and the known underlying mechanisms characterizing the genetic component and the interplay with environmental factors.

3.3.4 Description of the genome-wide association study summary statistics chosen

An important part of this thesis was to construct two PGS based on GWAS summary statistics from both adults and children. Considering the discussion from the last section, the discovery GWAS needed to assess BMI and be composed of individuals from European ancestry. At the time of PGS construction, the adult and child GWAS chosen had the highest sample size that corresponded to those characteristics. The adult GWAS chosen combined ~700,000 adults of European ancestry in a meta-analysis of UK Biobank participants and the GIANT consortium [14]. The study tested a total of 16,652,994 variants and identified 941 that were independently associated with BMI and explained 6.0% of BMI variance in an independent sample [14]. Similar to previous evidence [19], the genes associated with BMI were involved in the development of the nervous system.

The child GWAS chosen combined ~60,000 children aged 2-10 years old of European ancestry in a two-stage meta-analysis of 40 individual studies [21]. The study tested a total of 8,228,795 individual variants across the 40 studies. A total of 47 variants were significantly associated with BMI in the discovery phase, and 25 of those also reaching statistical significance in the combined meta-analysis, including a replication sample. The study observed a total proportion of BMI variance explained by all the variants tested of 23% and a high genetic correlation between childhood and adult BMI ($r_g = 0.76$) [21]. The lesser number of variants identified compared to the adult GWAS is indicative of the lower statistical power compared to the adult GWAS. Importantly, the prediction accuracy of PGS increases as GWAS sample size grows [117], which is important to keep in mind when comparing the adult-derived and child-derived PGS. The summary statistics of both GWAS were publicly available and were retrieved from the GIANT consortium website and the GWAS Catalog [83] (study ID: GCST90002409).

3.3.5 PRS-CS procedure

The PRS-CS [119] procedure was used to construct the two PGS. PRS-CS was chosen to construct the two PGS because approaches relying on shrinkage procedures typically achieve better prediction accuracy compared to pruning and thresholding approaches [117]. Furthermore, PRS-CS improves

other Bayesian shrinkage procedures by 1) allowing the inclusion of diverse genetic networks using continuous shrinkage priors (instead of discrete mixture priors) on genetic effect sizes and 2) modelling local LD patterns more efficiently by updating effect sizes for markers in LD in concert rather than separately [119]. A summary of the procedure is given here, but readers interested in additional details are directed to the original article presenting the complete methodology [119].

The method introduces a prior for the variant effect sizes, which has a global and local component. The local shrinkage parameter follows an absolutely continuous density function (the Strawman-Berger prior). More specifically, this type of prior is a gamma-gamma prior with a=1 and b=1/2 specifications. The global scaling parameter shared for all variants controlling for the degree of spareness in the model can either be 1) inferred from a small number of values tested in an independent data set and chosen based on its predictive performance, or 2) found directly in the data (auto option) using a fully Bayesian approach with a standard half-Cauchy prior ($\phi^{1/2} \sim C^{+}(0,1)$) applied to the parameter. The authors of the method suggest that the automatic approach is optimal with a large sample size, and that fixing the parameter at 0.01 can be sufficient for highly polygenic traits [119].

The process of constructing the two PGS was performed on the Digital Research Alliance of Canada (alliancecan.ca) servers, formerly known as Compute Canada. The steps are based and adapted from an internal description of the use of PRS-CS for constructing a PGS of educational attainment with QLSCD data. First, the GWAS summary statistics file was modified to reflect the formatting used in PRS-CS. Second, the PRS-CS procedure was applied to the GWAS summary statistics file, the target genotypic files, and the 1000 genomes reference panel [188] with the auto global scaling parameter. Third, the variant effects were coded to all be positive by flipping the alleles with negative effects. Fourth, the PGS was finalized by adjusting the weights based on the imputation dosage and by combining each chromosomal score into a single final PGS per individual. The final adult-derived PGS contained 613,732 variants and the final child-derived PGS contained 689,789 variants. Additionally, both scores were adjusted by the first 10 principal components of ancestry and then standardized to obtain a normal distribution centred at 0 (Supplementary Text 4-1) using a method described elsewhere [194]. Both the GWAS summary statistics data and the target genetic data were following the GRCh37 assembly build.

3.4 Dietary habit variables

The interplay between dietary habits and genetic susceptibility to obesity is the main subject of the second and third objectives of this thesis. This section describes the data collection process and the development of the variables investigated in the studies associated with the thesis.

3.4.1 Eating behaviours

Data on eating behaviours were gathered by individual questionnaire items asked to the most knowledgeable person (generally the mother) about the child on five occasions from ages 2 to 6 years. The behaviours evaluated were eating too much, eating too fast, eating enough, being fussy with food, eating between meals, eating at regular hours, refusing to eat, and eating different meals than their parents. Those individual item questions were adapted and translated into French from the Avon Longitudinal Study of Parents and Children [195]. Individually, these items have been found to have moderate to strong heritability in a previous twin study of children from Québec [196], especially at younger ages. A total of 2014 participants (95%) had eating behaviour data available at least once between 2 and 6 years of age.

Using individual items, we created two scores to represent over-eating and fussy eating to evaluate them as potential mediators of genetic susceptibility to obesity in childhood. Those scores and their composite items are based on problematic behaviours described in an early report of the QLSCD [197]. Details about how the scores were constructed, and the specific composite items are summarized in Table 3-1. Briefly, an over-eating score (range: 2-8) was composed of two individual items (eating too much and eating too fast) summed at each collection wave and then averaged over available waves. The fussy eating score (range: 3-12) was composed of three individual items (eating different meals, being fussy with food and refusing to eat) summed at each collection wave and then averaged over available waves. Incidentally, although data was collected over time, each participant with eating behaviour data available at least once has a singular invariable value for each score.

These eating behaviour categories were devised post priori for analytic purposes, since when questionnaires were constructed, the area of eating behavioural research was still developing [198].

However, a more recent analysis involving QLSCD participants observed that higher over-eating scores in childhood were negatively associated with slowness in eating and satiety responsiveness in adulthood measured using the adult version of the CEBQ [161]. Similarly, children characterized by high fussy eating were more likely to exhibit food fussiness in adulthood [161].

Table 3-1. Description of eating behaviour scores and individual questionnaire items.

Eating behaviour	Questionnaire items	Response choices	Waves	Scoring process
Fussy eating	When [name of the child] is at home with you for the main meal of the day, how often does [name of the child] (1) eat different meals? In general, does [name of the child] … (2) is [name of the child] fussy about food? (3) refuse to eat?	Never (1) Rarely (2) Sometimes (3) Often (4)	2 years 3 years 4 years 5 years 6 years	Step 1. Three numerical response values (1-4) summed at each individual wave to obtain fussy eating score at each available wave. Step 2. Average fussy eating scores over waves available.
Over-eating	In general, does [name of the child] … (1) over-eat? (2) eat too fast?	Never (1) Rarely (2) Sometimes (3) Often (4)	2 years 3 years 4 years 5 years 6 years	Step 1. Two numerical response values (1-4) summed at each individual wave to obtain over-eating score at each available wave. Step 2. Average over-eating scores over waves available.

Note: Response choices for "eat different meals?" were: "Always" (1), "Almost always" (2), "Sometimes" (3), "Almost never" (4). The values were reversed before Step 1.

3.4.2 Dietary intakes

Four types of food intake (juice and fruit drinks, snacks, meats, and fruits and vegetables), three macronutrient intakes (proteins, lipids, and carbohydrates), and total energy intake were considered. Data regarding food intakes were gathered on multiple occasions between ages 1 and 13 years from food frequency questionnaires administered to the person most knowledgeable about the child (generally the mother). Since we were concerned about dietary intakes in early childhood, we retained the assessments made at 1, 2, 3, and 4 years of age. With the reference being last week, the parents were asked how many times per day/week the participant eats different types of foods. Some choices of answer were not number specific and were averaged before obtaining the amount per day (e.g., "3 to 4 times a week" to 0.5 times per day). The data collection process and details regarding the individual items composing the food intake categories are summarized in Table 3-2.

Outside of regular data collection waves, the QLSCD also incorporated special studies on specific subjects, including a nutrition focused data collection when children were four years of age. The data regarding macronutrient and energy intake was gathered during that special nutrition study through 24-hour dietary recall interviews administered by a registered dietician to the parent and a daycare attendant when the child attended daycare. The consumption values of proteins, lipids, carbohydrates and total energy were estimated based on the Canadian Nutrient File [199] and the USDA recipe file [200]. A total of 2057 (97%) and 1549 (73%) participants had food intake and macronutrient intake data available, respectively. As a summary, Figure 3-1 depicts the number of participants at various stages of the PGS construction process and the number of participants that also have exposure and outcome data available.

Table 3-2. Description of food intake variables and individual questionnaire items.

Dietary intake	Questionnaire items	Response choices and recoding per day	Waves	Scoring process
Juice and fruit drink	When [name of the child] is at home with you for the main meal of the day, how often does [name of the child] eat the following foods: (1) Juices and fruit drinks.	Never (0) Once or twice a week (1.5/7) 3 to 4 times a week (0.5) 5 to 6 times a week (5.5/7) Once a day (1) Twice a day or more (2)	1 year 2 years 3 years 4 years	Step 1. Recode item to times consumed per day. Step 2. Average consumption over available wave.
Snacks	When [name of the child] is at home with you for the main meal of the day, how often does [name of the child] eat the following foods: (1) Snack foods (pastries, sweet, cookies, potato chips).	Never (0) Once or twice a week (1.5/7) 3 to 4 times a week (0.5) 5 to 6 times a week (5.5/7) Once a day (1) Twice a day or more (2)	1 year 2 years 3 years 4 years	Step 1. Recode item to times consumed per day. Step 2. Average consumption over available wave.
Meats	When [name of the child] is at home with you for the main meal of the day, how often does [name of the child] eat the following foods: (1) Poultry. (2) Meat. (3) Fish.	Never (0) Once or twice a week (1.5/7) 3 to 4 times a week (0.5) 5 to 6 times a week (5.5/7) Once a day (1) Twice a day or more (2)	1 year 2 years 3 years 4 years	Step 1. Recode item to times consumed per day. Step 2. Sum three item values at each wave. Step 3. Average consumption over available wave.
Fruits and vegetables	When [name of the child] is at home with you for the main meal of the day, how often does [name of the child] eat the following foods: (1) Fruit (excluding juice). (2) Vegetables and potatoes.	Never (0) Once or twice a week (1.5/7) 3 to 4 times a week (0.5) 5 to 6 times a week (5.5/7) Once a day (1) Twice a day or more (2)	1 year 2 years 3 years 4 years	Step 1. Recode item to times consumed per day. Step 2. Sum two item values at each wave. Step 3. Average consumption over available wave.

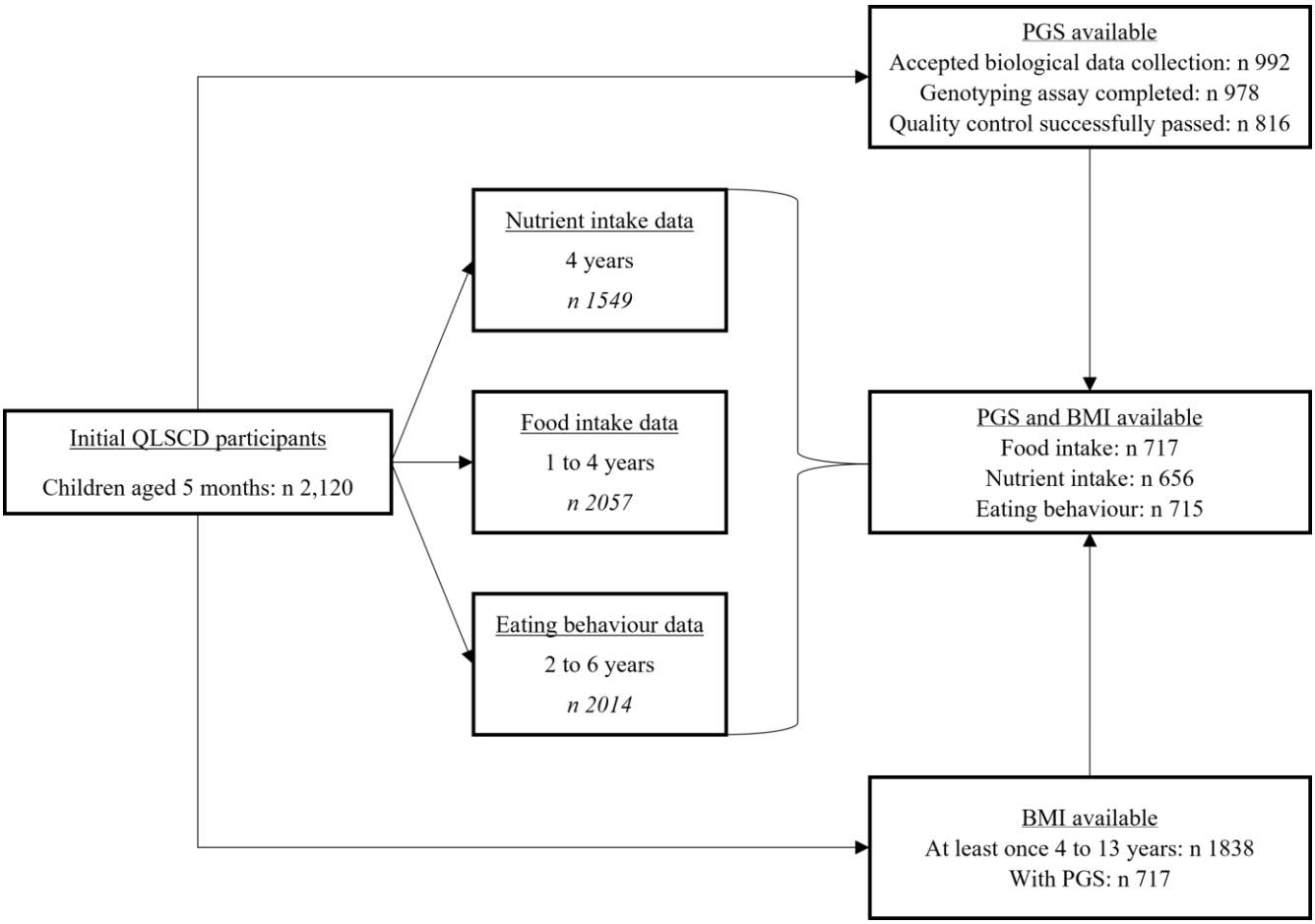


Figure 3-1. Flowchart of the number of participants with specific data available from the QLSCD.

3.5 Data analyses

This section presents the analysis methods that were used to investigate the three objectives of the thesis. The types of analyses are introduced and discussed here broadly, while specific model specifications, diagnostics, and additional sensitivity analyses are left for the manuscripts in Chapters 4, 5, and 6.

3.5.1 Objective 1

The first objective relates to the comparison of the association of a PGS trained with GWAS summary statistics from adults and children with BMI z-score and the risk of obesity in children. Typically, standard measures of associations can be used for such an evaluation. The phenotypic variance explained (R^2) and the effect size estimate (Beta) per unit of a PGS are generally used for continuous phenotypes (BMI z-score) [104]. For categorical phenotypes (risk of obesity), the effect size estimate (odds ratio) per unit of a PGS is typically used alongside measures of discrimination and calibration [106]. Consequently, the adult-derived PGS and the child-derived PGS were compared based on their association with BMI z-scores and the risk of obesity. Linear regression was used to obtain beta and R^2 estimates for the association with BMI z-scores, and logistic regression was used to estimate the association with the risk of obesity with odds ratios. Additionally, discrimination and calibration of the logistic models were evaluated using the area under the receiver operator curve and the Brier score. These analyses were conducted at each time point (4, 6, 7, 8, 10, 12, 13 years old) individually to ascertain trends of associations.

Additionally, graphical representations of the relation between a PGS and the trait of interest, such as strata plots [104], are presented. These plots are constructed by separating the study sample based on the PGS value (e.g., 10 equally sized strata representing 10% of the PGS distribution) and plotted against the mean (continuous) or prevalence (categorical) of the trait. Incidentally, the study sample with available PGS and anthropometric data was separated into deciles of the two PGS, and then plotted against the mean BMI z-score and the prevalence of obesity at each studied time point (4, 6, 7, 8, 10, 12, 13 years old).

3.5.2 Objective 2

Multiple methods to assess mediation have been developed, such as the Sobel test [201], which is commonly used with outcome measures at a single time point. When repeated outcome observations are available, longitudinal methods are favoured. Autoregressive models, latent growth curves and latent difference score models have been described previously and are commonly used in such designs [202]. More recently, methods have been described to evaluate the presence of mediation with multiple mediators [203], or with time-varying exposures and mediators [204].

In order to take advantage of the availability of repeated measures of BMI through time in the QLSCD, we opted to use longitudinal growth curve mediation analysis (LGCMA). LGCMA are structural models where variables measured at multiple time points are represented by two latent factors. The intercept refers to the starting point at time 1, and the slope defines the linear growth of the variable over time [202]. These models can take many forms depending on the presence, or not, of independent or mediator variables that are also available at multiple time points. The causal framework and the assumptions of these models were recently clarified in a publication [205]. The assumptions of LGCMA include 1) “no unmeasured confounding for the exposure-outcome relationship” (p.4) [205], 2) “no unmeasured confounding for the mediator-outcome relationship” (p.4) [205], 3) “no unmeasured confounding for the exposure-mediator relationship” (p.4) [205], and 4) “no mediator-outcome confounders which are affected by the exposure” (p.4) [205]. The general goal is to estimate the direct and indirect effects of the exposure on the outcome at each time point of the outcome.

Our study included two time-invariant continuous exposures (adult-derived PGS and child-derived PGS), a single continuous time-invariant mediator (fussy eating or over-eating measured over 2 to 6 years of age), and a continuous outcome (BMI z-score) measured at 6 time points (6, 7, 8, 10, 12, 13 years of age). Thus, the latent variables constructed were the intercept and the linear slope for the BMI z-score. Adapting the methodology described earlier [205], we were able to estimate the proportion of the association between both PGS and BMI z-score that is mediated through each eating behaviour separately from 6 to 13 years. The complete formulas and a description of the models, including covariate adjustments, are presented in the appropriate manuscript section (Chapter 5).

3.5.3 Objective 3

Statistical interactions are usually evaluated using statistical models by adding a product term between the two exposures of interest [23]. For continuous outcomes, the coefficient of the product term in linear models reflects the interaction effect on the additive scale. For this objective we are interested in the interaction between continuous PGS and continuous dietary intakes (food intakes: mean score between 1 and 4 years old, macronutrient and energy intakes: at 4 years old) in relation to BMI (4, 6, 7, 8, 10, 12, 13 years old). Considering repeated measures of BMI are available, we opted for linear mixed models in order to account for the clustering per participant. Those models allowed to investigate the influence of the exposures (and their interaction) on both the mean level of BMI (intercept) and the growth with time (slope) by adding a product term for the time (or age of participants). Thus, we were able to examine the influence of food, macronutrients, and energy intakes on BMI, and if these effects varied by the child-derived PGS, the adult-derived PGS, and the age of the participants.

CHAPTER 4 - MANUSCRIPT 1

Article preface

This manuscript introduces the adult-derived and child-derived PGS used throughout this thesis. The scores are compared based on their association with BMI and the risk of obesity from 4 to 13 years of age using data collected from the QLSCD. Incidentally, this first manuscript relates to the first objective of the thesis: to measure the genetic susceptibility to obesity in children using two PGS derived from adults and children GWAS summary statistics, respectively, and compare their association with BMI and discrimination of obesity.

DG, LD and MB contributed to the conception and design of the study. LD and MB were involved in the data collection from QLSCD participants. DG performed the statistical analyses with insights and revisions from CG. DG was responsible for the first draft of the manuscript. All authors participated in discussions concerning the development of the study and were involved in the revisions to the manuscript and approved the submitted version.

Ethics approval (secondary analysis of QLSCD data): University of Ottawa Research Ethics board (ethics file number: H-01-23-8018, approval date: 10/02/2023, Appendix 1).

This manuscript was resubmitted after review to the Canadian Journal of Physiology and Pharmacology on December 11th, 2024. A copy of the confirmation email is available in Appendix 3. A PRS-RS reporting checklist for manuscript 1 is also available in Appendix 4. The following manuscript is harmonized to the thesis format.

Title page

Title: Polygenic scores of obesity in childhood based on summary statistics from adults vs. children

Danick Goulet[1*], Michel Boivin[2], Christopher Gravel[1, 3, 4], Julian Little[1], Isabelle Ouellet-Morin[5], Jean-Philippe Gouin[6], Lise Dubois[1]

[1]School of Epidemiology and Public Health, University of Ottawa, Ottawa, ON, Canada

[2]Université Laval, Québec, QC, Canada

[3]Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montréal, QC, Canada

[4]Department of Mathematics and Statistics, University of Ottawa, Ottawa, ON, Canada

[5]École de criminologie, Université de Montréal, Montréal, QC, Canada

[6]Département de psychologie, Université Concordia, Montréal, QC, Canada

***Co-correspondant authors**:

Lise Dubois : ldubois@uottawa.ca

Michel Boivin: michel.boivin@psy.ulaval.ca

Word count (main text): 5794 (main text), 194 (abstract)

Number of tables: 1

Number of figures: 4

Number of Supplementary materials: 11

Competing interests: The authors declare that there are no competing interests.

Abstract

The lack of polygenic scores (PGSs) developed for body-mass index (BMI) in children may be problematic because the genetic architecture characterizing BMI changes throughout life. This study aims to describe the genetic susceptibility to obesity in children and to compare two PGSs based on data from adults and children and their association with BMI and discrimination of obesity. The study sample comprises 717 participants aged 4 to 13. Adult- and child-based PGS were evaluated by examining 1) mean BMI across polygenic score risk categories, 2) the capacity to identify obesity with logistic regression and 3) the linear association with BMI z-scores using linear regression. Increases in one standardized unit of adult-based PGS were related to a stronger increase in BMI z-score (β=0.24-0.39) than PGS derived in children (β=0.21-0.30). The association between obesity and the child score was higher (OR = 1.75-2.33) than for the adult score (OR = 1.74-2.06) for the age group 4-7 years. The inverse was observed for the age group 8-13 years (OR_{child} 1.56-1.79 vs. OR_{adult} 1.78-2.54). Both adult- and child-based PGS show strong associations with BMI and risk of obesity, with the adult-based score standing out from 8 years old.

Abbreviations: BMI – body-mass index; GWAS – genome-wide association study; PGS – polygenic score; QLSCD – Quebec Longitudinal Study of Child Development; WHO – World Health Organization

Keywords: Adiposity, obesity, polygenic score, longitudinal study, childhood

Introduction

In 2015, obesity was among the deadliest health problems, responsible for approximately 4 million deaths and a loss of 120 million healthy years of life in 195 countries [1]. Starting at conception, genetic and environmental factors have been shown to dynamically and synergistically contribute to obesity, and their influence accumulates throughout life [2]. Since childhood obesity tends to persist into adulthood, a clear understanding of its underlying genetic and environmental etiology is paramount to inform prevention, ideally starting early in life [3].

The heritability of body mass index (BMI) is estimated to range between 40 and 80% [4]. As for most complex phenotypes, the genetic architecture of obesity is polygenic, characterized by a large array of genetic variants, each exerting a minor effect. Since the introduction of high-throughput and efficient genotyping technologies, genome-wide association studies (GWASs) have examined the potential role of millions of genetic variants and have identified hundreds of variants associated with BMI in participants of European ancestry [5]. Polygenic scores (PGSs) build on the findings of GWASs to estimate continuous indexes of genetic susceptibility to a given phenotype by aggregating the effects of genetic variants located across the genome. Many PGSs for BMI have been proposed for adult populations, typically based on the GWAS of hundreds of thousands of participants. However, there is only a limited number of PGSs developed to be applied to samples of children or adolescents because there are fewer GWAS for children and adolescents that are also based on a smaller number of participants ($N < 60\,000$) [6-10].

Evidence suggests that the genetic architecture of obesity may change throughout life [4]. For instance, the heritability of BMI is known to vary depending on age, with higher heritability found in childhood and adolescence compared to adulthood [4]. Also, while specific genetic variants are associated with BMI in both childhood and adulthood, their effects may nevertheless differ across developmental periods [11]. However, only a few studies have used a PGS to predict BMI on more than one occasion throughout childhood or adolescence. Khera et al. assessed BMI at six points in time between birth and 18 years of age using the Avalon Longitudinal Study of Parents and Children. The study showed that among youth, differences in weight and BMI between those more and those less genetically susceptible increase from birth to 18 years [10]. This is supported by another study observing that the effect of a PGS on BMI increases in the transition between adolescence and adulthood [12]. Another study [13]

followed over 1000 participants from birth to 38 years of age. The study showed that children with a higher PGS were characterized by adiposity rebound displaying prematurely and accrued more weight between birth and 3 years of age. Those characteristics were then shown to mediate the association between the PGS and obesity in later years [13]. Hüls et al. measured BMI at two time points between ages 2 and 16 and showed that the correlation between a PGS and BMI increased with age [8]. Given the small number of studies using BMI-PGS in children and adolescent populations, more studies need to describe the value of PGSs estimated from adult populations on children and adolescents' BMI to determine if they differ from those from child populations. Additionally, more investigation is needed to estimate how those PGSs relate to obesity as well, since the scores may affect adiposity differently at the higher end of the distribution.

With one exception [9], all studies using a BMI-PGS among children or adolescents were based on GWAS summary statistics from adult samples, either the GIANT meta-analysis [14] (n=339 224) or a meta-analysis of both GIANT and the UK Biobank [5] (n~700 000). In contrast, the study by Fang et al. used GWAS summary statistics from a joint analysis of 20 GWASs uniting over 35,000 children [15]. Because PGSs directly stem from the GWAS summary statistics, the referenced sample on which the GWAS is based may affect the capacity for the resulting PGS to predict the phenotype of interest. For phenotypes that are likely to have a distinct genetic etiology across developmental periods or for which genes may be differently linked with the phenotype depending on age, careful consideration must be given to the age composition of the reference GWAS [16]. This is especially true in the case of BMI, for which the genetic variants, their respective weights, and their genetic expressions might differ with age [16]. Overall, the majority of variants associated with childhood BMI also have evidence for association in comparative GWAS of adult BMI [11, 15, 17]. Conversely, only 22 out of 97 variants known for their association with adult BMI were also found to be associated with childhood BMI [15], which could be explained by the lower sample size used in GWAS for children compared to adults. This raises the question of the appropriateness of adult-based PGSs in child samples, opposing adult-based scores that are derived from GWASs with higher sample sizes to child-based scores that would include variants more specifically associated with BMI in childhood.

To date, few analyses have compared the use of childhood-based and adult-based BMI-PGS in children. A recent study [18] compared four PGSs derived from summary statistics estimated for different adiposity-related traits, including childhood BMI and adult BMI. Both scores showed similar patterns of BMI change through age and genetic risk, where BMI differences between those in the

highest and lowest decile of genetic susceptibility grew over time. However, the adult BMI-PGS better predicted BMI, especially after the age of 3 [18]. These findings suggest that GWASs conducted with a sample consisting of a majority of adults yield better performance, even at an early age, than a PGS estimated from a GWAS that includes children and adolescent participants. This pattern of findings was likely due to the larger sample sizes of the adult-based GWAS and PGS, which allowed for more precision in the estimates of BMI [18]. Incidentally, this study addresses 1) the need for longitudinal assessments of BMI-PGS throughout childhood and adolescence and 2) the uncertainty surrounding the use of adult- vs. child-based PGSs in childhood. The purpose of the study is to measure genetic susceptibility using PGSs with summary statistics from children and adults in a longitudinal cohort of children and to compare their respective performances.

Materials and methods

Study design and participants

A representative sample of children born in 1998 in the province of Québec, Canada (n=2120) accepted to take part in the Quebec Longitudinal Study of Child Development (QLSCD) [19], for which the first data collection took place at the age of 5 months. These children and their families have been followed longitudinally, and follow-up is still ongoing. Data on parents (e.g., demographic, breastfeeding, parental depression) and child characteristics (e.g., child behaviour, mental/physical health, diet, height, and weight assessments) were gathered annually or biennially up to 23 years of age using questionnaires and interviews [19]. DNA collection for genetic analyses was extracted when the child was 10 years old. Informed consent was obtained from the parents along with children's assent from the age of 10, while the participating youth consented for themselves when they reached adolescence. DNA collection required additional approval from the study participants. Both the initial QLSCD formation and additional genetic data collection were approved by the ethics committee of the Centre Hospitalier Universitaire Sainte-Justine Mère-Enfant (#2104).

Genotyping and polygenic score construction

A subsample of the 1334 QLSCD participants at the 10-year-old collection were approached to collect DNA samples, and 992 accepted. Genotyping was performed using the Illumina Infinium PsychArray-24. The array was chosen because the cognitive development of children was an area of research of interest in the QLSCD. Nonetheless, most markers are proven tag single nucleotide polymorphisms from core arrays. Quality control and imputation steps applied to the QLSCD genotypic data have been previously described [20]. Briefly, checks for variant allele frequency, genotyping rate, and Hardy Weinberg Equilibrium test p-values were implemented before and after the imputation using the 1000 genomes project reference data [21], SHAPEIT v2 (r837) [22], and impute2 [23]. Additionally, a cryptic relatives check was implemented to identify genetically related pairs, and individuals with a call rate below 0.95, those flagged for gender mismatch, and genetic duplicates were excluded. After the quality control and imputation, genetic data on 8,407,807 variants were available for 816 participants.

Weighted and standardized PGSs were constructed using the PRS-CS method, which has been shown to outperform other procedures, especially in large sample sizes [24]. We used the model-tuning method PUMAS [25] to identify the best value for the global shrinkage parameter. The procedure only requires GWAS summary statistics and a linkage disequilibrium reference panel. The GWAS summary statistics were partitioned for PGS training (0.75) and PGS parameter evaluation (0.25) purposes, which was repeated 4 times using Monte-Carlo cross-validation. We tested the global shrinkage parameter values of 0.01, 0.0001, and auto. The parameter auto outperformed (mean $R^2 = 0.207$) the other two values (0.01: mean $R^2 = 0.054$, 0.0001: mean $R^2 = 0.185$) for the adult-derived PGS. For the child-derived PGS, the 0.0001 and auto values had similar performance (mean $R^2 = 0.022$). We opted to use the auto parameter to reduce the difference between the adult and the child scores. We regressed each score on the first 10 principal components of genetic ancestry to account for population stratification. We used the subsequent residuals to obtain adjusted scores, as previously described [26] (R script in Supplementary Text 4-1). The most populated GWAS meta-analysis summary statistics for BMI conducted for adults [5] and children [11] of European ancestry were used alongside the 1000 Genomes reference panel [21] to construct the scores. The final PGS derived from adult and child GWAS summary statistics included 613,732 and 689,789 individual variants, respectively. Almost all variants that were included in the adult PGS were also in the child PGS (613,386 common variants). Out of the 100 variants with the highest weights in the child PGS, 85 had similar directionality as with the adult PGS. Similarly, 81 out of the 100 variants with the highest weights in the adult PGS had

effects in the same direction as the child PGS. Additionally, there were 11 common variants out of the 100 variants with the highest effect sizes in both scores.

Anthropometric measures

Trained technicians performed height and weight measurements seven times between 4 and 13 years of age (4, 6, 7, 8, 10, 12, and 13 years). The standardized protocol followed by the technicians involved a measuring tape, a ruler and a scale, where the children were weighed without shoes with only light clothing. BMI z-scores were calculated using the World Health Organization (WHO) Growth reference data [27] to account for the children's age and sex. Weight status was established based on the WHO classification. Children aged 5-19 with a BMI above 1 standard deviation from the WHO growth reference were considered overweight, and those above 2 standard deviations were considered in situation of obesity. Children under 5 years of age with a BMI above 2 standard deviations were considered overweight, while those with a BMI above 3 standard deviations were considered in situation of obesity. Of the 816 children with available genotype data, 717 had anthropometric data available at at least one time point.

Statistical analysis

The sociodemographic characteristics of the total sample were estimated. We also compared the sociodemographic characteristics of the QLSCD participants included in the study (n=717) with those excluded (n=1403) (Supplementary Table 4-1). Comparisons were tested using Chi-square test for proportions, and one-way ANOVA for normally distributed variables.

The association between child- and adult-based PGSs, and BMI z-scores was modelled using linear regressions for BMI measurements taken from the age of 4 to 13 years. The linear effect estimates were computed with 95% confidence intervals, and the proportion of variance explained by each score was evaluated using R^2. Next, participants were stratified into PGS risk categories to observe how the mean BMI z-score and proportion of individuals with overweight and obesity vary across genetic risks. These analyses were conducted separately at each BMI collection time. The unpaired T-test was used to evaluate the difference in BMI z-score between the first and fifth quintiles to examine if those with the highest genetic risk have significantly higher BMI than those with the lowest risk. Finally, we

compared the capacity of the PGSs to identify youth with obesity at each time point using logistic regressions (ORs, AUC, Brier score).

Additionally, we obtained "overall" effect estimates across all ages accounting for repeated measurements by performing linear mixed models for the continuous BMI z-scores (4, 6, 7, 8, 10, 12, and 13 years) outcome and generalized linear mixed models for the obesity outcome. Mixed models were constructed first by evaluating how best to represent the age-related difference in BMI z-scores using the Akaike Information Criterion. The best model was deemed to include a random intercept and random slope for age, an autoregressive of order 1 covariance structure, and a smooth term for age. Adult and child PGSs were then added separately as continuous variables. We further examined the residuals to verify the assumptions about the linearity of the models. Scaled residuals for generalized linear mixed models were used to evaluate model misspecification and over/under dispersion using the DHARMa R package [28]. Removing the random slope for age in generalized mixed models appeared to mitigate underdispersion and heteroscedasticity issues.

In additional analyses, we 1) adjusted the linear, logistic and mixed models for age, sex, birth weight, maternal education (post-secondary education or not), household income (less than $60,000 or at least $60,000), and whether mothers were born in Canada, and 2) estimated sex-specific results by including in the model an interaction term between PGS and sex. Using the R statistical package "simr" [29] we estimated power between 0.89 and 0.95 to identify an effect of 0.15 per PGS standard deviation in simple linear regressions, which is very conservative considering that in a sample of 11-year-old children [30] each increase of 1 standard deviation of a PGS was associated with 0.40 increase in BMI z-score. Statistical analyses were performed using R [31] version 4.2.2.

Results

Table 4-1 shows the baseline characteristics of the 717 study participants. Overall, slightly more participants were females (55.2%), with mothers of average age at birth (20-34 years, 84.0%), and with medium household income ($30,000 to <$60,000, 41.2%) at baseline. The mother's education was evenly spread from a secondary-school to a university diploma (24.1 to 30.5%) at baseline. We observed that the QLSCD participants included who had both valid BMI and PGSs tended to be more likely to be female (55.2% vs. 45.8%), had mothers with higher levels of education (59.9% vs. 53.6%

with post-secondary education), grew up in families with higher levels of income at 5 months (33.8% vs. 28.4% >$ 60,000), and were born in Canada (97.8% vs. 83.8%) compared with those excluded (Supplementary Table 4-1).

Linear association between PGSs and BMI z-score

Figure 4-1 presents the results of the simple and mixed linear associations between PGSs and BMI z-scores from 4 to 13 years of age. Using linear mixed models, we estimate that the adult PGS explained 7.2% (ranging from 3.9 to 10.4% across time points) of BMI z-score variance, compared to the 5.4% (ranging from 3.5% to 6.9% across time points) proportion explained by the child PGS (Figure 4-1). Considering all repeated observations with the linear mixed model, the increase of one standardized deviation of the adult PGS was associated with an increase of 0.32 standardized units of BMI (β=0.32, 95% CI: [0.25, 0.39]), compared to 0.25 for the child PGS (β=0.25, 95% CI: [0.18, 0.32]). The effect of the adult PGS on BMI z-scores increased from 4 years (β=0.24, 95% CI: [0.15, 0.33], R^2=3.9%) to 8 years (β=0.37, 95% CI: [0.28, 0.46], R^2=8.8%), after which it plateaued ($0.34 \leq \beta \leq 0.39$) with the highest estimate noted at 12 years (β=0.39, 95% CI: [0.30, 0.47], R^2=10.4%). In contrast, the highest per-standardized-unit increase of the child PGS was observed at 7 years (β:0.30, 95% CI: [0.22, 0.39], R^2=6.9%), and without further improvement noted subsequently ($0.24 \leq \beta \leq 0.29$). In sum, the linear association between the adult PGS and BMI z-score was numerically higher than that of the child PGS at all time points, but the effects of both PGSs on BMI z-score were within each other's confidence intervals at all time points. All estimates related to the linear associations between PGSs and BMI z-scores are available in Supplementary Table 4-2.

Both PGSs identify a BMI gradient across quintiles and distinguish those at higher obesity risk

Figure 4-2 shows a positive linear gradient of the PGS quintiles for each PGS on the BMI z-scores collected from 4 to 13 years of age. For both PGS and at all time points, the mean BMI z-score in the fifth quintile was significantly higher than in the first quintile ($p<0.001$). Although comparable at earlier time points, the difference in mean BMI z-score between the first and fifth quintiles appeared to increase starting at 8 years of age for the adult PGS compared with the child PGS (Figure 4-2 E-G), indicating that genetic variants characterizing adult obesity may already emerge at that age.

Figure 4-3 presents the proportion of children classified with a healthy weight, overweight, or obesity for each PGS quintile from 4 to 13 years of age, where both PGSs seemed to distinguish the weight status of children. The proportion of children classified with obesity generally increased with higher PGS at each time point. Differences in obesity proportion between youth at the lowest and highest child-derived genetic risk remained stable (i.e., from 9.2% to 13.6%) over the age range. In contrast, that difference markedly increased for ages 8-13 years (15.4%-18.2%) for the genetic risk derived from the adult PGS, in comparison with BMI measurements collected at an earlier age (i.e., 4 and 6 years: 6.8%-9.7%). Similar to results observed for mean BMI z-scores, the adult-based score seems to get better at identifying children at higher risk of obesity with age.

Association with obesity

Figure 4-4 shows the odds ratio of obesity for PGS derived from adults and children GWASs. Considering all observations in a generalized mixed model, the odds of obesity increased by a 2.45-fold (95% CI:1.45, 4.15) for every standardized unit increase in the adult PGS, compared to a 1.88-fold increase (95% CI:1.16, 3.03) for the PGS derived from GWAS conducted in children and adolescent samples. We observed that from 4 to 7 years of age, odds ratios were numerically higher for the child PGS, whereas the inverse was observed from 8 to 13 years of age. The highest ORs were observed at 7 years for the child-based PGS (β: 2.33, 95% CI:1.67, 3.31), and at 8 years for the adult-based PGS (β: 2.54, 95% CI:1.89, 3.49). Assessing the accuracy of each score to identify obesity, the AUC and the accuracy of probabilistic predictions measured by the Brier score were similar across time points for the child (AUC:0.62-0.72, Brier:0.05-0.12) and adult PGS (AUC:0.65-0.74, Brier:0.5-0.12). Complete numerical results for logistic regression models are available in Supplementary Table 4-3, and Receiver-operator curves and calibration plots are available in Supplementary Figures 4-1 and 4-2.

Additional analyses showed that adjusting for sociodemographic factors did not affect the conclusions gathered from the associations between genetic scores and BMI z-scores or obesity risk throughout childhood (Supplementary Figures 4-3 and 4-4). The proportion of BMI z-score variance explained by the linear models increased slightly (1.2 to 3.3% increase). Similarly, the AUC increased for obesity risk models (0.03 to 0.10 increase). Furthermore, including an interaction term between each of the PGSs and sex into the previously described models showed that the linear effect of the adult PGS on BMI z-scores was significantly stronger in males compared to females at all time points except ages 4

and 12. The stronger adult PGS effect associated with being male was constant across time points, ranging from an additional 0.17 to 0.20 standardized units of BMI z-scores (Supplementary Table 4-4). Considering the higher proportion of girls at latter time points due to attrition this could impact prior results. Since the impact of the adult-based PGS on BMI z-scores is lower in girls, differences in BMI z-scores between quintiles and effect estimates for the association between the adult-based PGS and BMI z-scores could be underestimated in older children (more conservative estimates). Effect modification by sex on the linear effect of the adult PGS did not generalize to obesity (Supplementary Table 4-5).

Discussion

This study aimed to compare two BMI-PGSs derived from adult and child GWAS summary statistics in a longitudinal cohort of children. Previous evidence suggests that due to a larger sample size, BMI-PGS derived from studies on adult populations may be better suited to identify children at the higher end of BMI distribution [18]. Our results mostly agree with prior documentation, as the adult-based PGS displays a stronger association across most analyses, although the child-based PGS still showed promising results, especially when it comes to identifying younger children at risk of obesity. More specifically, the PGS derived from summary statistics from adults' linear association with BMI z-score was numerically higher than that of a PGS derived from children summary statistics across all time points, and from 8 years of age onward it captured higher discrepancies in mean BMI z-scores and in the proportion of obesity between high and low genetic risk groups, and it showed a stronger ability to identify obesity. Nonetheless, the closeness in the associations with BMI and risk of obesity displayed by both scores indicates they should be valid to use in various circumstances. For instance, the PGSs can be used as instrument variables to represent genetic susceptibility to obesity in Mendelian randomization studies, or etiological analyses assessing the intricacies of the development of obesity in the presence of genetic determinants. These findings are essential because BMI-PGSs are rarely applied in childhood or adolescence. Only a handful of other studies have explored how adult- and child-based PGSs compare, and we showed that a child-based PGS could outperform an adult-based PGS in specific circumstances.

First, only a few studies have examined the associations of PGSs with BMI across childhood and adolescence. Our PGS based on adult summary statistics compares favourably with two previously published PGSs [6, 7] that also used adult summary statistics. A score of 2321 variants explained 4.7% of the variance in BMI in 3414 11-year-old children [6], compared with 8.8% from our adult-based PGS. Similarly, a score including 941 variants explained 6.5% of the variance in BMI in 1289 individuals of 11 to 22 years of age [7], compared with 8.8-10.4% from 10 to 13 years of age for our adult-based PGS. Another study had access to BMI measured longitudinally in children and adolescents [8]. Considering all repeated measurements, the proportion of variance explained by our adult-based PGS (7.2%) was at the low end of what was observed for the score used in Hüls et al. (11%), which comprised over 2.1 million variants in 3098 children aged 2 to 16 years [8]. A possible explanation is the maximum age in Hüls et al. being higher than that of our study, where the adult-based PGS explained a higher proportion of the variance in BMI at older ages.

Moreover, the ability of both our scores to classify obesity aligns with previous research. We observed an area under the curve ranging from 0.65 to 0.74 and 0.62 to 0.72 for the PGS derived from adult and child summary statistics, respectively. This implies that, at most, the probability that someone with an obesity phenotype presents a higher score than someone without obesity is 0.74 for the adult-based PGS. This accuracy is at the high end of what was observed using the PGS in Hüls et al., with an AUC of 0.64 in adults from the UK Biobank [32]. One of the potential uses of PGSs in a clinical setting is to help identify individuals at higher risk of developing obesity. This information can have benefits for both individuals that have not yet developed obesity and those who have already developed obesity. A recent review [33] of the utility of PGS in relation to obesity highlights the potential benefit of implementing preventive measures early in life for those at higher risk of developing obesity, and to inform treatment practices, such as the introduction of pharmacological treatment in those at higher risk of persisting obesity.

However, the current clinical utility of genetic scores for population screening and individual risk prediction is lacklustre for complex diseases. A recent study [34] assessing 926 scores for 310 diseases observed a 5% false positive rate median of only 11% over all the scores. Considering that to obtain a 5% false positive rate of 80% it would require observing an area under the curve of 0.96 [34], the lack of utility in population screening applies to the use of PGS in the context of obesity as well. The current classification ability of PGSs for obesity (including in the present study) is still not high enough to be effectively employed for population screening. Other suggested clinical uses for PGS have been

discussed, such as encouraging patients to start an intervention plan [35]. However, it remains unclear whether delivering genetic risk information to individuals might produce effective behavioural change. In some cases, there is concern that genetic risk information may even lower motivation and provoke anxiety [36]. Additionally, there are concerns about furthering health disparities since most of the PGS predicting BMI have been constructed using GWAS summary statistics from samples of European ancestry [33]. Still, considering the high involvement of environmental factors in the etiology of obesity, incorporating non-genetic factors may improve the clinical utility of PGS for obesity [32].

Only a handful of studies compared the use of PGSs derived from adult versus child populations concerning their association with BMI during childhood. We observed similar results to those obtained in a previous study that compared BMI z-score trends in children based on PGS using both adult and children summary statistics [18]. In both our research and the study by Helgeland et al., the child-based PGS showed similar patterns in BMI z-scores depending on PGS risk compared with the adult-based PGS, albeit with a lower proportion of variance explained and lower differences in mean BMI z-score between the extremes of genetic risk. In Helgeland et al., the difference in BMI z-score between the lowest and highest decile of the PGS started growing rapidly from the ages of 3 to 8. In contrast, this difference was more pronounced beginning at 8 years of age in our study. This suggests that genetic variants characterizing adult obesity are already present in early life. This is supported by the observation that out of 20 variants (or their proxy variants) that were significantly associated with BMI in the child GWAS [11] used for our child-derived PGS, 18 were also significantly associated with BMI in the comparative adult GWAS [5]. This also aligns with results obtained from the Avon Longitudinal Study of Parents and Children, where the discrepancies in BMI and BMI z-score between the lowest and highest decile of an adult-derived PGS comprised of 2.1 million variants grew over time from birth to 18 years of age [10].

Another study [37] comparing PGSs derived from adult and child summary statistics in the context of obesity found that the proportion of variance of BMI explained by a child-based PGS was higher than that of the adult-based PGS between 12 and 24 years of age. This difference in results relative to our study could be explained by the fact that the child-based PGS used in that study was constructed using recall data from adult participants about their height and weight in childhood and early adulthood [38], a method prone to misclassification bias. Since this information bias is unlikely to be related to an individual's genetic profile, this likely results in reduced differences between exposures and thus attenuation of the associations found. However, the long recall period makes it difficult to infer the

direction of bias. Another possible explanation is that the child PGS captures more genetic variation specific to childhood weight regulation. Although computed using a high sample size (n=453 169), both scores underperformed compared with the PGS calculated by Khera et al. [10]. The fact that the effect of the child-based PGS on BMI is somewhat constant throughout life, contrary to the adult-based PGS, which increased with time, further suggests that a child-based PGS could have a stronger association with BMI than an adult-based PGS in early life, given an equal sample size to compute the scores. Indeed, adult GWAS for BMI have significantly higher sample size compared to child GWAS. Generally, a higher GWAS sample size increases the power to identify variants associated with BMI [39], which leads to better accuracy of the resulting PGS [40]. BMI GWAS using adult data are getting close to the theoretical upper bound of proportion of variance explained by common genetic variants [33]. It remains unclear if a higher populated child GWAS could produce PGSs surpassing the predictive ability of adult-derived PGSs in childhood.

We found that the association between the adult-based PGS and BMI z-score was stronger in males than in females in childhood. In addition to the presence of a different array of genetic variants affecting weight in females and males, differences between sexes in adipose tissue distribution, or in gonadal hormone pathways [41], especially in childhood and adolescence, may present a possible explanation. Our observation contradicts the results of an earlier study. A PGS comprising 97 variants was found to have a stronger effect on BMI and BMI change in females vs. males in early adulthood (18 to 45 years of age) [42]. A possible explanation is the different age ranges studied, which could involve other biological mechanisms. However, it is worth noting that this observation stems from a secondary analysis with a relatively small sample size that may lead to false positives. Rather, our results should be used as a hypothesis for larger studies that could replicate this finding.

The primary strength of this study is the availability of longitudinal anthropometric data. Since childhood is an important time in the genetic architecture of obesity, the possibility to examine multiple time points helps to better characterize how the relationship between genetic risk and BMI changes with age. Also, in QLSCD, the height and weight data collected were directly measured with the participants. Parent-reported anthropometric measures of children are known to cause bias due to underestimation of weight and overestimation of height [43]. Directly measuring this information helps avoid possible information bias. Another strength of the study is using PRS-CS to construct both PGSs. The method, which considers information on all the available variants, is known to outperform thresholding and pruning methods using only a fraction of them.

The study also had a few limitations. First, we used BMI as a measure of adiposity. BMI can be used to measure overall levels of body fat but is limited when it comes to adequately representing the location of excess adiposity and does not distinguish between lean and fat mass. While we acknowledge the use of BMI may obscure the relationship between a PGS and adiposity, this measure is the most widely used in GWASs and PGS studies of adiposity traits and is strongly associated with adverse health consequences due to excess weight. Second, differences between study participants and total QLSCD participants may induce a selection bias. Characteristics related to both genetic risk and BMI may influence exclusion from the study, either by loss of follow-up or by unavailable genetic information. Third, the genotyping array (PsychArray-24) was created to study psychiatric genetic susceptibility by containing markers (~50,000) specifically associated with psychiatric disorders. Although the number of variants contained in an array or the genome-wide coverage does not typically influence imputation quality (especially for common variants of minor allele frequency > 5%), an uneven spread of the variants may diminish imputation results [44]. However, the PsychArray-24 also contains 271,000 markers from the Core array which has shown to have a coverage of variants present in the GWAS Catalog after imputation similar to that of other comparable arrays [44]. Nonetheless, it is possible we have missed known variants associated with BMI variation by using a psychiatric specific array. Fourth, our study was conducted in a sample of French-Canadians, a population of European ancestry, which limits the generalizability of our results to populations of other ancestry.

Our investigation indicates that BMI-PGS derived from both adult and child GWAS summary statistics show strong associations with BMI and obesity over 4-13 years of age. The adult-based PGS was superior at predicting continuous BMI over the age range. A shift was observed around 8 years of age. Before that point, the child-based PGS was able to better predict obesity risk, whereas starting at 8 years the adult-based PGS displayed better prediction ability. This suggests that the mechanisms underlying genetic propensity for obesity in adulthood already emerge in early childhood. Still, the ability of the child-based PGS to identify children at elevated risk of obesity, in low-age groups especially, suggests it should be readily applied in investigations regarding younger children.

ACKNOWLEDGEMENTS

We acknowledge the contribution of Till Andlauer, Stéphane Paquin, Geneviève Morneau-Vaillancourt, Isabelle Ouellet-Morin and Michel Boivin, who were involved in the quality control of the genetic data of the QLSCD participants that are used in the research. We are grateful to the QLSCD participants and their families who took part in the various data collection rounds over the years.

AUTHOR CONTRIBUTIONS

DG, LD and MB contributed to the conception and design of the study. LD and MB were involved in the data collection from QLSCD participants. DG performed the statistical analyses with insights and revisions from CG. DG was responsible for the first draft of the manuscript. All authors participated in discussions concerning the development of the study and were involved in the revisions to the manuscript and approved the submitted version.

FUNDING

This work was supported by a CIHR operating grant (#165964). The funders were not involved in the design or conduct of the study; collection, management, analysis, or interpretation of the data; or preparation, review, or approval of the manuscript. The analyses were performed using data from the Quebec Longitudinal Study of Child Development (QLSCD), conducted by Sante Quebec, a division of the Institut de la Statistique du Quebec (ISQ) and funded by the Ministry of Health and Social Services of Quebec.

DATA AVAILABILITY STATEMENT

The data used to support the findings of this study was obtained from the Québec Longitudinal Study of Child Development, conducted by Santé Québec, a division of the Institut de la Statistique du Québec and may be released upon application to the Institut de la Statistique du Québec, through the Zone de recherche at: https://statistique.quebec.ca/fr/institut/services-recherche#/accueil

References

1. GBD 2015 Obesity Collaborators, et al., *Health Effects of Overweight and Obesity in 195 Countries over 25 Years.* The New England journal of medicine, 2017. **377**(1): p. 13-27 DOI: 10.1056/NEJMoa1614362.
2. Kansra, A.R., S. Lakkunarajah, and M.S. Jay, *Childhood and Adolescent Obesity: A Review.* Frontiers in pediatrics, 2021. **8**: p. 581461-581461 DOI: 10.3389/fped.2020.581461.
3. Styne, D.M., et al., *Pediatric Obesity-Assessment, Treatment, and Prevention: An Endocrine Society Clinical Practice Guideline.* J Clin Endocrinol Metab, 2017. **102**(3): p. 709-757 DOI: 10.1210/jc.2016-2573.
4. Silventoinen, K. and H. Konttinen, *Obesity and eating behavior from the perspective of twin and genetic research.* Neuroscience & Biobehavioral Reviews, 2020. **109**: p. 150-165 DOI: https://doi.org/10.1016/j.neubiorev.2019.12.012.
5. Yengo, L., et al., *Meta-analysis of genome-wide association studies for height and body mass index in ∼700000 individuals of European ancestry.* Human molecular genetics, 2018. **27**(20): p. 3641-3649 DOI: 10.1093/hmg/ddy271.
6. Coleman, J.R.I., et al., *Individual and shared effects of social environment and polygenic risk scores on adolescent body mass index.* Scientific reports, 2018. **8**(1): p. 6344-6344 DOI: 10.1038/s41598-018-24774-5.
7. Xie, T., et al., *Genetic Risk Scores for Complex Disease Traits in Youth.* Circulation: Genomic and Precision Medicine, 2020. **13**(4): p. e002775 DOI: 10.1161/CIRCGEN.119.002775.
8. Hüls, A., et al., *Polygenic risk for obesity and its interaction with lifestyle and sociodemographic factors in European children and adolescents.* International Journal of Obesity, 2021. **45**(6): p. 1321-1330 DOI: 10.1038/s41366-021-00795-5.
9. Fang, J., et al., *Polygenic risk, adherence to a healthy lifestyle, and childhood obesity.* Pediatric Obesity, 2019. **14**(4): p. e12489 DOI: https://doi.org/10.1111/ijpo.12489.
10. Khera, A.V., et al., *Polygenic Prediction of Weight and Obesity Trajectories from Birth to Adulthood.* Cell, 2019. **177**(3): p. 587-596.e9 DOI: 10.1016/j.cell.2019.03.028.
11. Vogelezang, S., et al., *Novel loci for childhood body mass index and shared heritability with adult cardiometabolic traits.* PLoS genetics, 2020. **16**(10): p. e1008718-e1008718 DOI: 10.1371/journal.pgen.1008718.

12. Sanz-de-Galdeano, A., A. Terskaya, and A. Upegui, *Association of a genetic risk score with BMI along the life-cycle: Evidence from several US cohorts.* PloS one, 2020. **15**(9): p. e0239067-e0239067 DOI: 10.1371/journal.pone.0239067.
13. Belsky, D.W., et al., *Polygenic risk, rapid childhood growth, and the development of obesity: evidence from a 4-decade longitudinal study.* Archives of pediatrics & adolescent medicine, 2012. **166**(6): p. 515-521 DOI: 10.1001/archpediatrics.2012.131.
14. Locke, A.E., et al., *Genetic studies of body mass index yield new insights for obesity biology.* Nature, 2015. **518**(7538): p. 197-206 DOI: 10.1038/nature14177.
15. Felix, J.F., et al., *Genome-wide association analysis identifies three new susceptibility loci for childhood body mass index.* Human molecular genetics, 2016. **25**(2): p. 389-403 DOI: 10.1093/hmg/ddv472.
16. Turner, J.D., et al., *Twin Research in the Post-Genomic Era: Dissecting the Pathophysiological Effects of Adversity and the Social Environment.* Int J Mol Sci, 2020. **21**(9) DOI: 10.3390/ijms21093142.
17. Bradfield, J.P., et al., *A genome-wide association meta-analysis identifies new childhood obesity loci.* Nature genetics, 2012. **44**(5): p. 526-531 DOI: 10.1038/ng.2247.
18. Helgeland, Ø., et al., *Characterization of the genetic architecture of infant and early childhood body mass index.* Nature Metabolism, 2022 DOI: 10.1038/s42255-022-00549-1.
19. Orri, M., et al., *Cohort Profile: Quebec Longitudinal Study of Child Development (QLSCD).* Social Psychiatry and Psychiatric Epidemiology, 2020 DOI: 10.1007/s00127-020-01972-z.
20. Perret, L.C., et al., *Polygenic risk score and peer victimisation independently predict depressive symptoms in adolescence: results from the Quebec Longitudinal Study of Children Development.* Journal of Child Psychology and Psychiatry, 2022. **n/a**(n/a) DOI: https://doi.org/10.1111/jcpp.13706.
21. Auton, A., et al., *A global reference for human genetic variation.* Nature, 2015. **526**(7571): p. 68-74 DOI: 10.1038/nature15393.
22. Delaneau, O., J.-F. Zagury, and J. Marchini, *Improved whole-chromosome phasing for disease and population genetic studies.* Nature Methods, 2013. **10**(1): p. 5-6 DOI: 10.1038/nmeth.2307.
23. Howie, B.N., P. Donnelly, and J. Marchini, *A Flexible and Accurate Genotype Imputation Method for the Next Generation of Genome-Wide Association Studies.* PLOS Genetics, 2009. **5**(6): p. e1000529 DOI: 10.1371/journal.pgen.1000529.

24. Ge, T., et al., *Polygenic prediction via Bayesian regression and continuous shrinkage priors.* Nature communications, 2019. **10**(1): p. 1776-1776 DOI: 10.1038/s41467-019-09718-5.
25. Zhao, Z., et al., *PUMAS: fine-tuning polygenic risk scores with GWAS summary statistics.* Genome Biology, 2021. **22**(1): p. 257 DOI: 10.1186/s13059-021-02479-9.
26. Khera, A.V., et al., *Whole-Genome Sequencing to Characterize Monogenic and Polygenic Contributions in Patients Hospitalized With Early-Onset Myocardial Infarction.* Circulation, 2019. **139**(13): p. 1593-1602 DOI: 10.1161/circulationaha.118.035658.
27. de Onis, M., et al., *Development of a WHO growth reference for school-aged children and adolescents.* Bulletin of the World Health Organization, 2007. **85**(9): p. 660-667 DOI: 10.2471/blt.07.043497.
28. Hartig, F., *_DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models_.* R package version 0.4.6, 2022, <https://CRAN.R-project.org/package=DHARMa>.
29. Green, P. and C.J. MacLeod, *SIMR: an R package for power analysis of generalized linear mixed models by simulation.* Methods in Ecology and Evolution, 2016. **7**(4): p. 493-498 DOI: https://doi.org/10.1111/2041-210X.12504.
30. Abdulkadir, M., et al., *Polygenic Score for Body Mass Index Is Associated with Disordered Eating in a General Population Cohort.* J Clin Med, 2020. **9**(4) DOI: 10.3390/jcm9041187.
31. R Core Team, *R: A language and environment for statistical computing*. 2021, R Foundation for Statistical Computing: Vienna, Austria, https://www.R-project.org/.
32. Loos, R.J.F. and G.S.H. Yeo, *The genetics of obesity: from discovery to biology.* Nature Reviews Genetics, 2022. **23**(2): p. 120-133 DOI: 10.1038/s41576-021-00414-z.
33. Jansen, P.R., et al., *The utility of obesity polygenic risk scores from research to clinical practice: A review.* Obesity Reviews, 2024. **25**(11): p. e13810 DOI: https://doi.org/10.1111/obr.13810.
34. Aroon, D.H., et al., *Performance of polygenic risk scores in screening, prediction, and risk stratification: secondary analysis of data in the Polygenic Score Catalog.* BMJ Medicine, 2023. **2**(1): p. e000554 DOI: 10.1136/bmjmed-2023-000554.
35. Lewis, A.C.F., et al., *Patient and provider perspectives on polygenic risk scores: implications for clinical reporting and utilization.* Genome Medicine, 2022. **14**(1): p. 114 DOI: 10.1186/s13073-022-01117-8.
36. O'Donovan, C.B., et al., *Knowing your genes: does this impact behaviour change?* Proceedings of the Nutrition Society, 2017. **76**(3): p. 182-191 DOI: 10.1017/S0029665116002949.

37. Brandkvist, M., et al., *Separating the genetics of childhood and adult obesity: a validation study of genetic scores for body mass index in adolescence and adulthood in the HUNT Study.* Human Molecular Genetics, 2020. **29**(24): p. 3966-3973 DOI: 10.1093/hmg/ddaa256.
38. Richardson, T.G., et al., *Use of genetic variation to separate the effects of early and later life adiposity on disease risk: mendelian randomisation study.* BMJ (Clinical research ed.), 2020. **369**: p. m1203-m1203 DOI: 10.1136/bmj.m1203.
39. Cai, J., et al., *Priors, population sizes, and power in genome-wide hypothesis tests.* BMC Bioinformatics, 2023. **24**(1): p. 170 DOI: 10.1186/s12859-023-05261-9.
40. Wu, T., et al., *Polygenic power calculator: Statistical power and polygenic prediction accuracy of genome-wide association studies of complex traits.* Front Genet, 2022. **13**: p. 989639 DOI: 10.3389/fgene.2022.989639.
41. Link, J.C. and K. Reue, *Genetic Basis for Sex Differences in Obesity and Lipid Metabolism.* Annu Rev Nutr, 2017. **37**: p. 225-245 DOI: 10.1146/annurev-nutr-071816-064827.
42. Song, M., et al., *Longitudinal Analysis of Genetic Susceptibility and BMI Throughout Adult Life.* Diabetes, 2017. **67**(2): p. 248-255 DOI: 10.2337/db17-1156.
43. Shields, M., et al., *Obesity estimates for children based on parent-reported versus direct measures.* Health Rep, 2011. **22**(3): p. 47-58.
44. Verlouw, J.A.M., et al., *A comparison of genotyping arrays.* European Journal of Human Genetics, 2021. **29**(11): p. 1611-1624 DOI: 10.1038/s41431-021-00917-7.

Tables and figures

Table 4-1. Characteristics of participants (n=717)

Characteristic	*% (n) or mean ± SD*
Sex, female	55.2 (396)
Birth weight, g	3418.1 ± 487.9
Preterm birth	4.9 (35)
First child	41.6 (298)
Maternal age at birth	
≤ 20 years	2.0 (14)
20-34 years	84.0 (602)
≥ 35 years	14.1 (101)
Maternal education	
< Secondary school diploma	15.9 (114)
Secondary school diploma	24.1 (173)
Post-sec. except university	29.4 (211)
University diploma	30.5 (219)
Household income	
< 30 000$	25.0 (178)
30 000 – <60 000$	41.2 (293)
60 000 – <80 000$	17.2 (122)
≥ 80 000$	16.6 (118)
Born in Canada	97.8 (701)

Baseline characteristics of study participants. N: sample size, SD: standard deviation.

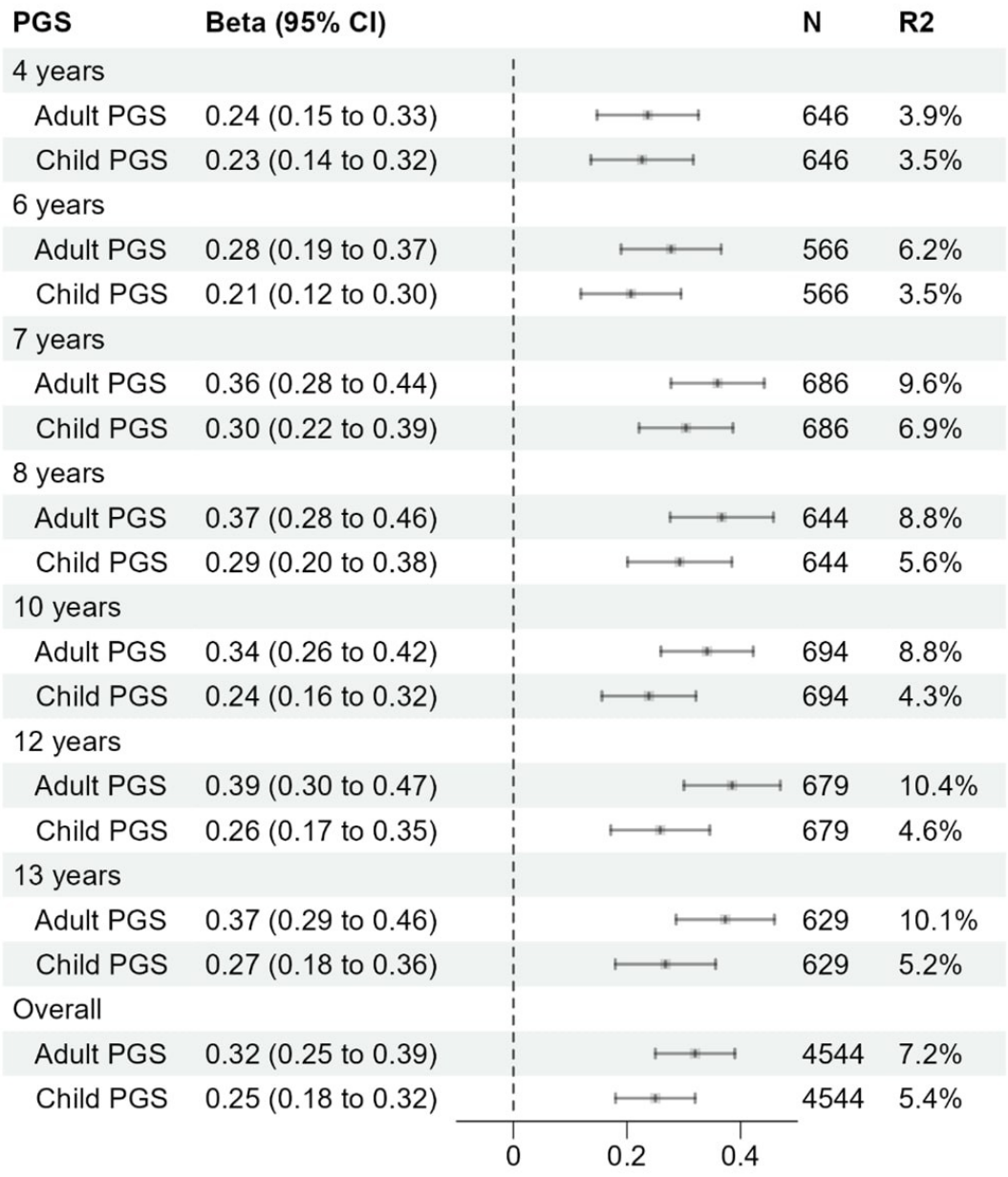


Figure 4-1. Linear association between PGS and BMI z-scores, 4 to 13 years old. Beta ± 95% CI. N: sample size, R2: coefficient of determination, PGS: polygenic score, BMI: body mass index**.**

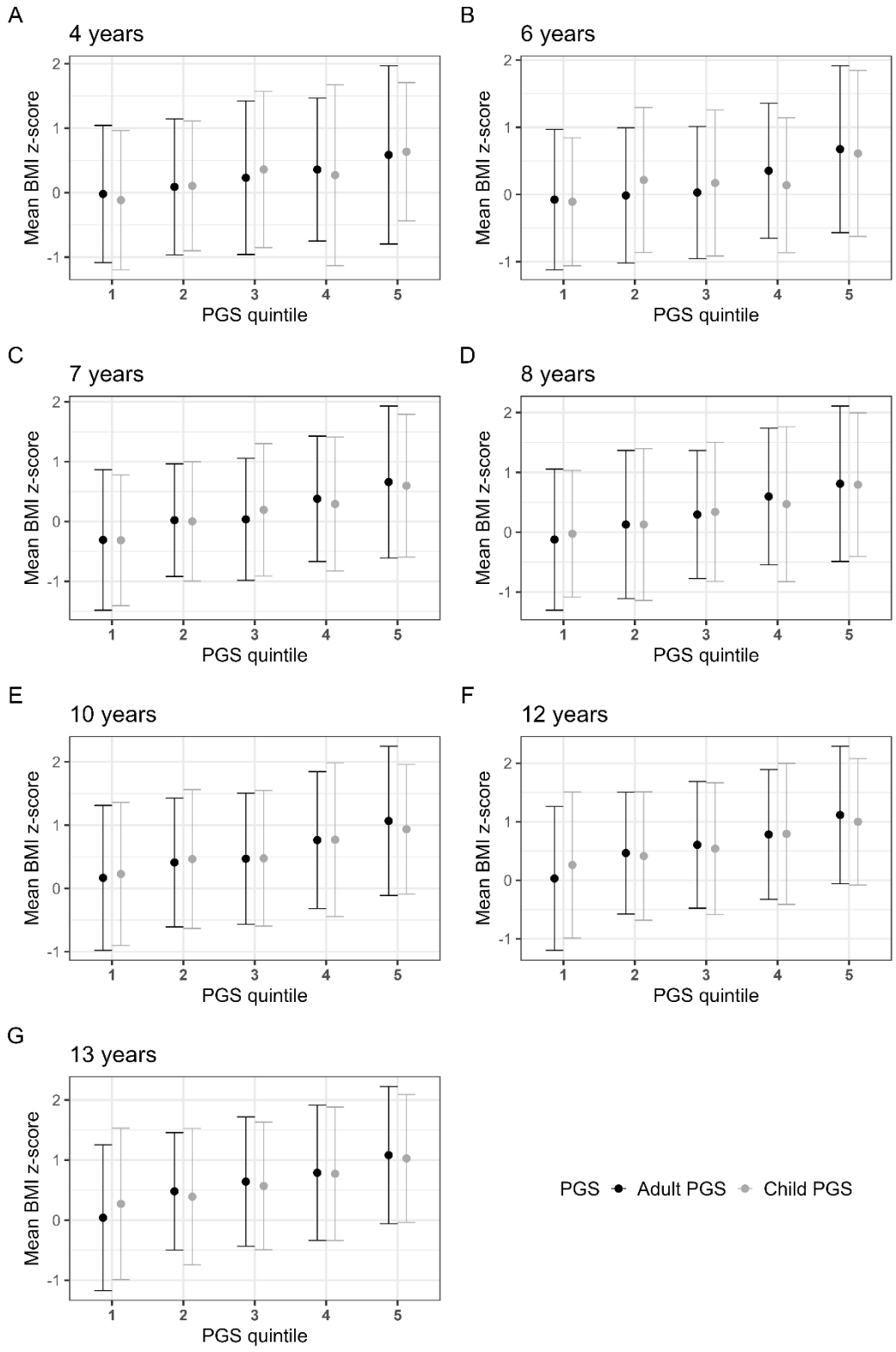


Figure 4-2. Mean BMI z-scores per PGS quintile at (A) 4 years, (B) 6 years, (C) 7 years, (D) 8 years, (E) 10 years, (F) 12 years, and (G) 13 years. PGS: polygenic score, BMI: body mass index.

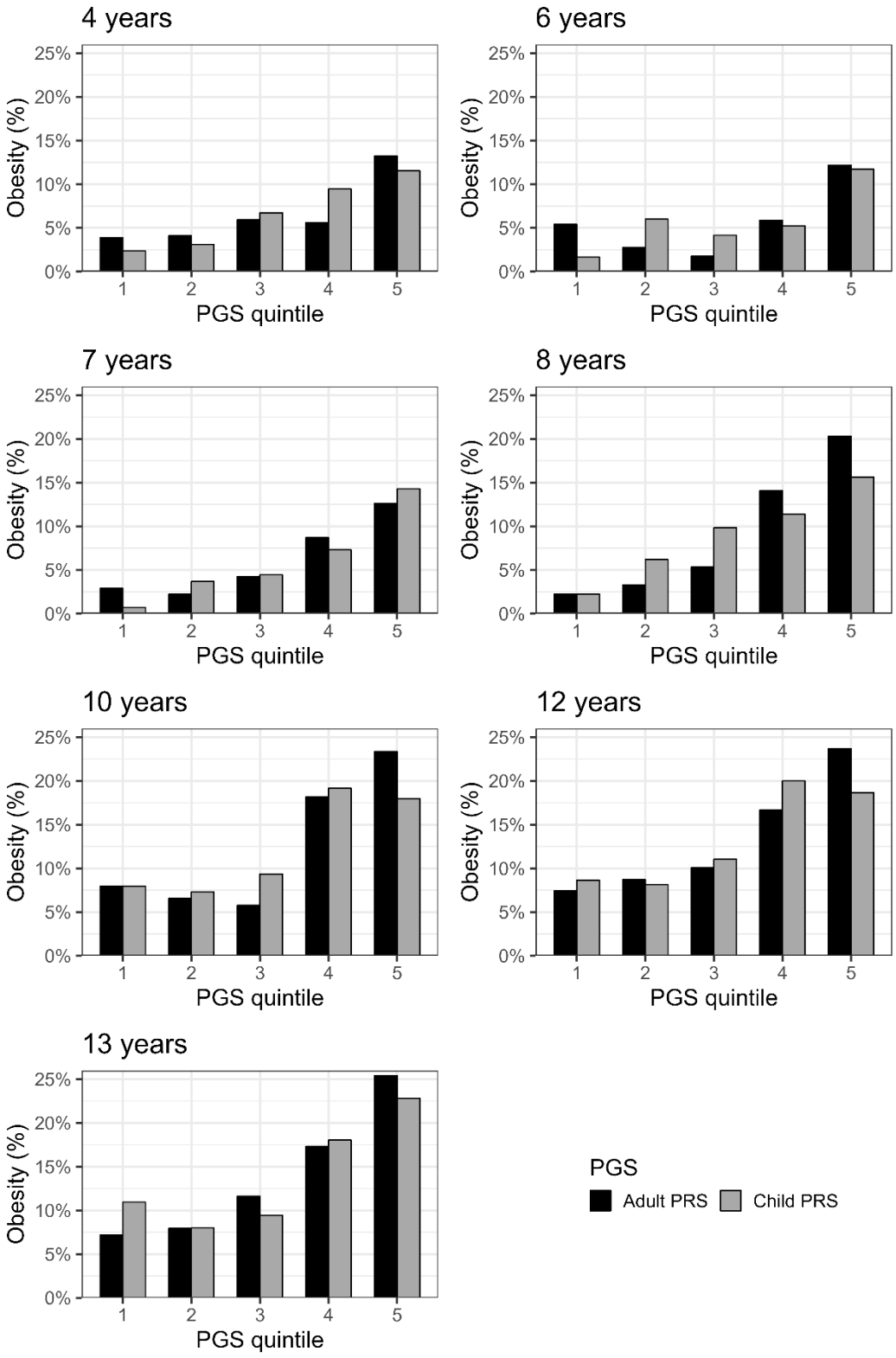


Figure 4-3. Proportion of obesity per PGS quintile at (A) 4 years, (B) 6 years, (C) 7 years, (D) 8 years, (E) 10 years, (F) 12 years, and (G) 13 years. PGS: polygenic score.

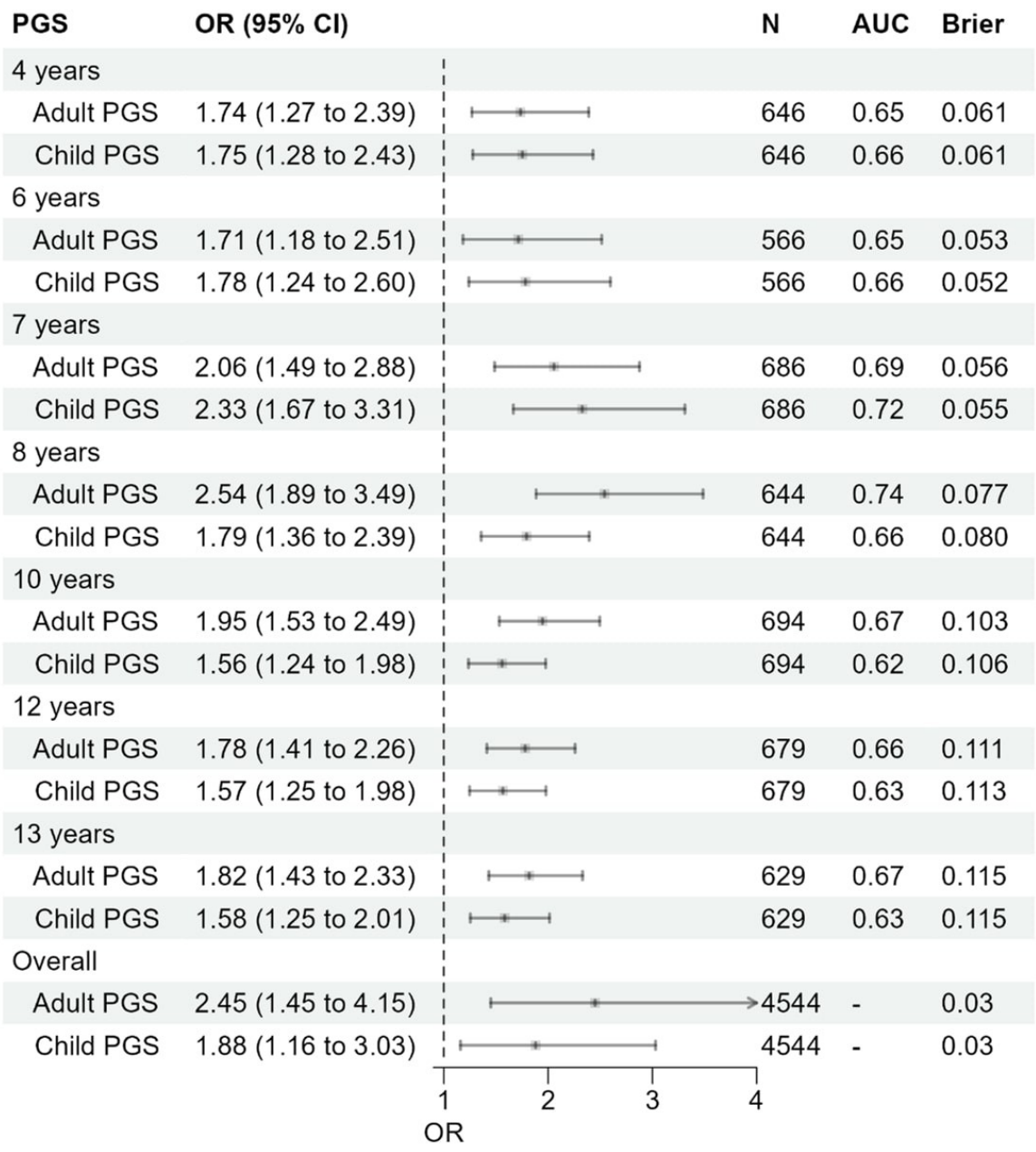


Figure 4-4. Obesity odds ratio for adult and child PGS, 4 to 13 years old. **N: sample size, AUC: area under the receiver operator curve, Brier: Brier score, PGS: polygenic score, BMI: body mass index.**

Supplementary materials

Supplementary Tables

Supplementary Table 4-1. Characteristics of participants, included vs. excluded.

Characteristic ***% (n) or mean ± SD***	***Included*** ***N=717***	***Excluded*** ***N=1403***	***p-value***
Sex, female	55.2 (396)	45.9 (644)	**<0.001**
Birth weight, g	3.42 ± 0.49	3.40 ± 0.50	0.348
Preterm birth	4.9 (35)	4.6 (64)	0.825
First child	41.6 (298)	45.5 (638)	0.095
Maternal age at birth			0.220
≤ 20 years	2.0 (14)	3.2 (45)	
20-34 years	84.0 (602)	83.7 (1173)	
≥ 35 years	14.1 (101)	13.1 (184)	
Maternal education			**0.012**
< Secondary school diploma	15.9 (114)	19.4 (271)	
Secondary school diploma	24.1 (173)	27.3 (382)	
Post-sec. except university	29.4 (211)	28.6 (400)	
University diploma	30.5 (219)	24.8 (347)	
Household income			**0.001**
< 30 000$	25.0 (178)	32.4 (444)	
30 000 – <60 000$	41.2 (293)	40.1 (550)	
60 000 – <80 000$	17.2 (122)	15.0 (205)	
≥ 80 000$	16.6 (118)	12.5 (172)	
Born in Canada	97.8 (701)	83.1 (1164)	**<0.001**

N: sample size, SD: standard deviation. Statistical differences between included and excluded participants assessed with Chi-square test for proportions, and one-way ANOVA for normally distributed variables.

Supplementary Table 4-2. PGS association with BMI z-score by age.

Model	Beta (95% CI)	p-value	R^2	Cor
4 years	N=646			
Adult PGS	0.24 (0.15, 0.33)	**<0.001**	0.039	0.20
Child PGS	0.23 (0.14, 0.32)	**<0.001**	0.035	0.19
6 years	N=566			
Adult PGS	0.28 (0.19, 0.37)	**<0.001**	0.062	0.25
Child PGS	0.21 (0.12, 0.30)	**<0.001**	0.035	0.19
7 years	N=686			
Adult PGS	0.36 (0.28, 0.44)	**<0.001**	0.096	0.31
Child PGS	0.30 (0.22, 0.39)	**<0.001**	0.069	0.27
8 years	N=644			
Adult PGS	0.37 (0.28, 0.46)	**<0.001**	0.088	0.30
Child PGS	0.29 (0.20, 0.38)	**<0.001**	0.056	0.24
10 years	N=694			
Adult PGS	0.34 (0.26, 0.42)	**<0.001**	0.088	0.30
Child PGS	0.24 (0.16, 0.32)	**<0.001**	0.043	0.21
12 years	N=679			
Adult PGS	0.39 (0.30, 0.47)	**<0.001**	0.104	0.32
Child PGS	0.26 (0.17, 0.35)	**<0.001**	0.046	0.22
13 years	N=629			
Adult PGS	0.37 (0.29, 0.46)	**<0.001**	0.101	0.32
Child PGS	0.27 (0.18, 0.36)	**<0.001**	0.052	0.23
Overall	N=4544 observations			
Adult PGS	0.32 (0.25, 0.39)	**<0.001**	0.072	
Child PGS	0.25 (0.18, 0.32)	**<0.001**	0.054	

R^2, adjusted proportion of BMI z-score variance explained by the PGS; Cor, Pearson's correlation between BMI z-score predicted by the model and observed values; N, sample size at specific time point. From 4 to 13 years old, estimated effect of the increase of 1 standard deviation of PGS on BMI z-score with 95% CI using simple linear regression. For overall model, linear mixed model is used to account for all repeated observations. Correlation is not reported for linear mixed model because clustering by participant resulted in near 1.0 correlation.

Supplementary Table 4-3. Obesity prediction for PGS, by age.

Model	OR (95% CI)	p-value	R^2	AUC
4 years	N=646			
Adult PGS	1.74 (1.27, 2.39)	**<0.001**	0.038	0.65
Child PGS	1.75 (1.28, 2.43)	**<0.001**	0.039	0.66
6 years	N=566			
Adult PGS	1.71 (1.18, 2.51)	**0.005**	0.033	0.65
Child PGS	1.78 (1.24, 2.60)	**0.002**	0.040	0.66
7 years	N=686			
Adult PGS	2.06 (1.49, 2.88)	**<0.001**	0.061	0.69
Child PGS	2.33 (1.67, 3.31)	**<0.001**	0.082	0.72
8 years	N=644			
Adult PGS	2.54 (1.89, 3.49)	**<0.001**	0.104	0.74
Child PGS	1.79 (1.36, 2.39)	**<0.001**	0.044	0.66
10 years	N=694			
Adult PGS	1.95 (1.53, 2.49)	**<0.001**	0.059	0.67
Child PGS	1.56 (1.24, 1.98)	**<0.001**	0.028	0.62
12 years	N=679			
Adult PGS	1.78 (1.341, 2.26)	**<0.001**	0.046	0.66
Child PGS	1.57 (1.25, 1.98)	**<0.001**	0.028	0.63
13 years	N=629			
Adult PGS	1.82 (1.43, 2.33)	**<0.001**	0.049	0.67
Child PGS	1.58 (1.25, 2.01)	**<0.001**	0.030	0.63
Overall	N=4544 observations			
Adult PGS	2.45 (1.45, 4.15)	**<0.001**	0.031	
Child PGS	1.88 (1.16, 3.03)	**0.010**	0.022	

R^2, pseudo R^2 for the PGS variable; AUC, area under the receiver operator curve; N, sample size at specific time point. From 4 to 13 years old, estimated odds ratio of obesity for the increase in 1 standard deviation in PGS with 95% CI using logistic regression. For overall model, generalized linear mixed model is used to account for all repeated observations. AUC is not reported for generalized linear mixed model because clustering by participant resulted in near 1.0 AUC.

Supplementary Table 4-4. PGS association with BMI z-score by age, stratified by sex.

Model	Female		Male		
	Beta (95% CI)	**p-value**	**Beta (95% CI)**	**p-value**	**P PGS x Sex**
4 years	N=358		N=288		
Adult PGS	0.16 (0.04, 0.28)	**0.011**	0.33 (0.16, 0.46)	**<0.001**	0.065
Child PGS	0.18 (0.06, 0.30)	**0.004**	0.30 (0.16, 0.43)	**<0.001**	0.210
6 years	N=320		N=246		
Adult PGS	0.19 (0.07, 0.31)	**0.002**	0.39 (0.19, 0.46)	**<0.001**	**0.027**
Child PGS	0.16 (0.04, 0.27)	**0.004**	0.26 (0.13, 0.40)	**<0.001**	0.232
7 years	N=381		N=305		
Adult PGS	0.28 (0.17, 0.39)	**<0.001**	0.46 (0.33, 0.58)	**<0.001**	**0.036**
Child PGS	0.26 (0.15, 0.37)	**<0.001**	0.36 (0.24, 0.49)	**<0.001**	0.230
8 years	N=352		N=292		
Adult PGS	0.29 (0.17, 0.40)	**<0.001**	0.48 (0.34, 0.62)	**<0.001**	**0.038**
Child PGS	0.24 (0.12, 0.37)	**<0.001**	0.35 (0.22, 0.49)	**<0.001**	0.235
10 years	N=384		N=310		
Adult PGS	0.26 (0.15, 0.37)	**<0.001**	0.44 (0.32, 0.56)	**<0.001**	**0.031**
Child PGS	0.23 (0.11, 0.34)	**<0.001**	0.25 (0.12, 0.37)	**<0.001**	0.828
12 years	N=377		N=302		
Adult PGS	0.31 (0.20, 0.42)	**<0.001**	0.48 (0.35, 0.61)	**<0.001**	0.052
Child PGS	0.23 (0.11, 0.45)	**<0.001**	0.29 (0.16, 0.42)	**<0.001**	0.491
13 years	N=358		N=271		
Adult PGS	0.29 (0.17, 0.40)	**<0.001**	0.49 (0.35, 0.62)	**<0.001**	**0.026**
Child PGS	0.22 (0.10, 0.34)	**<0.001**	0.32 (0.19, 0.45)	**<0.001**	0.273
Overall	N=2533		N=2014		
Adult PGS	0.23 (0.14, 0.32)	**<0.001**	0.43 (0.33, 0.53)	**<0.001**	**0.005**
Child PGS	0.21 (0.11, 0.30)	**<0.001**	0.31 (0.21, 0.41)	**<0.001**	0.132

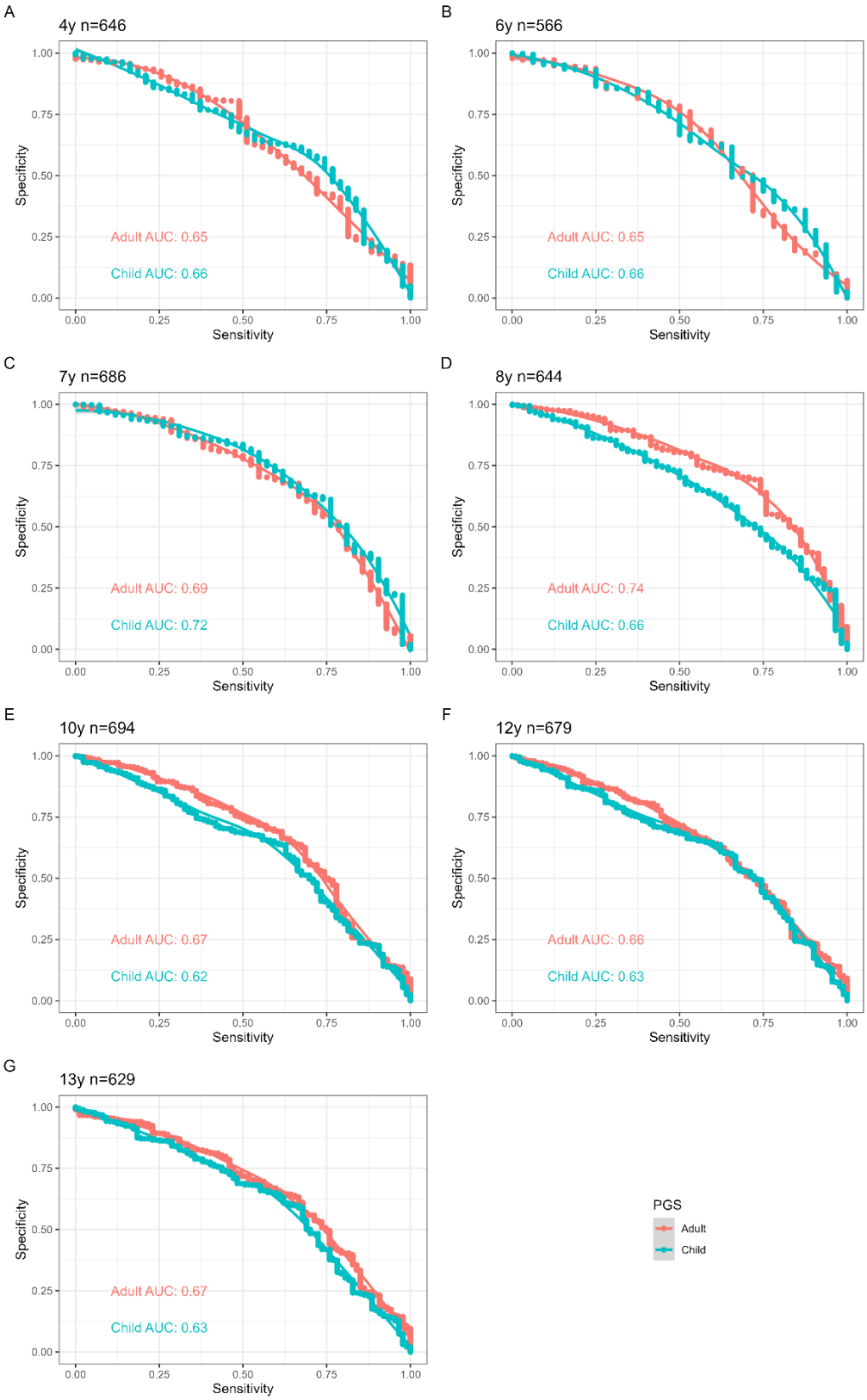
P PGS x Sex, p-value for the interaction term between PGS and sex; N, sample size at specific time point. From 4 to 13 years old, estimated marginal effect of the increase of 1 standard deviation of PGS on BMI z-score with 95% CI using simple linear regression. For overall model, linear mixed model is used to account for all repeated observations.

Supplementary Table 4-5. Obesity odds ratio by age, stratified by sex.

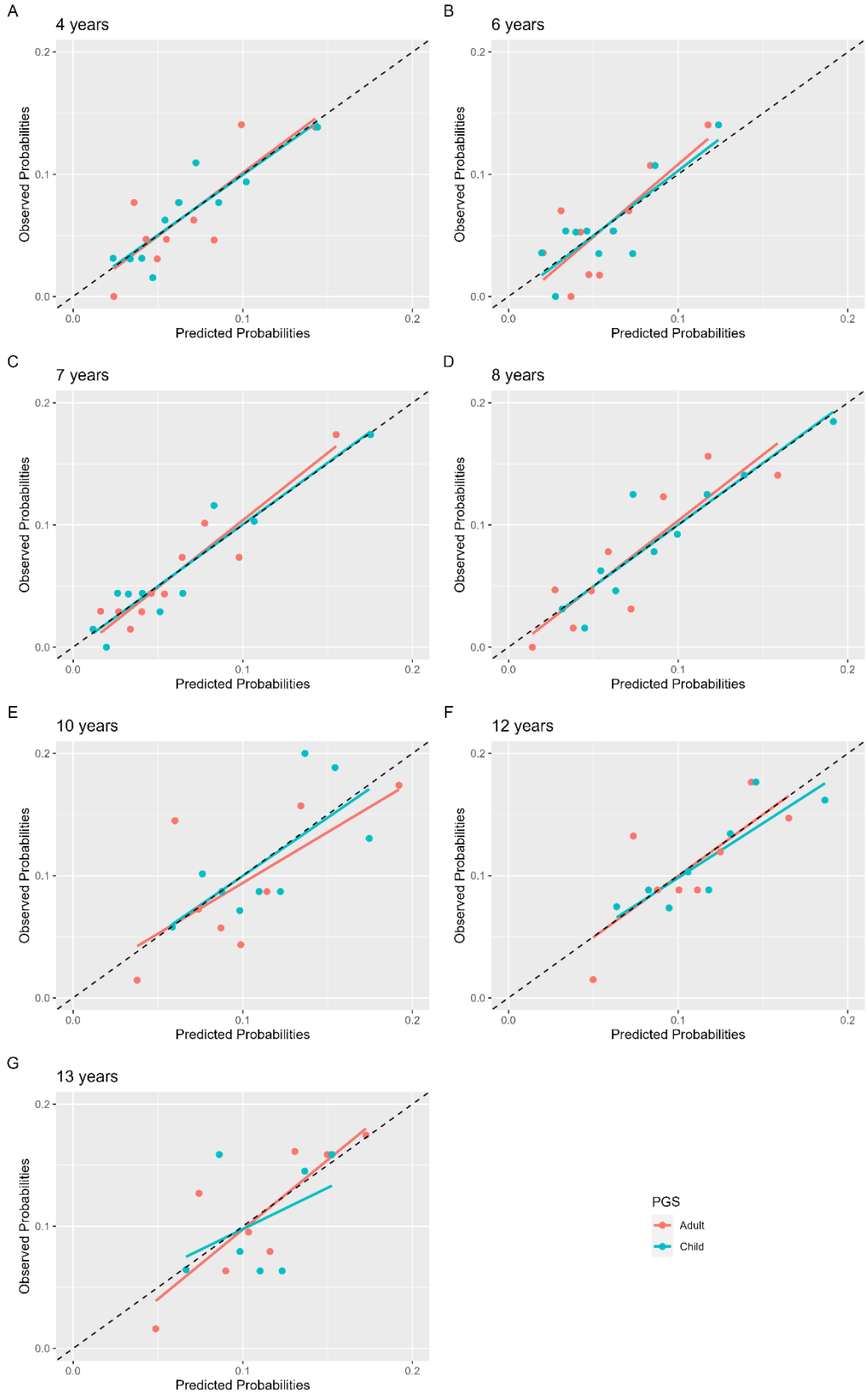
Model	Female		Male		
	OR (95% CI)	**p-value**	**OR (95% CI)**	**p-value**	**P PGS x Sex**
4 years	N=358		N=288		
Adult PGS	1.33 (0.86, 2.06)	0.201	2.37 (1.48, 3.97)	**<0.001**	0.085
Child PGS	1.76 (1.12, 2.80)	**0.016**	1.79 (1.14, 2.87)	**0.014**	0.959
6 years	N=320		N=246		
Adult PGS	1.43 (0.84, 2.44)	0.184	2.03 (1.19, 3.62)	**0.012**	0.371
Child PGS	1.60 (0.94, 2.75)	0.085	2.05 (1.22, 3.55)	**0.008**	0.516
7 years	N=381		N=305		
Adult PGS	2.02 (1.32, 3.13)	**0.001**	2.11 (1.28, 3.62)	**0.005**	0.893
Child PGS	2.80 (1.76, 4.64)	**<0.001**	2.05 (1.26, 3.48)	**0.005**	0.385
8 years	N=352		N=292		
Adult PGS	2.59 (1.72, 4.05)	**<0.001**	2.61 (1.69, 4.22)	**<0.001**	0.979
Child PGS	2.26 (1.49, 3.56)	**<0.001**	1.50 (1.03, 2.22)	**0.036**	0.163
10 years	N=384		N=310		
Adult PGS	1.86 (1.32, 2.67)	**<0.001**	2.09 (1.49, 3.01)	**<0.001**	0.648
Child PGS	1.68 (1.18, 2.44)	**0.005**	1.46 (1.08, 2.01)	**0.017**	0.560
12 years	N=377		N=302		
Adult PGS	1.85 (1.31, 2.66)	**<0.001**	1.75 (1.28, 2.45)	**<0.001**	0.821
Child PGS	1.72 (1.20, 2.51)	**0.004**	1.48 (1.10, 2.01)	**0.011**	0.525
13 years	N=358		N=271		
Adult PGS	2.01 (1.38, 2.97)	**<0.001**	1.75 (1.27, 2.46)	**<0.001**	0.590
Child PGS	1.59 (1.09, 2.35)	**0.017**	1.60 (1.18, 2.20)	**0.003**	0.987
Overall	N=2533		N=2014		
Adult PGS	2.13 (1.08, 4.18)	**0.029**	3.39 (1.50, 7.67)	**<0.001**	0.386
Child PGS	1.82 (0.93, 3.58)	0.0823	2.10 (1.06, 4.18)	**<0.001**	0.769

P PGS x Sex, p-value for the interaction term between PGS and sex; N, sample size at specific time point. From 4 to 13 years old, estimated marginal odds ratio of obesity for the increase in 1 standard deviation in PGS with 95% CI using logistic regression. For overall model, generalized linear mixed model is used to account for all repeated observations.

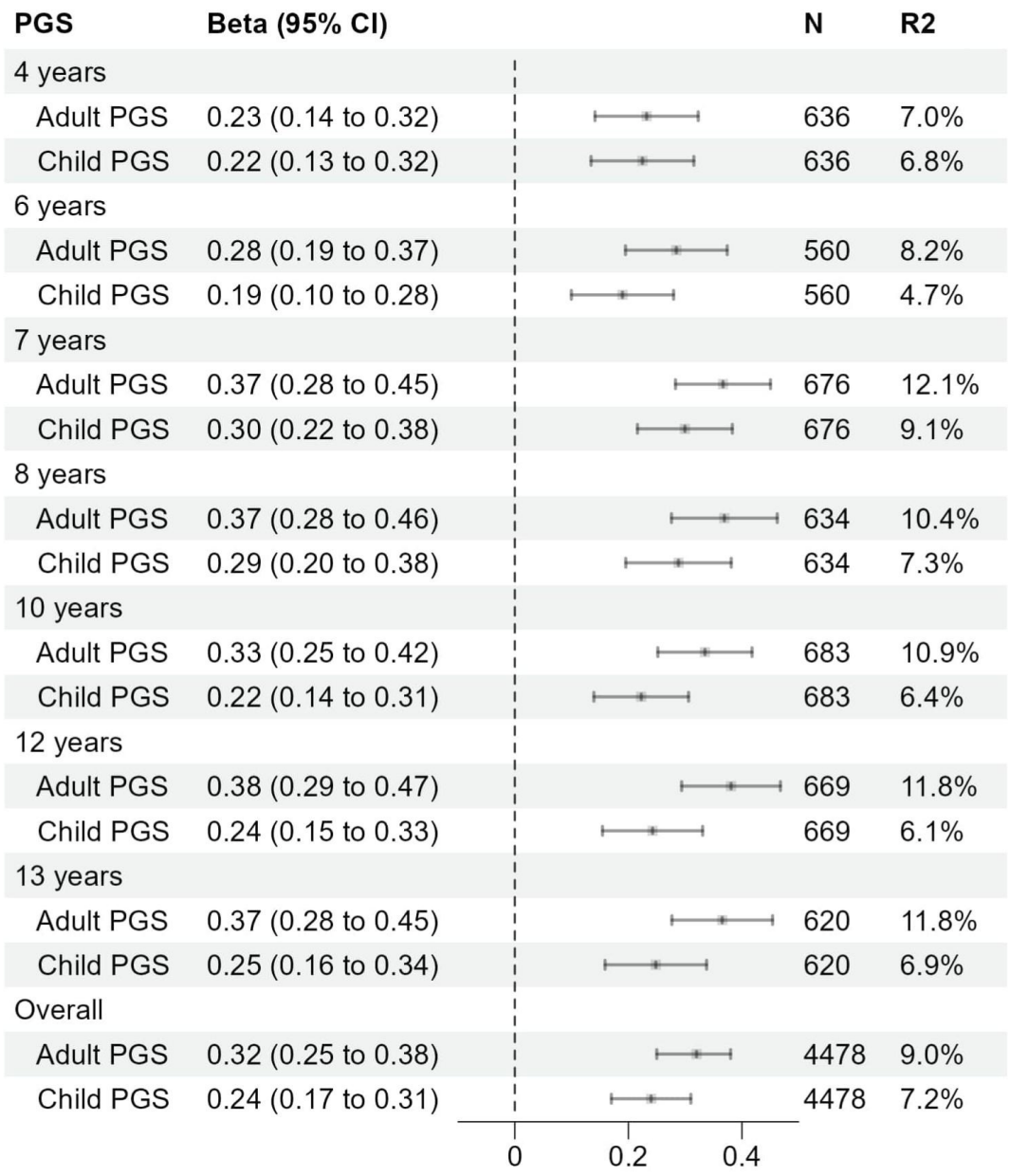
Supplementary Figures



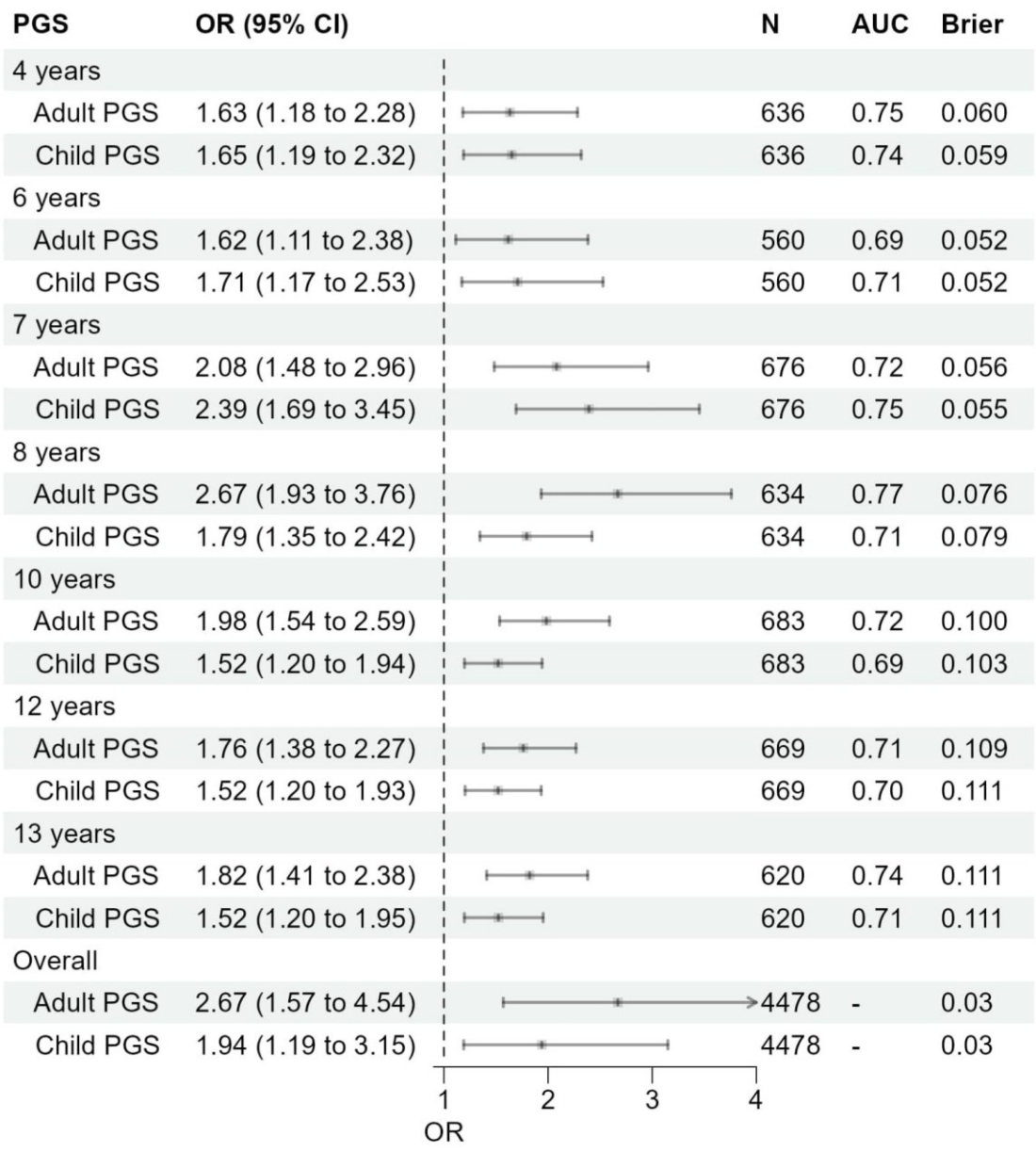
Supplementary Figure 4-1. AUC plot, logistic regression, QLSCD 4-13 years.



Supplementary Figure 4-2. Calibration plot, logistic regression, QLSCD 4-13 years.



Supplementary Figure 4-3. Linear association between PGS and BMI z-scores adjusted for covariates, 4 to 13 years old. **Beta ± 95% CI. N: sample size, R2: coefficient of determination, PGS: polygenic score, BMI: body mass index.**



Supplementary Figure 4-4. Obesity odds ratio for adult and child PGS adjusted for covariates, 4 to 13 years old. N: sample size, AUC: area under the receiver operator curve, Brier: Brier score, PGS: polygenic score, BMI: body mass index.

Supplementary Text 4-1. R script for principal components adjusted scores.

```
require("broom")  # tidy regression output
require("mosaic")  # standardizing variables

load(file="BMI Yengo Vogel scores and BMI")
load(file="PCA717.Rdata")

PCA717 <- PCA717[, 1:11]
dataPRS <- merge(dataPRS, PCA717, by="IID")

# Adult score
Yengo=lm(PRS_Yengo~PC1+PC2+PC3+PC4+PC5+PC6+PC7+PC8+PC9+PC10,data=dataPRS)

dataPRS$Adultpred=predict(Yengo,dataPRS)
dataPRS$PRS_Yengo_adj = dataPRS$PRS_Yengo-dataPRS$Adultpred
dataPRS <- dataPRS %>%
  mutate(PRS_Yengo_adj = scale(PRS_Yengo_adj))

# Child score
Vogel=lm(PRS_Vogel~PC1+PC2+PC3+PC4+PC5+PC6+PC7+PC8+PC9+PC10,data=dataPRS)

dataPRS$Childpred=predict(Vogel,dataPRS)
dataPRS$PRS_Vogel_adj = dataPRS$PRS_Vogel-dataPRS$Childpred
dataPRS <- dataPRS %>%
  mutate(PRS_Vogel_adj = scale(PRS_Vogel_adj))
```

CHAPTER 5 - MANUSCRIPT 2

Article preface

After introducing the genetic component of obesity with PGS in the first manuscript, this second manuscript adds a behavioural component by delving into the role of fussy eating and over-eating behaviours expressed from 2 to 6 years in translating genetically induced differences in BMI from 6 to 13 years in the QLSCD. Incidentally, the second manuscript relates to the second objective of the thesis: to determine the extent to which the association between adult- and child-derived PGS and BMI is mediated through eating behavioural traits (over-eating, fussy eating) assessed in early childhood (i.e., preschool age).

DG, LD and MB contributed to the development of the research question and the design of the study. LD and MB were involved in the data collection from QLSCD participants. DG performed the statistical analyses with insights and revisions from CG. DG was responsible for the first draft of the manuscript. All authors participated in discussions concerning the development of the study and were involved in the revisions to the manuscript and approved the submitted version.

Ethics approval (secondary analysis of QLSCD data): University of Ottawa Research Ethics board (ethics file number: H-01-23-8018, approval date: 10/02/2023, Appendix 1).

This manuscript was published in the journal Pediatric Obesity on October 10th, 2024. The following manuscript is harmonized to this thesis format. The original published manuscript is available in Appendix 5, alongside the proof of authorization to reuse the material in the thesis in Appendix 6. A STROBE reporting checklist for manuscript 2 is available in Appendix 7. The article should be cited as:

Goulet D, Boivin M, Gravel CA, Little J, Potter BK, Dubois L. Mediation of genetic susceptibility to obesity through eating behaviours in children. *Pediatric Obesity*. 2024;e13180. doi:10.1111/ijpo.13180.

Title page

Title: Mediation of genetic susceptibility to obesity through eating behaviours in children

Danick Goulet[1*], Michel Boivin[2], Christopher A. Gravel[1, 3, 4], Julian Little[1], Beth K Potter[1], Lise Dubois[1]

[1]School of Epidemiology and Public Health, University of Ottawa, Ottawa, ON, Canada

[2]Université Laval, Québec, QC, Canada.

[3]Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montréal, QC, Canada

[4]Department of Mathematics and Statistics, University of Ottawa, Ottawa, ON, Canada

Correspondence:

Lise Dubois : ldubois@uottawa.ca

Adress: 600 Peter Morand Crescent, Ottawa (ON), Canada

Michel Boivin : michel.boivin@psy.ulaval.ca

Word count (main text): 5021

Word count (abstract): 200

Number of tables: 3

Number of figures: 3

Number of Supplementary materials: 3

Abbreviations: BMI – body-mass index; PGS – polygenic score; QLSCD – Quebec Longitudinal Study of Children Development; LGCMA – Longitudinal Growth Curve Mediation Analysis; GWAS – Genome-Wide Association Study; IPW – Inverse probability weighting.

Keywords: Obesity, polygenic score, mediation, eating behaviours, children's health, longitudinal analysis.

Running title: Mediation of PGS for BMI through eating behaviours.

Abstract

Background/Objectives: Few studies have examined the putative mediating role of eating behaviours linking genetic susceptibility and body weight. The goal of this study was to investigate the extent to which two polygenic scores (PGSs) for body mass index (BMI), based on child and adult data, predicted BMI through over-eating and fussy eating across childhood.

Subjects/Methods: The study sample involved 692 participants from a birth-cohort study. Height and weight were measured on six occasions between ages 6 and 13 years. Over-eating and fussy eating behaviours were assessed five times between ages 2 and 6 years. Longitudinal growth curve mediation analysis was used to estimate the contributions of the PGSs to BMI z-scores mediated by over-eating and fussy eating.

Results: Both PGSs predicted BMI z-scores (PGS_{child}: β=0.26, 95% CI: 0.19-0.33; PGS_{adult}: β=0.34, 95% CI: 0.27-0.41). Over-eating significantly mediated these associations, but this mediation decreased over time from 6 years (PGS_{child}: 18.0%, 95% CI: 3.1-32.9, p-value=0.018; PGS_{adult}: 14.2%, 95% CI: 2.8-25.5, p-value=0.014) to 13 years (PGS_{child}: 11.4%, 95% CI: -0.4-23.1, p-value=0.057; PGS_{adult}: 6.2%, 95% CI: 0.4-12.0, p-value=0.037). Fussy eating did not show any mediation.

Conclusions: Our results support the view that appetite is key to translating genetic susceptibility into changes in body weight.

Introduction

Since the World Health Organization declared obesity as a global epidemic in 1997 [1], the prevalence of obesity has continued to rise worldwide. Most worrisome is the rise of obesity in childhood. Children with overweight or obesity between the ages of 5 and 11 years accounted for more than a quarter of Canadian children between 2012 and 2013 [2]. This is especially problematic because childhood obesity tends to persist into adulthood. Compared to children of a healthy weight, those with obesity are estimated to have a fivefold increase in risk of presenting with obesity in adulthood [3]. Accordingly, understanding the factors underlying obesity in childhood is a public health priority.

A substantial portion of the observed variability in individual body weight can be accounted for by variation in genetic susceptibility to weight gain. According to twin studies, genetic factors are responsible for 40% to 80% of the variance in body mass index (BMI) in populations of European ancestry [4]. This strong genetic underpinning of BMI can already be observed in early childhood. Dozens of genetic variants have been found to be associated with childhood BMI, each with small independent contributions [5, 6]. Combining those individual genetic risks results in a significant association with BMI in childhood [5].

According to the Behavioural susceptibility theory, the genes involved in shaping obesity phenotypes exert their influence mainly through neurological processes regulating appetite, which are translated through various eating behaviours [7]. This is consistent with the observation that many genetic variants associated with BMI are involved in brain regions related to behavioural control [8]. Large-scale genome-wide association studies (GWAS) have revealed that variants associated with BMI are predominantly expressed in the central nervous system in regions involved in cognition and emotion regulation (e.g. hippocampus and limbic system), and in areas responsible for appetite regulation (e.g. hypothalamus and pituitary gland) [9, 10]. Empirical data support the view that appetite has a strong role in obesity and that appetite is characterized by a strong genetic component as shown by restrained, uncontrolled, and emotional eating having moderate heritability [8]. However, due to the scarcity of genetically informed studies in childhood, there is limited empirical evidence on the possible role of eating behaviours as mediating genetic susceptibility to weight gain.

Observational studies have leveraged genotype data to examine the degree to which eating behaviours (e.g., satiety responsiveness, appetite, uncontrolled eating, emotional eating, susceptibility to hunger),

usually assessed by questionnaires, mediate the association between genetic susceptibility, assessed through polygenic scores (PGSs), and BMI in adults. For instance, uncontrolled eating and emotional eating partly mediated the association between three different PGSs and BMI in French, UK, and Finnish cohorts [11, 12]. Additionally, disinhibition mediated the association between a PGS and BMI in French Canadian adults [13] and UK adults [14]. Although investigated in all three cohorts, susceptibility to hunger was only identified as a mediator in the French Canadian study [13]. These studies used PGSs comprising <100 variants. Additionally, another study observed that the association between a PGS comprised of close to one million variants and BMI was mediated through infrequent and unhealthy eating, emotional and external eating, and snacking [15].

Three studies have examined the role of behavioural traits related to appetite in the genetic susceptibility of BMI in children. These studies used the Child Eating Behaviour Questionnaire, one of the most commonly used questionnaires assessing behavioural traits related to appetite [7]. The questionnaire measures eight behaviours related to appetite that can be categorized into food approach (food responsiveness, enjoyment of food, emotional over-eating, and desire to drink) and food avoidance (satiety responsiveness, food fussiness, emotional undereating, and slowness in eating) [16]. Satiety responsiveness was found to mediate the association between a 28 variant PGS and BMI in 2258 children from the UK (mean age = 10 years old) [17]. In contrast, satiety responsiveness did not mediate the effect of the 32 variant PGS on BMI in younger (4 to 8-year-old) Norwegian children [18]. Food responsiveness, emotional over-eating, enjoyment of food and slowness in eating were also not found to be mediators in the Norwegian study [18]. In the third study, high appetite at two years mediated the association between a 16 variant PGS and BMI in French Canadian children aged 2 to 5 years, where the proportion mediated decreased from 47% at two years to 24% at five years [19]. These studies have the same limitations as the mediation analyses performed on adults; they were mostly cross-sectional and used PGS comprising a small number of variants. Cross-sectional designs are problematic in mediation analyses. Ideally, these mediation studies should be longitudinal to ensure the observed eating behaviours precede changes in body weight. More recent pan-genomic approaches to the construction of PGS include hundreds of thousands of variants yielding stronger BMI prediction [20].

In summary, previous studies suggest that eating behaviours could play a role in translating genetic susceptibility to higher BMI, but the evidence should be improved by examining this mediation 1) in childhood and adolescence for which the evidence is scarce and mixed, 2) using more recent and state-

of-the-art PGS that include a large number of variants to improve prediction, and 3) through longitudinal designs to assess the directionality of associations. Moreover, the bulk of evidence has been based on PGSs derived from adult anthropometric data, rather than data from children. Considering the genetic susceptibility to body weight variation changes throughout life [8], the identity and effect of specific genetic variants may also vary over the life course. Using PGSs derived from both adults and children is likely to yield more insights than using a single score. Therefore, the goal of the present study was to examine (1) the prediction of school-aged child BMI by two BMI-PGSs, respectively derived from child and adult GWAS summary statistics, and (2) the extent to which this prediction is mediated through eating behavioural traits (over-eating, fussy eating) assessed in early childhood (i.e., preschool age).

Methods

Study design and participants

We used data from the Quebec Longitudinal Study of Child Development (QLSCD), a longitudinal birth-cohort study from Quebec, Canada. The study recruited 2120 children born in 1997-1998 by a region-based stratified random selection from the Master birth register to ensure the sample was representative of the province's birth population. The children and their families were assessed annually or biannually starting at 5 months, and follow-up is ongoing. A variety of information was gathered on the children (e.g., behaviour, mental/physical health, diet, height, and weight assessments) and their families using questionnaires and interviews with the parents, teachers, and children themselves, as well as through direct assessments. More comprehensive details about the QLSCD, including exclusion criteria, attrition at each data collection, main findings, or a more exhaustive look at the data gathered, are available in a published cohort profile [21]. Biological samples were collected from a subset of QLSCD participants when the children were 10 years of age, and DNA was extracted. These DNA samples were later genotyped using the Illumina Infinium PsychArray-24 (see below). The present study included extensive BMI and eating behaviour assessments which were collected during follow-up. Informed consent was obtained from the parents at each step of the data collection until age 10, along with the children's permission. Starting at age 10, informed consent was obtained directly from the participants.

Anthropometric measures

Height and weight data were longitudinally collected through different means depending on the wave of data collection. Anthropometric data was reported by the parent, directly measured, or reported by the participant. In this study, we focus on the data measured directly. Specifically, trained research assistants measured height and weight on six occasions when participants were aged 6, 7, 8, 10, 12, and 13 years, following a standardized protocol, with the use of a measuring tape, ruler and scale. Two to three assessments were performed, and an average of the two closest values was computed as the final measure. BMI z-scores were then calculated using the World Health Organization Growth reference data [22], which considers both the sex and age of children using the R package "childsds" [23]. Considering the low prevalence of obesity in our sample, we decided against separating our sample in weight categories and proceeded with the BMI z-score variable as the outcome of interest.

Eating behaviours

Two eating behaviour scores (fussy eating and over-eating) were considered as the putative mediators. The scores were based on five eating behavioural traits assessed five times from age 2 to 6 years. The individual items were translated to French from those used in the Avon Longitudinal Study of Parents and Children [24]. Questions were modified when necessary based on recommendations from an expert advisory group and a pretest conducted in a sample of parents that were not part of the QLSCD [25]. For more details, see a previous publication using eating behaviour data from the QLSCD [26], and the QLSCD website (www.jesuisjeserai.stat.gouv.qc.ca). The most knowledgeable person about the child (the mother in most cases) answered items including (1*) When [name of the child] is at home with you for the main meal of the day, how often does [name of the child] (1) eat different meals… (*2) *…is [name of the child] fussy about food?* (3) *…refuse to eat?* (4) *… over-eat?* (5) *…eat too fast?* Answers were graded *on* a 4-point *Likert scale from*: "Almost never (1) to "Always (4)", for the first question, and "Never (1 point), Rarely (2 points), Sometimes (3 points), Often (4 points)" for the last four questions. The sum of "eating different meals", "fussy about food", and "refusing to eat" was calculated at each time point for the "fussy eating" score, and the sum of "the items "over-eating" and "eating too fast" was used to compute the "over-eating" score. The final two scores were obtained by averaging the sum at each time point with complete data per behaviour available per participant to obtain a mean score between 2 and 6 years. The internal consistency of the eating behaviour scores was satisfactory, with Cronbach's alpha of 0.82 for fussy eating and 0.74 for overeating.

Genotyping and polygenic score construction

DNA collection was proposed to a subsample of 1334 QLSCD participants remaining in the study at the 10-year-old data collection timepoint. A total of 992 participants agreed to DNA sampling and genotyping, which was performed using the Illumina Infinium PsychArray-24. Quality control of genetic data was performed as described in an earlier publication [27]. Variants with minor allele frequency below 0.01 and genotyping rate below 0.98 were excluded. Individuals with a call rate below 0.95 and those flagged for sex mismatch were excluded. Additional checks were also implemented for genetic duplicates, cryptic relatives, genetic outliers, and heterozygosity deviations. Variants in the remaining individuals were then re-evaluated to exclude non-autosomal variants, those with call rates below 98%, those with a minor allele frequency below 5% and Hardy-Weinberg Equilibrium test p-values below 1×10^{-3}. Imputation was performed using the 1000 genomes project reference data [28] and programs SHAPEIT v2 (r837) [29] and impute2 [30]. Finally, additional quality control was performed post-imputation where variants with a minor allele frequency below 1%, a Hardy-Weinberg Equilibrium test p-values below 1×10^{-6}, and an INFO metric below 0.8 were removed. After completing all quality control and imputation steps, 816 participants had genetic data, including 8,465,216 variants.

The PGSs were derived from the GWAS summary statistics obtained from prior studies performed on ~700,000 adults [10] and ~60,000 children of European ancestry [5] using the PRS-CS method [20] and the 1000 Genomes reference panel [28] as the external linkage disequilibrium reference panel. The global shrinkage parameter was set to 0.01 since BMI is a polygenic trait. Both PGSs were then regressed on the first 10 principal components of genetic ancestry to account for population stratification, with the subsequent residuals used to obtain adjusted scores, following the method described by Khera et al. [31] (R script in Supplementary Text 5-1). The final PGSs included 689,789 (child-based) and 613,732 (adult-based) individual variants. A total of 715 participants had genetic, eating behaviour, and anthropometric data available and passed genotypic quality control processes. An additional 23 participants did not have all covariate data available, leaving 692 participants as the sample size for the main mediation analysis. A description of how the participants were parsed from the initial QLSCD sample to the final study sample is available in Supplementary Figure 5-1.

Statistical analysis

The sociodemographic characteristics and details about birth (including birth weight, presence of preterm birth, maternal age at birth, maternal education, household income and whether the mother was born in Canada) of study participants are presented as proportions for categorical variables, mean (standard deviation) for symmetric continuous variables and median (IQR) for skewed variables. Study participants' characteristics were compared to those of QLSCD non-participants using a Chi-square test, one-way ANOVA and Wilcoxon rank sum test, depending on the variable type. Linear mixed models were used to estimate the association between PGSs and BMI z-scores to account for the repeated anthropometric measurements. Models included a smooth term for age, a random slope for age at measurement, and an autoregressive order 1 covariance structure. We also assessed the association between PGSs and the two eating behaviours using linear regression. Statistical significance was determined using the Student's t-test.

Longitudinal growth curve mediation analysis (LGCMA) was used to obtain the proportion of the variance of BMI z-score accounted for by both PGSs that is mediated by fussy eating and over-eating. The general model (Figure 5-1) included two latent variables (the intercept and the slope of the longitudinal outcome), a mediator regression model (equation A), an intercept regression model (equation B), and a slope regression (equation C):

$$EB = \beta_0 + \beta_1 PGSc_i + \beta_2 PGSa_i + \varepsilon_i \quad \text{(A)}$$

$$Intercept = \varphi_0 + \varphi_1 PGSc_i + \varphi_2 PGSa_i + \varphi_3 EB_i + \varepsilon_i \quad \text{(B)}$$

$$Slope = \gamma_0 + \gamma_1 PGSc_i + \gamma_2 PGSa_i + \gamma_3 EB_i + \varepsilon_i \quad \text{(C)}$$

where EB denotes eating behaviour, PGSc denotes the child-based PGS, and PGSa denotes the adult-based PGS. The proportion mediated (e.g. for the child-based PGS) can be obtained by solving the following (equation D) at each time-point:

$$\%\ Mediated = Mediated\ effect / Total\ effect \quad \text{(D)}$$

With: $$Mediated\ effect = \beta_1 PGS_i \times (\varphi_3 EB_i + \gamma_3 EB_i \times t)$$

And: $$Total\ effect = Mediated\ effect + \varphi_1 PGS_i + \gamma_1 PGS_i \times t$$

where *t* denotes age in years.

The methodology and equations A through D were adapted and applied to our study from previously described work [32]. We adjusted for potential baseline characteristics that could confound the association between eating behaviours and BMI z-scores. We considered variables available in the QLSCD which may affect eating behaviours and BMI based on available literature [33-38]. Those variables include birth weight, maternal BMI, preterm birth (yes/no), being the mother's first child (yes/no), maternal age at birth (20 years and under, 21 to 34 years, 35 years and over), maternal education (did not complete secondary school, completed secondary school, completed post-secondary diploma, university diploma), household income (less than $30 000, 30 000$ to 60 000$, 60 000$ to 80 000$, more than 80 000$), and immigration situation (mother born in Canada or not) and were added to the mediator and intercept regression models. We decided against including energy intake as a covariate in the analyses considering it more likely plays the role of intermediate variable between eating behaviours and BMI, rather than a confounder. There is evidence that eating behaviours are associated with dietary intakes (including energy intake) [39-41] in children, which is in turn associated with variations in BMI [39, 42]. Adjusting for energy intake would result in mitigating one of the mechanisms by which eating behaviours can influence BMI. We also assessed whether the mediation differed between girls and boys as supplementary analyses. We compared a constrained model where all the coefficients were restricted to be equal in boys and girls alike to an unconstrained model where the coefficients were allowed to vary between sexes using ANOVA.

In sensitivity analysis, inverse probability weighting (IPW) was used to assess the robustness of the primary analysis to informative loss to follow-up. The R package "twang" [43] was used to calculate a propensity score per participant based on a model where presence in the current study was predicted by baseline covariates likely to affect attrition (sex, maternal education, maternal BMI, birth weight, preterm birth, maternal age at birth, household income and whether the mother was born in Canada). The average treatment effect was used so that weighted study participants would resemble more

accurately the initial QLSCD participants. All statistical analyses were completed using R version 4.3.0 [44].

Results

More than half of the participants were female (54.8%). The majority (86.3%) had mothers aged 20 - 34 years at time of birth, a household income above $30,000 (74.7%), and almost all had a mother born in Canada (97.7%) (Table 5-1). When compared to QLSCD participants excluded from the study, the study sample included more females (54.8% vs. 46.3%), had mothers with higher levels of education (60.4% vs. 53.2% with post-secondary education), and lived in households with higher income (33.7% vs. 27.7% ≥60,000$ income) (Table 5-1).

The child- and adult-based PGSs were substantially associated with BMI z-scores (Table 5-2) after accounting for correlations between measures taken in the same children over the range of 6 to 13 years of age. An increase of one standard deviation of the child and adult-based PGS was associated with an increase in BMI of 0.26 standard deviations (95% CI: 0.19-0.33, p<0.001) and 0.34 standard deviations (95% CI: 0.27-0.41, p<0.001), respectively. Both PGSs were inversely associated with fussy eating (child: β=-0.14, 95% CI: -0.24 - -0.04, p=0.007; adult: β=-0.11, 95% CI: -0.21 - -0.01, p=0.030), and positively associated with over-eating (child: β=0.13, 95% CI: 0.07-0.20, p<0.001; adult: β=0.15, 95% CI: 0.08-0.22, p<0.001). Both PGSs were significantly associated with BMI z-scores in boys and girls (Table 5-2). PGSs were always significantly associated with eating behaviours in boys, but not in girls, although the confidence intervals overlapped (Table 5-2).

Table 5-3 shows the proportion of the predictive association between both PGSs and BMI z-scores mediated by fussy eating and over-eating from 6 to 13 years of age. The mediation by over-eating was statistically significant at all but one time point and was higher for the child-based PGS at all six time points compared to the adult-based PGS. Additionally, the mediation decreased over time for both PGSs. Thus, the strongest mediation was observed at age 6 (child-based PGS: 18.0%, 95% CI: 3.1-32.9, p-value=0.018; adult-based PGS: 14.2%, 95% CI: 2.8-25.5, p-value=0.014) and the lowest at age 13 (child-based PGS: 11.4%, 95% CI: -0.4-23.1, p-value = 0.057; adult-based PGS: 6.2%, 95% CI: 0.4-12.0, p-value = 0.037) adjusting for covariates. Fussy eating did not mediate the association between both PGSs and BMI z-scores.

Figures 5-2 and 5-3 describes in more detail the paths from the two LGCMA models, one for each behaviour. Each growth model estimated the intercept (IS), analog to stability, and the slope (SL) of BMI from ages 6 to 13. Each model also estimated the putative direct and indirect contributions of each PGS, the latter through over-eating and fussy eating, respectively. The two models revealed unique direct and indirect (i.e., mediated) contributions of both the child-based PGS and adult-based PGS to BMI stability and slope. In the case of fussy eating (Figure 5-2), direct associations were the rule and mediation pathways were not conclusive. For instance, both the child-based PGS and the adult-based PGS were directly associated with the BMI intercept, but not with the BMI slope. The absence of mediation by fussy eating is indicated by the association between fussy eating and the BMI intercept being close to the null. Figure 5-3 illustrates that both the child-based PGS and the adult-based PGS were associated with over-eating, which was in turn associated with the BMI intercept. Over-eating was inversely associated with the BMI slope, which explains the decreasing mediation observed for both PGSs through time. The association between the adult-based PGS and BMI slope (β: 0.012, i.e., increase over time in the direct effect on BMI) was higher compared to the child-based PGS (β: 0.000), which illustrates the higher decrease in mediation observed through time for the adult-based PGS compared to the child-based PGS. In additional analysis, we assessed whether the proportion mediated by over-eating differed between boys and girls. The models did not differ by sex (Δ Chi-square: 10.50, Δ DF: 8, p-value: 0.232).

To assess the robustness of our findings, we performed IPW to ensure the study sample resembled the initial sample in terms of baseline characteristics. As seen in Supplementary table 5-2, IPW did not change the overall results, with the estimated proportion of effect mediated by over-eating decreasing with time for both PGSs. In terms of magnitude, the proportion mediated in the IPW model was higher at all but two time point compared to the original analyses for both PGSs. The main difference observed was that the confidence intervals related to the adult PGS using IPW were slightly larger compared to the initial analyses, leading to non-significant results. After adjusting for attrition, the proportion mediated by fussy eating was still close to null for both PGSs.

Discussion

The goal of the present study was to assess (1) the concurrent associations between two PGSs derived from child and adult GWAS and childhood BMI and (2) if problematic eating behaviours, such as fussy eating and over-eating, mediated these associations. Using LGCMA we found that the child-derived and the adult-derived PGSs were independently associated with BMI z-scores in childhood. This suggests that albeit being correlated, both PGS capture unique genetic contributions to obesity. Furthermore, we confirmed that over-eating mediates the association between both PGSs and BMI z-scores. This result is consistent with recent research suggesting that weight-related genes act by changing the expression of proteins in brain regions involved in appetite and satiety regulation [45].

Our study adds to growing evidence of the emerging role of common genetic variants influencing BMI In childhood in two ways: (1) the presence of independent effects on BMI from PGSs derived from child and adult GWAS, and (2) the increase over time in the direct effect on BMI for the adult-based PGS, but not the child-based PGS. This is consistent with evidence from large-scale GWAS analysis and twin studies. Increasing differences in weight between high- and low-genetic-risk individuals throughout childhood in 300,000 individuals suggest that new genetic sources of variation seem to emerge in childhood [46]. This is also supported by the observation that a PGS developed on an adult GWAS presented an increasing contribution to the variation in BMI throughout childhood, especially between 3 and 8 years [47]. Furthermore, results from a large-scale analysis of over 38,000 twin pairs support that genetic factors affecting variations in BMI change throughout childhood. The study [48] showed that genetic correlations between BMI measures were smaller with increasing age from 1 to 19 years. More specifically, our study added evidence for unique contributions to BMI variation from genetic factors present in early childhood and those emerging later in development. This suggests that PGSs derived from GWAS statistics obtained in child and adult populations contain overlapping, but distinct genetic factors.

The role of appetite as a mediator of the genetic contribution to body weight in childhood is a central tenet of the Behavioural susceptibility theory. Changes in the food environment contribute to the risk of obesity, with genetic susceptibility responsible for shaping the individual's response to such changes. Our study supports this hypothesis. Our over-eating score is composed of two items referring to eating too much and eating too fast. The well-known Child Eating Behavior Questionnaire subscales food responsiveness and satiety responsiveness/slowness in eating includes similar items as "If allowed to,

my child would eat too much", and "My child finishes his/her meal very quickly", respectively [7]. It is likely that our over-eating score relates more to the appetite and satiety compared to emotional behaviours that can also be studied. This also aligns with another study completed using QLSCD data that found an association between the over-eating score and satiety responsiveness measured at 22 years using the adult version of the Child Eating Behavior Questionnaire [49].

Prior studies have used a similar approach to test the mediation of the contribution of genetic factors to BMI by appetite-related traits in children. One study [17] found mediation by satiety responsiveness, while another [18] did not. Differences between prior results could be explained through a contrast in methods. Mainly, cross-sectional data was used in the study that observed mediation at a single point in time, while the other used a longitudinal method designed to identify mediation for BMI growth (slope), rather than BMI at a specific time point. With its longitudinal design focused on identifying mediation at different points in time, our study allowed us to shed light on previous mixed results. We identified mediation through over-eating at sequential time points that was largely driven by the intercept of BMI. The use of structural models allowed us to observe the strong association between over-eating and the intercept of BMI responsible for the high mediation at 6 years, which then decreased due to an inverse effect of over-eating on the slope of BMI.

This result is consistent with another study of children that found decreasing mediation through time. High appetite at two years mediated the effect of a 16 variant PGS on BMI z-score from 2 to 5 years in over 1,000 French Canadian children [19]. The proportions mediated by high appetite decreased from 47% at the 2-year mark to 24% at the 5-year mark. Like our study, the reduction in mediation at older time points likely stems from the passage of time, where appetite at 2 years becomes progressively less influential. Indeed, the inverse effect of over-eating on BMI growth could be explained by the BMI observations becoming further apart from the index measurement of over-eating as time passes. The concordance of our results and those obtained in Lauzon-Guillain et al. [19] is compelling since one study gathered anthropometric data before adiposity rebound (typically around 5 years) and the other after. This suggests that mediation through over-eating remains present at different developmental stages.

The presence of mediation of genetic susceptibility to obesity through over-eating aligns with molecular genetic literature that emphasizes the importance of the brain in translating the effect of genetic factors on body weight regulation. RNA-based analyses applied to large-scale GWAS have shown that most variants associated with BMI are predominantly expressed in the central nervous

system in regions such as the hippocampus and limbic system, involved in cognition and emotion regulation, but also the hypothalamus and pituitary gland, responsible for appetite regulation [9, 10]. Additionally, a recent study identified 60 instances where protein concentration in the brain was linked with genetic variants known to influence obesity [45]. More specifically, that study used a combination of gene colocalization and mendelian randomization analyses focused on the left dorsolateral prefrontal cortex. This brain region notably influences appetite, satiety regulation, and cognitive functions, including decision-making and executive functioning [45]. These results provide a biological basis for the theory that genetic susceptibility to obesity is expressed through variations in appetite, resulting from changes in protein concentration. We did not observe mediation through fussy eating. This result complements the available but scarce literature on the subject. Avoidant eating and slowness in eating were not found to mediate the effect of a PGS on BMI in a Finnish adult cohort and a Norwegian longitudinal children study, respectively [15, 18].

The additional evidence favouring the Behavioural susceptibility theory further provides incentives for supporting obesity prevention efforts that focus on creating a healthy food environment rather than centred on personal responsibility. This is an important public health issue since a shift from the global food system leading to deteriorating nutritional choices is considered a driving factor behind the rise of obesity [50]. For example, the current food environment in Canada does not favour healthy food consumption patterns on most fronts [51]. A recent evaluation of the policies and actions of multiple levels of government in Canada for creating a healthy food environment highlighted 1) many limitations to Canada's food environment policy landscape, and 2) restrictions on marketing aimed at children as a key policy area to improve a healthier food landscape in Canada [51].

The primary strength of the present study is the use of a longitudinal design. This allows a better characterization of the mediation of genetic susceptibility to obesity by over-eating throughout childhood and adolescence. Many studies only use a singular time point to describe this process, while we were able to detail how the mediation evolves through time. Furthermore, the use of PRS-CS to calculate the PGSs is another strength of the study considering the method captures more genetic variability compared with methods only incorporating a few variants into the scores. The study also has a few limitations. First, even though we adjusted our analysis models for potential confounders, there could still be unmeasured confounders, such as physical activity, blurring the association between the eating behaviours and BMI z-scores specifically. Second, we used inverse probability weighting to attenuate potential selection bias through attrition. However, we could only produce a weighted study

sample similar to the initial QLSCD sample in terms of available baseline characteristics, therefore the study sample could still differ from the initial sample in some capacity. Third, BMI is known to have limits in distinguishing between lean and fat mass and does not distinguish between the location of fat in the body. Fourth, eating behaviour data was not available throughout the childhood period with measured anthropometric collection. Eating behaviours might change as children's age increases, and the strength of the association between eating behaviours exhibited in early childhood and BMI would be expected to diminish as time passes.

Conclusion

In our investigation of French-Canadian children, over-eating throughout childhood, but not fussy eating, mediated genetic susceptibility to obesity. This result aligns with the prevailing theoretical construct behind the genetic architecture of obesity that hinges on the role of appetite. The few previous studies identified appetite-related traits as mediators of genetic susceptibility to weight gain but were mostly performed in adult populations and had a cross-sectional design. The present study analysed data from a longitudinal study of Canadian children and used dense genotypic scores as measures of genetic susceptibility to obesity. The finding of mediation by over-eating suggests that global and wide-reaching policies aimed at introducing a healthier food environment will be key in implementing obesity prevention efforts.

Tables and Figures

Table 5-1. Characteristics of study participants.

Characteristic *% (n) or mean ± SD*	*Included* *N=692*	*Excluded* *N=1428*	*p-value*
Sex, female % (n)	54.8 (379)	46.3 (661)	**<0.001**
Birth weight, kg mean ± SD	3.42 ± 0.49	3.40 ± 0.50	0.351
Preterm birth % (n)	5.1 (35)	4.5 (64)	0.631
First child % (n)	41.8 (289)	45.3 (647)	0.135
Maternal age at birth % (n)			0.098
≤ 20 years	1.7 (12)	3.3 (47)	
20-34 years	84.0 (581)	83.7 (1194)	
≥ 35 years	14.3 (99)	13.0 (186)	
Maternal education % (n)			**0.009**
< Secondary school diploma	15.6 (110)	19.3 (275)	
Secondary school diploma	24.3 (164)	27.4 (391)	
Post-sec. except university	29.5 (206)	28.4 (405)	
University diploma	30.5 (212)	24.8 (354)	
Household income % (n)			**0.003**
< 30 000$	25.3 (175)	32.2 (447)	
30 000 – <60 000$	41.0 (284)	40.2 (559)	
60 000 – <80 000$	17.2 (119)	15.0 (208)	
≥ 80 000$	16.5 (114)	12.7 (176)	
Born in Canada % (n)	97.7 (676)	83.4 (1191)	**<0.001**
Fussy eating med (IQR)	5.60 (4.80, 6.60)	5.67 (4.75, 6.67)	0.745
Over-eating med (IQR)	2.80 (2.33, 3.60)	2.75 (2.20, 3.50)	0.207
Obesity at 6 years % (n)			0.685
Healthy weight	80.2 (441)	80.0 (496)	
Overweight	14.0 (77)	13.1 (81)	
Obesity	5.8 (32)	6.9 (43)	
Obesity at 13 years % (n)			0.202
Healthy weight	63.9 (389)	67.1 (416)	

Overweight	22.0 (134)	22.1 (137)	
Obesity	14.1 (86)	10.8 (67)	

Study participants included those (N = 692) who had complete anthropometric, genetic, eating behaviour, and covariate data available, and were thus part of the main mediation analysis. Chi-square test used for categorical variables, ANOVA test used for normally distributed continuous variables, and Wilcoxon ran sum test used for non-normally distributed continuous variables. Significance threshold (**bold**) set at 0.05.

Table 5-2. Association of the PGS with eating behaviors and BMI in children.

Outcome	**Child PGS**		**Adult PGS**	
	β (95% CI)	p-value	β (95% CI)	p-value
BMI z-score				
Total	0.26 (0.19, 0.33)	**<0.001**	0.34 (0.27, 0.41)	**<0.001**
Girls	0.22 (0.12, 0.32)	**<0.001**	0.25 (0.16, 0.34)	**<0.001**
Boys	0.31 (0.20, 0.42)	**<0.001**	0.45 (0.34, 0.56)	**<0.001**
Fussy eating				
Total	-0.14 (-0.24, -0.04)	**0.007**	-0.11 (-0.21, -0.01)	**0.030**
Girls	-0.07 (-0.20, 0.07)	0.340	-0.02 (-0.16, 0.11)	0.750
Boys	-0.22 (-0.37, -0.07)	**0.003**	-0.23 (-0.38, -0.07)	**0.004**
Overeating				
Total	0.13 (0.07, 0.20)	**<0.001**	0.15 (0.08, 0.22)	**<0.001**
Girls	0.07 (-0.02, 0.17)	0.113	0.10 (0.01, 0.19)	**0.033**
Boys	0.20 (0.10, 0.31)	**<0.001**	0.22 (0.12, 0.32)	**<0.001**

Estimated effect (β) of increase in one standard deviation of the child- and adult-based PGS and BMI z-score (linear mixed model), fussy eating (linear regression), and over-eating (linear regression) with 95% confidence interval. BMI z-score model includes age, age-squared, and sex as covariates with random intercept and slope for age. Eating behavior models include sex as a covariate. Significance threshold (**bold**) set at 0.05.

Table 5-3. Proportion of the effect of PGS on BMI z-score mediated by eating behaviours.

Model	**Child PGS**		**Adult PGS**	
	% (95% CI)	**p-value**	**% (95% CI)**	**p-value**
Fussy eating	N=692			
6 years	1.9 (-2.1, 5.8)	0.360	1.3 (-1.6, 4.3)	0.369
7 years	1.8 (-2.1, 5.6)	0.362	1.2 (-1.4, 3.9)	0.370
8 years	1.7 (-2.1, 5.5)	0.372	1.1 (-1.4, 3.6)	0.379
10 years	1.6 (-2.2, 5.3)	0.421	0.9 (-1.3, 3.1)	0.425
12 years	1.4 (-2.7, 5.5)	0.504	0.7 (-1.4, 2.9)	0.505
13 years	1.3 (-3.0, 5.7)	0.554	0.7 (-1.5, 2.8)	0.554
Over-eating	N=692			
6 years	18.0 (3.1, 32.9)	**0.018**	14.2 (2.8, 25.5)	**0.014**
7 years	17.1 (3.0, 31.3)	**0.018**	12.7 (2.6, 22.9)	**0.014**
8 years	16.2 (2.7, 29.8)	**0.019**	11.4 (2.3, 20.5)	**0.015**
10 years	14.4 (1.8, 26.9)	**0.025**	9.1 (1.5, 16.6)	**0.018**
12 years	12.4 (0.4, 24.3)	**0.042**	7.1 (0.8, 13.3)	**0.028**
13 years	11.4 (-0.4, 23.1)	0.057	6.2 (0.4, 12.0)	**0.037**

Estimated proportion mediated by eating behaviours (fussy eating, over-eating) in the association between PGS and BMI z-score with 95% CI using latent growth curve mediation analysis. Models are all adjusted for birth weight, preterm birth, maternal BMI, maternal age, maternal education, household income, whether the mother was born in Canada, and whether the participant is the first child as covariates in both the eating behavior and BMI z-score (effect on the initial state) outcome models. Significance threshold (**bold**) set at 0.05.

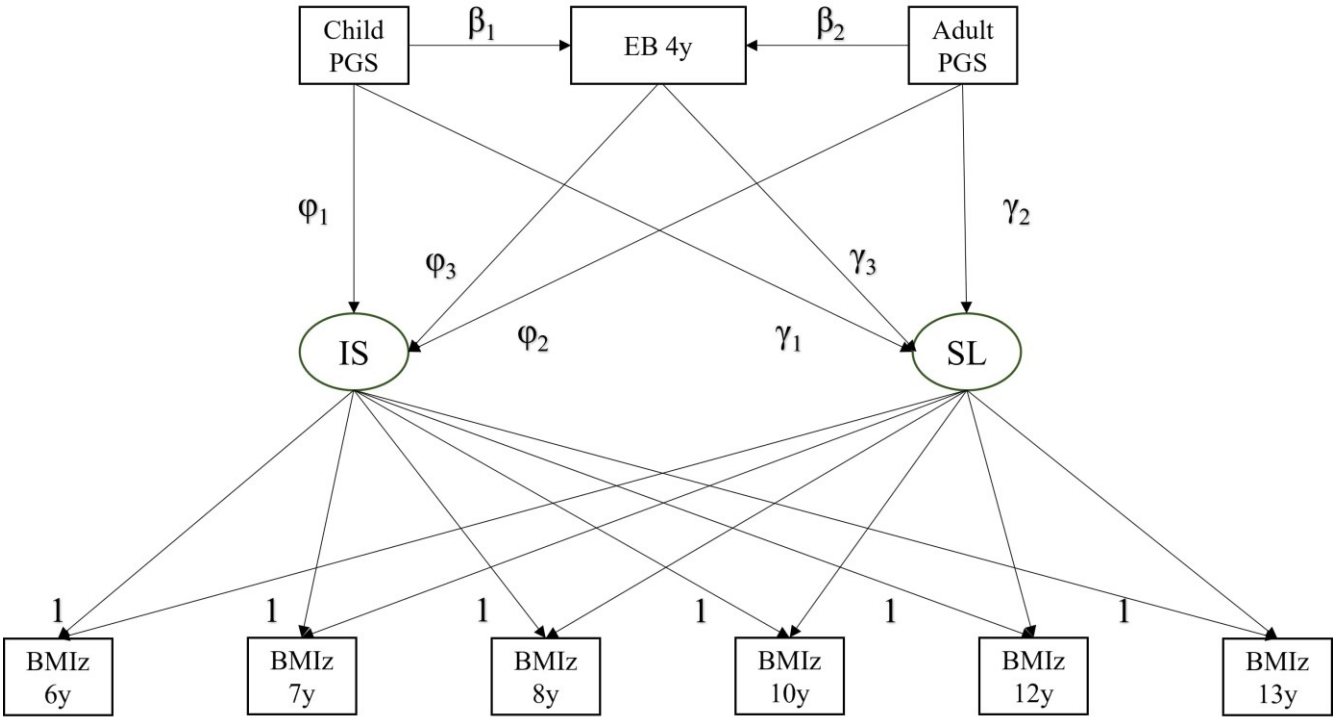


Figure 5-1. Latent growth curve mediation model. **EB = eating behaviour, IS = initial state, SL = slope. Figure depicts three regression models with: B = coefficients for the eating behaviour ourcome model, φ = coefficients for the IS outcome model, and ϒ = coefficients for the SL outcome model.**

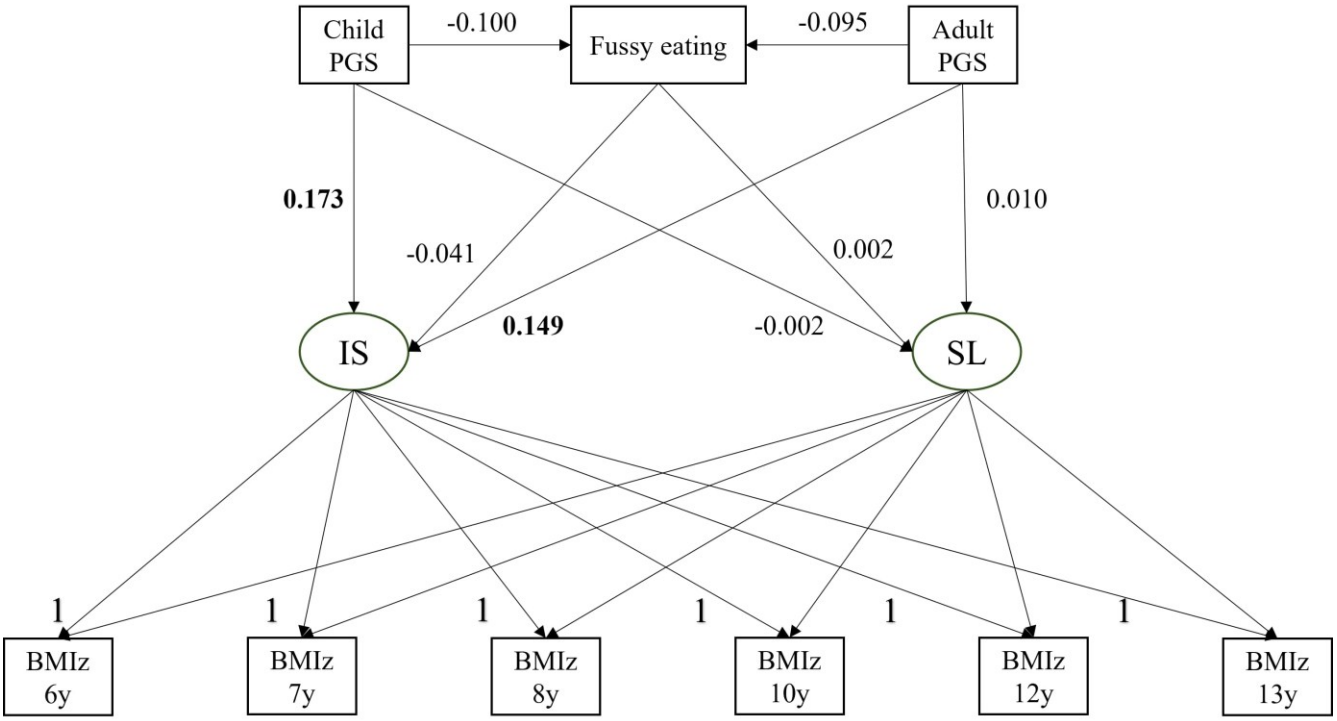


Figure 5-2. Latent growth curve mediation model, for fussy eating. **IS = initial state, SL = slope. The figure depicts four direct associations between PGSs and BMI z-scores: from the child PGS to the initial state of BMI z-scores (1) and the slope of BMI z-scores (2), and from the adult PGS to the initial state (3) and the slope (4). Four indirect associations between PGSs and BMI z-scores are presented: From the child PGS through fussy eating to the initial state (1) and the slope (2), and from the adult PGS through fussy eating to the initial state (3) and the slope (4). Paths in bold are significant at p<0.05.**

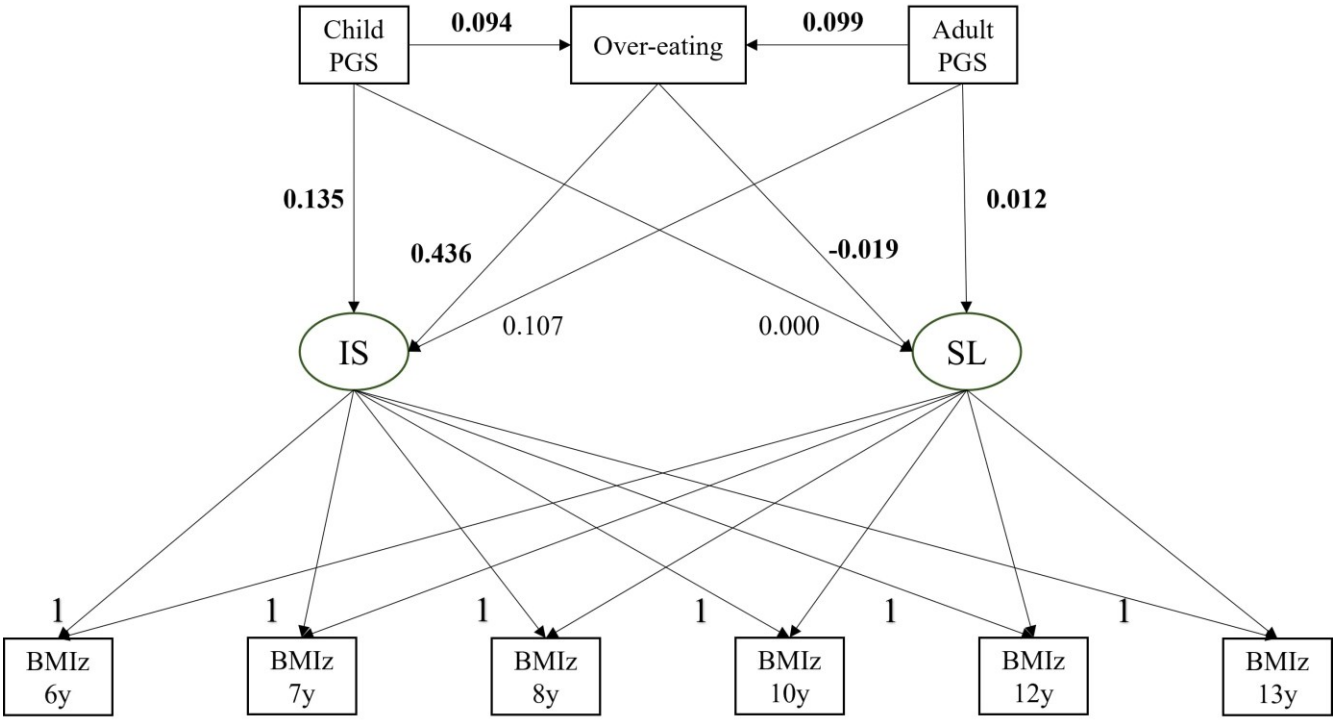


Figure 5-3. Latent growth curve mediation model, for over-eating. **IS = initial state, SL = slope. The figure depicts four direct associations between PGSs and BMI z-scores: from the child PGS to the initial state of BMI z-scores (1) and the slope of BMI z-scores (2), and from the adult PGS to the initial state (3) and the slope (4). Four indirect associations between PGSs and BMI z-scores are presented: From the child PGS through over-eating to the initial state (1) and the slope (2), and from the adult PGS through over-eating to the initial state (3) and the slope (4). Paths in bold are significant at p<0.05.**

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

ACKNOWLEDGEMENTS

DG, LD and MB contributed to the conception and design of the study. LD and MB were involved in the data collection from QLSCD participants. DG performed the statistical analyses with insights and revisions from CG. DG was responsible for the first draft of the manuscript. All authors participated in discussions concerning the development of the study and were involved in the revisions to the manuscript and approved the submitted version.

We also acknowledge the contribution of Till Andlauer, Stéphane Paquin, Geneviève Morneau-Vaillancourt, Isabelle Ouellet-Morin and Michel Boivin, who were involved in the quality control of the genetic data of the QLSCD participants that are used in the research. We are grateful to the QLSCD participants and their families who took part in the various data collection rounds over the years.

FUNDING

This work was supported by a CIHR operating grant (#165964). The funders were not involved in the design or conduct of the study; collection, management, analysis, or interpretation of the data; or preparation, review, or approval of the manuscript. The analyses were performed using data from the Quebec Longitudinal Study of Child Development (QLSCD), conducted by Sante Quebec, a division of the Institut de la Statistique du Quebec (ISQ) and funded by the Ministry of Health and Social Services of Quebec.

DATA AVAILABILITY STATEMENT

The data used to support the findings of this study was obtained from the Québec Longitudinal Study of Child Development, conducted by Santé Québec, a division of the Institut de la Statistique du Québec and may be released upon application to the Institut de la Statistique du Québec, through the Zone de recherche at: https://statistique.quebec.ca/fr/institut/services-recherche#/accueil

References

1. Haththotuwa, R.N., C.N. Wijeyaratne, and U. Senarath, *Chapter 1 - Worldwide epidemic of obesity*, in *Obesity and Obstetrics (Second Edition)*, T.A. Mahmood, S. Arulkumaran, and F.A. Chervenak, Editors. 2020, Elsevier. p. 3-8.
2. Statistique Canada. *Indice de masse corporelle des enfants et des jeunes, 2012 à 2013*. 2014 2015 [cited 23 novembre 2020; Available from: https://www150.statcan.gc.ca/n1/pub/82-625-x/2014001/article/14105-fra.htm.
3. Simmonds, M., et al., *Predicting adult obesity from childhood obesity: a systematic review and meta-analysis.* Obesity Reviews, 2016. **17**(2): p. 95-107 DOI: https://doi.org/10.1111/obr.12334.
4. Silventoinen, K., et al., *Genetic and environmental effects on body mass index from infancy to the onset of adulthood: an individual-based pooled analysis of 45 twin cohorts participating in the COllaborative project of Development of Anthropometrical measures in Twins (CODATwins) study.* The American journal of clinical nutrition, 2016. **104**(2): p. 371-379 DOI: 10.3945/ajcn.116.130252.
5. Vogelezang, S., et al., *Novel loci for childhood body mass index and shared heritability with adult cardiometabolic traits.* PLoS genetics, 2020. **16**(10): p. e1008718-e1008718 DOI: 10.1371/journal.pgen.1008718.
6. Felix, J.F., et al., *Genome-wide association analysis identifies three new susceptibility loci for childhood body mass index.* Human molecular genetics, 2016. **25**(2): p. 389-403 DOI: 10.1093/hmg/ddv472.
7. Wardle, J., et al., *Development of the Children's Eating Behaviour Questionnaire.* Journal of Child Psychology and Psychiatry, 2001. **42**(7): p. 963-970 DOI: https://doi.org/10.1111/1469-7610.00792.
8. Silventoinen, K. and H. Konttinen, *Obesity and eating behavior from the perspective of twin and genetic research.* Neuroscience & Biobehavioral Reviews, 2020. **109**: p. 150-165 DOI: https://doi.org/10.1016/j.neubiorev.2019.12.012.
9. Locke, A.E., et al., *Genetic studies of body mass index yield new insights for obesity biology.* Nature, 2015. **518**(7538): p. 197-206 DOI: 10.1038/nature14177.

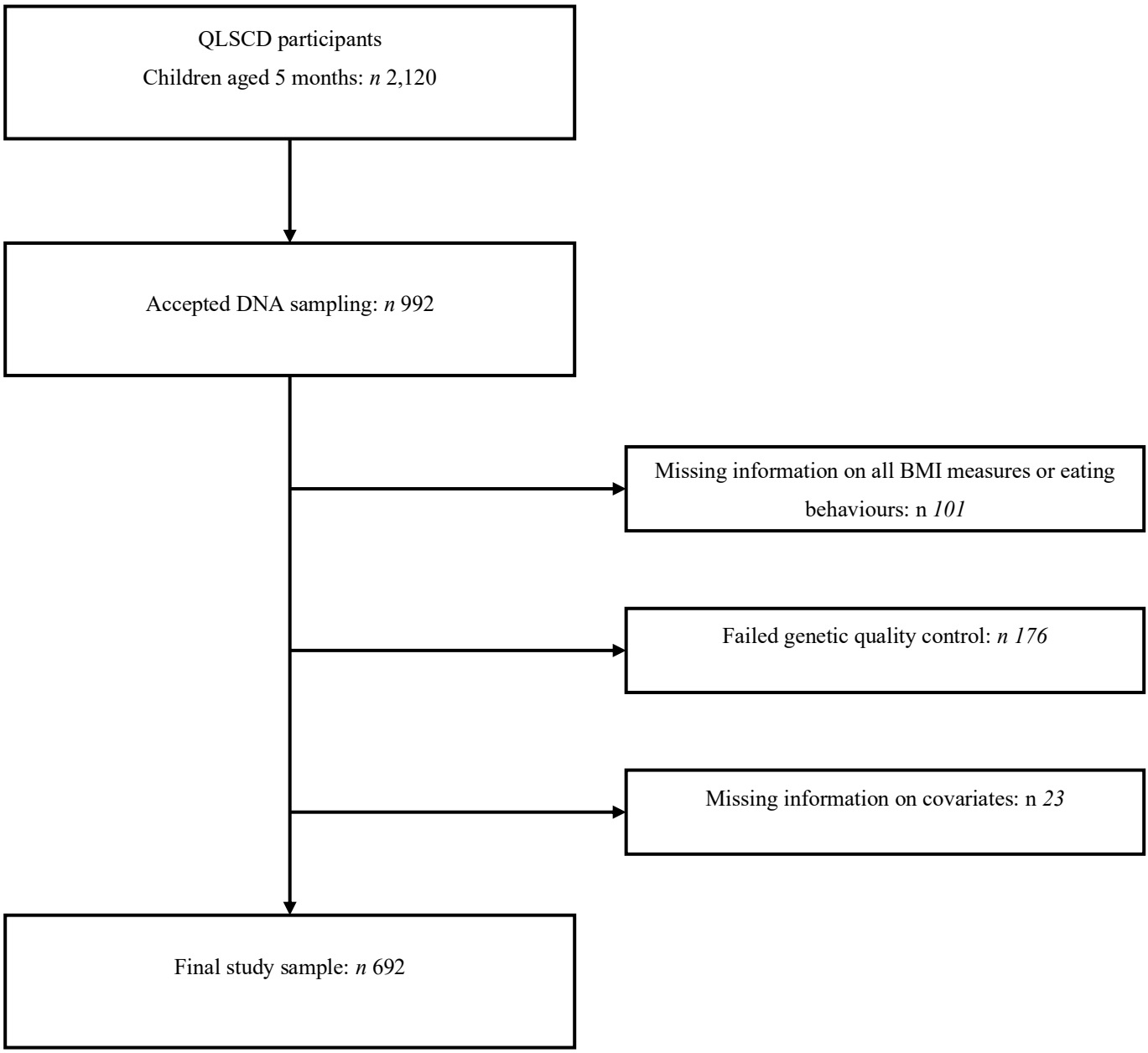
10. Yengo, L., et al., *Meta-analysis of genome-wide association studies for height and body mass index in ~700000 individuals of European ancestry.* Human molecular genetics, 2018. **27**(20): p. 3641-3649 DOI: 10.1093/hmg/ddy271.
11. Konttinen, H., et al., *Appetitive traits as behavioural pathways in genetic susceptibility to obesity: a population-based cross-sectional study.* Scientific Reports, 2015. **5**(1): p. 14726 DOI: 10.1038/srep14726.
12. de Lauzon-Guillain, B., et al., *Mediation and modification of genetic susceptibility to obesity by eating behaviors.* The American Journal of Clinical Nutrition, 2017. **106**(4): p. 996-1004 DOI: 10.3945/ajcn.117.157396 %J The American Journal of Clinical Nutrition.
13. Jacob, R., et al., *The role of eating behavior traits in mediating genetic susceptibility to obesity.* The American Journal of Clinical Nutrition, 2018. **108**(3): p. 445-452 DOI: 10.1093/ajcn/nqy130 %J The American Journal of Clinical Nutrition.
14. Brunner, E.J., et al., *Appetite disinhibition rather than hunger explains genetic effects on adult BMI trajectory.* International Journal of Obesity, 2021. **45**(4): p. 758-765 DOI: 10.1038/s41366-020-00735-9.
15. Masip, G., et al., *The genetic architecture of the association between eating behaviors and obesity: combining genetic twin modeling and polygenic risk scores.* The American Journal of Clinical Nutrition, 2020. **112**(4): p. 956-966 DOI: 10.1093/ajcn/nqaa181.
16. Herle, M., et al., *The Role of Eating Behaviours in Genetic Susceptibility to Obesity.* Current Obesity Reports, 2020. **9**(4): p. 512-521 DOI: 10.1007/s13679-020-00402-0.
17. Llewellyn, C.H., et al., *Satiety Mechanisms in Genetic Risk of Obesity.* JAMA Pediatrics, 2014. **168**(4): p. 338-344 DOI: 10.1001/jamapediatrics.2013.4944 %J JAMA Pediatrics.
18. Steinsbekk, S., et al., *Polygenic Risk, Appetite Traits, and Weight Gain in Middle Childhood: A Longitudinal Study.* JAMA Pediatrics, 2016. **170**(2): p. e154472-e154472 DOI: 10.1001/jamapediatrics.2015.4472 %J JAMA Pediatrics.
19. de Lauzon-Guillain, B., et al., *Association between genetic obesity susceptibility and mother-reported eating behaviour in children up to 5 years.* 2019. **14**(5): p. e12496 DOI: https://doi.org/10.1111/ijpo.12496.
20. Ge, T., et al., *Polygenic prediction via Bayesian regression and continuous shrinkage priors.* Nature communications, 2019. **10**(1): p. 1776-1776 DOI: 10.1038/s41467-019-09718-5.
21. Orri, M., et al., *Cohort Profile: Quebec Longitudinal Study of Child Development (QLSCD).* Social Psychiatry and Psychiatric Epidemiology, 2020 DOI: 10.1007/s00127-020-01972-z.

22. de Onis, M., et al., *Development of a WHO growth reference for school-aged children and adolescents.* Bulletin of the World Health Organization, 2007. **85**(9): p. 660-667 DOI: 10.2471/blt.07.043497.
23. Vogel, M., *childsds: Data and Methods Around Reference Values in Pediatrics*. 2022, https://CRAN.R-project.org/package=childsds.
24. Golding, et al., *ALSPAC–The Avon Longitudinal Study of Parents and Children.* Paediatric and Perinatal Epidemiology, 2001. **15**(1): p. 74-87 DOI: https://doi.org/10.1046/j.1365-3016.2001.00325.x.
25. Institut de la statistique du Québec, *Enquête sociale et de santé auprès des enfants et des adolescents québécois.* 1999, Les Publications du Québec. p. 520, https://statistique.quebec.ca/fr/document/enquete-sociale-et-de-sante-aupres-des-enfants-et-adolescents-quebecois-1999.
26. Dubois, L., et al., *Preschool children's eating behaviours are related to dietary adequacy and body weight.* European Journal of Clinical Nutrition, 2007. **61**(7): p. 846-855 DOI: 10.1038/sj.ejcn.1602586.
27. Perret, L.C., et al., *Polygenic risk score and peer victimisation independently predict depressive symptoms in adolescence: results from the Quebec Longitudinal Study of Children Development.* Journal of Child Psychology and Psychiatry, 2022. **n/a**(n/a) DOI: https://doi.org/10.1111/jcpp.13706.
28. Auton, A., et al., *A global reference for human genetic variation.* Nature, 2015. **526**(7571): p. 68-74 DOI: 10.1038/nature15393.
29. Delaneau, O., J.-F. Zagury, and J. Marchini, *Improved whole-chromosome phasing for disease and population genetic studies.* Nature Methods, 2013. **10**(1): p. 5-6 DOI: 10.1038/nmeth.2307.
30. Howie, B.N., P. Donnelly, and J. Marchini, *A Flexible and Accurate Genotype Imputation Method for the Next Generation of Genome-Wide Association Studies.* PLOS Genetics, 2009. **5**(6): p. e1000529 DOI: 10.1371/journal.pgen.1000529.
31. Khera, A.V., et al., *Whole-Genome Sequencing to Characterize Monogenic and Polygenic Contributions in Patients Hospitalized With Early-Onset Myocardial Infarction.* Circulation, 2019. **139**(13): p. 1593-1602 DOI: 10.1161/circulationaha.118.035658.
32. Sullivan, A., et al., *Longitudinal Mediation Analysis with Latent Growth Curves.* 2021 DOI: https://doi.org/10.48550/arXiv.2103.05765.

33. Keane, E., et al., *Measured Parental Weight Status and Familial Socio-Economic Status Correlates with Childhood Overweight and Obesity at Age 9.* PLOS ONE, 2012. **7**(8): p. e43503 DOI: 10.1371/journal.pone.0043503.
34. Barclay, K. and M. Myrskylä, *Maternal age and offspring health and health behaviours in late adolescence in Sweden.* SSM - Population Health, 2016. **2**: p. 68-76 DOI: https://doi.org/10.1016/j.ssmph.2016.02.012.
35. Williams, A.S., et al., *Socioeconomic Status and Other Factors Associated with Childhood Obesity.* J Am Board Fam Med, 2018. **31**(4): p. 514-521 DOI: 10.3122/jabfm.2018.04.170261.
36. Gnawali, A., *Prematurity and the Risk of Development of Childhood Obesity: Piecing Together the Pathophysiological Puzzle. A Literature Review.* Cureus, 2021. **13**(12): p. e20518 DOI: 10.7759/cureus.20518.
37. Ochiai, H., et al., *Number of siblings, birth order, and childhood overweight: a population-based cross-sectional study in Japan.* BMC Public Health, 2012. **12**: p. 766 DOI: 10.1186/1471-2458-12-766.
38. White, P.A., et al., *Household income and maternal education in early childhood and risk of overweight and obesity in late childhood: Findings from seven birth cohort studies in six high-income countries.* International Journal of Obesity, 2022. **46**(9): p. 1703-1711 DOI: 10.1038/s41366-022-01171-7.
39. French, S.A., et al., *Eating behavior dimensions. Associations with energy intake and body weight. A review.* Appetite, 2012. **59**(2): p. 541-549 DOI: https://doi.org/10.1016/j.appet.2012.07.001.
40. Fogel, A., et al., *Oral processing behaviours that promote children's energy intake are associated with parent-reported appetitive traits: Results from the GUSTO cohort.* Appetite, 2018. **126**: p. 8-15 DOI: 10.1016/j.appet.2018.03.011.
41. Maneschy, I., et al., *Eating Behaviours and Dietary Intake in Children and Adolescents: A Systematic Review.* Current Nutrition Reports, 2024. **13**(3): p. 363-376 DOI: 10.1007/s13668-024-00544-w.
42. Kininmonth, A., et al., *The association between childhood adiposity and appetite assessed using the Child Eating Behavior Questionnaire and Baby Eating Behavior Questionnaire: A systematic review and meta-analysis.* Obes Rev, 2021. **22**(5): p. e13169 DOI: 10.1111/obr.13169.

43. Cefalu, M., et al., *twang: Toolkit for Weighting and Analysis of Nonequivalent Groups*. 2021, https://CRAN.R-project.org/package=twang.
44. R Core Team, *R: A Language and Environment for Statistical Computing*. 2024, R Foundation for Statistical Computing: Vienna, Autria, https://www.R-project.org/.
45. Gagnon, E., et al., *Genetic control of body weight by the human brain proteome.* iScience, 2023. **26**(4) DOI: 10.1016/j.isci.2023.106376.
46. Khera, A.V., et al., *Polygenic Prediction of Weight and Obesity Trajectories from Birth to Adulthood.* Cell, 2019. **177**(3): p. 587-596.e9 DOI: 10.1016/j.cell.2019.03.028.
47. Helgeland, Ø., et al., *Characterization of the genetic architecture of infant and early childhood body mass index.* Nature Metabolism, 2022 DOI: 10.1038/s42255-022-00549-1.
48. Silventoinen, K., et al., *Changing genetic architecture of body mass index from infancy to early adulthood: an individual based pooled analysis of 25 twin cohorts.* International Journal of Obesity, 2022 DOI: 10.1038/s41366-022-01202-3.
49. Dubois, L., et al., *Eating behaviors, dietary patterns and weight status in emerging adulthood and longitudinal associations with eating behaviors in early childhood.* Int J Behav Nutr Phys Act, 2022. **19**(1): p. 139 DOI: 10.1186/s12966-022-01376-z.
50. Swinburn, B.A., et al., *The global obesity pandemic: shaped by global drivers and local environments.* The Lancet, 2011. **378**(9793): p. 804-814 DOI: 10.1016/S0140-6736(11)60813-1.
51. Vanderlee, L., et al., *Policies to Create Healthier Food Environments in Canada: Experts' Evaluation and Prioritized Actions Using the Healthy Food Environment Policy Index (Food-EPI).* Int J Environ Res Public Health, 2019. **16**(22) DOI: 10.3390/ijerph16224473.

Supplementary materials



Supplementary Figure 5-1. Flow chart of study participants from the Québec Longitudinal Study of Child Development (QLSCD) participants.

Supplementary Table 5-1. Proportion of the effect of PGS on BMI z-score mediated by over-eating, IPW for attrition.

Model	Child PGS		Adult PGS	
	% (95% CI)	**p-value**	**% (95% CI)**	**p-value**
Fussy eating	N=692			
6 years	0.9 (-2.3, 4.2)	0.573	1.4 (-2.1, 5.0)	0.421
7 years	0.8 (-2.2, 3.9)	0.587	1.2 (-1.9, 4.2)	0.448
8 years	0.7 (-2.1, 3.5)	0.610	0.9 (-1.7, 3.6)	0.492
10 years	0.5 (-2.0, 3.0)	0.699	0.6 (-1.8, 2.9)	0.642
12 years	0.2 (-2.3, 2.8)	0.852	0.2 (-2.1, 2.6)	0.842
13 years	0.1 (-2.6, 2.8)	0.936	0.1 (-2.3, 2.5)	0.935
Over-eating	N=692			
6 years	19.0 (6.0, 32.1)	**0.004**	15.9 (-1.0, 32.8)	0.065
7 years	18.1 (5.6, 30.7)	**0.005**	13.9 (-0.4, 28.2)	0.057
8 years	17.2 (5.0, 29.4)	**0.006**	12.1 (-0.2, 24.4)	0.053
10 years	15.2 (3.5, 27.0)	**0.011**	9.2 (-0.1, 18.5)	0.052
12 years	13.1 (1.4, 24.8)	**0.028**	6.9 (-0.3, 14.1)	0.062
13 years	12.0 (0.3, 23.6)	**0.044**	5.9 (-0.5, 12.3)	0.073

Estimated proportion mediated by over-eating in the association between PGS and BMI z-score with 95% CI using latent growth curve mediation analysis. Significance threshold (**bold**) set at 0.05.

Supplementary Text 5-1. R script for principal components adjusted scores.

```
require("broom")  # tidy regression output

require("mosaic")  # standardizing variables

load(file="BMI Yengo Vogel scores and BMI")

load(file="PCA717.Rdata")

PCA717 <- PCA717[, 1:11]

dataPRS <- merge(dataPRS, PCA717, by="IID")


# Adult score

Yengo=lm(PRS_Yengo~PC1+PC2+PC3+PC4+PC5+PC6+PC7+PC8+PC9+PC10,data=dataPRS)


dataPRS$Adultpred=predict(Yengo,dataPRS)

dataPRS$PRS_Yengo_adj = dataPRS$PRS_Yengo-dataPRS$Adultpred

dataPRS <- dataPRS %>%

  mutate(PRS_Yengo_adj = scale(PRS_Yengo_adj))


# Child score

Vogel=lm(PRS_Vogel~PC1+PC2+PC3+PC4+PC5+PC6+PC7+PC8+PC9+PC10,data=dataPRS)


dataPRS$Childpred=predict(Vogel,dataPRS)

dataPRS$PRS_Vogel_adj = dataPRS$PRS_Vogel-dataPRS$Childpred

dataPRS <- dataPRS %>%

  mutate(PRS_Vogel_adj = scale(PRS_Vogel_adj))
```

CHAPTER 6 - MANUSCRIPT 3

Article preface

This manuscript further explores the interplay between genetic and environmental factors in relation to BMI variations by leveraging the longitudinal anthropometric data collected in the QLSCD to assess which food or macronutrient intake can interact with the two PGS. This third manuscript relates to the third objective of the thesis: to estimate the association between preschool dietary (food, macronutrient, and energy) intakes and BMI (mean and change with age) displayed in school-aged children and identify the ability of adult- and child-derived PGS to modify those associations.

DG, LD and MB contributed to the development of the research question and the design of the study. LD and MB were involved in the data collection from QLSCD participants. DG performed the statistical analyses with insights and revisions from CG. DG was responsible for the first draft of the manuscript. All authors participated in discussions concerning the development of the study and were involved in the revisions to the manuscript and approved the submitted version.

Ethics approval (secondary analysis of QLSCD data): University of Ottawa Research Ethics board (ethics file number: H-01-23-8018, approval date: 10/02/2023, Appendix 1).

This manuscript was submitted to the Journal of Nutritional Science on November 19th, 2024. A copy of the confirmation email is available in Appendix 8. A STROBE reporting checklist for manuscript 3 is available in Appendix 9. The following manuscript was harmonized to the thesis format.

Title page

Title: Interaction between polygenic scores and dietary intake in relation to body mass index in Canadian children

Short title: PGS diet interaction in relation to BMI

Authors: Danick Goulet[1], Michel Boivin[2*], Christopher A. Gravel[1, 3, 4, 5], Julian Little[1], Beth K Potter[1], Lise Dubois[1*]

[1]School of Epidemiology and Public Health, University of Ottawa, Ottawa, ON, Canada

[2]Université Laval, Québec, QC, Canada.

[3]Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montréal, QC, Canada

[4]Department of Mathematics and Statistics, University of Ottawa, Ottawa, ON, Canada

[5]Data Literacy Research Institute, University of Ottawa, Ottawa, ON, Canada

***Correspondence**:

Lise Dubois : ldubois@uottawa.ca

Address: 600 Peter Morand Crescent, Ottawa (ON), Canada

Michel Boivin : michel.boivin@psy.ulaval.ca

Address : 2325 Rue des Bibliothèques, Québec (QC), Canada

Keywords: Obesity, polygenic score, moderation, nutrition, children's health, longitudinal analysis.

Abbreviations: BMI – body mass index; PGS – polygenic score; QLSCD – Quebec Longitudinal Study of Child Development; SSB – sugar-sweetened beverages

Abstract

The aim was to identify possible genetic moderation through polygenic scores (PGS) for body mass index (BMI), of the association between dietary intakes and BMI in children with the goal of identifying individuals for whom early prevention of obesity would be the most effective. The study sample included children who were part of a French-Canadian birth-cohort study. Height and weight were measured on seven occasions between ages 4 and 13 years. Food (juice and fruit drinks, sweets and snack foods, meats, and fruits and vegetables), macronutrient (proteins, lipids, carbohydrates), and energy intakes were measured up to 4 years. Linear mixed models were used to account for repeated BMI measurements. The consumption of juice and fruit drinks (in girls), sweets and snack foods, fruits and vegetables, proteins, lipids, carbohydrates and total energy were associated with BMI levels. Associations with BMI increased with age (kg/m^2 per year) for fruits and vegetables (β: -0.03, 95%CI: -0.06;-0.01), lipids (β: 0.11, 95%CI: 0.01;0.22), carbohydrates (β: 0.05, 95%CI: 0.01;0.08), and total energy (β: 0.07, 95%CI: 0.02;0.12), and with higher values of a BMI-PGS (kg/m^2 per SD) for proteins (β: 0.54, 95%CI: 0.03;1.06), lipids (β: 0.63, 95%CI: 0.12;1.13), and total energy (β: 0.32, 95%CI: 0.06;0.58). Leveraging a prospective study, we showed that the associations with higher BMI related to a lower consumption of fruits and vegetables, and a higher consumption of lipids, carbohydrates and energy increased with age in childhood. Obesity prevention efforts targeting dietary intake may be more beneficial for genetically susceptible children.

Introduction

The prevalence of obesity has increased continuously over the past decade for both adults and children in Canada [1, 2]. This is concerning since obesity leads to numerous health comorbidities in children, including hypertension, type 2 diabetes, or obstructive sleep apnea, which were previously thought to be present mostly in adults [3]. Furthermore, obesity in childhood is predictive of obesity in adulthood [4], as well as of higher risk of premature mortality [5] and worse cardiometabolic outcomes [6].

Given that food consumption is directly related to energy input, promoting healthy dietary habits are at the centre of childhood obesity prevention efforts [7]. Specifically, the intake of sugar-sweetened beverages (SSB) is the food group that is most consistently associated with weight gain in childhood [8]. Indeed, prospective studies have established the long-term association between SSB intake and weight gain, and clinical trials have shown that a reduction of SSB intake leads to a reduction in body weight in intervention settings [9, 10]. However, investigations on the consumption of foods such as fruits, vegetables, or sweets, largely using retrospective study designs, are characterized by inconsistent results [8, 11]. The lack of prospective studies, especially in children, is an important limitation in the evidence.

In recent years, the customary conceptualization of obesity evolved from a perspective centred around energy balance to a perspective that highlights the interaction between polygenic predisposition for obesity and neurodevelopmental dysfunctions modulated by the environment [12]. This approach is supported by the observation that genetic variants associated with obesity are mainly expressed in brain regions related to food behavioural rewards and addiction [13], and can be mapped to proteins located in the brain, some of which involve eating behaviour regulation [14]. Twin studies estimate that additive genetic factors account for up to 80% of individual differences in BMI [15]. Additionally, previous evidence shows systematically more weight change variance between than within pairs of monozygotic twins in response to overfeeding and negative energy balance, suggesting that genetic factors modulate the weight change in response to dietary habits [16]. More specifically, the intake of various types of foods (e.g., SSB, or fried foods) has been shown to modify the effect of individual genetic variants on obesity risk [17].

Polygenic scores (PGSs) are often used to measure genetic susceptibility to a trait by aggregating the effect of individual genetic variants into a single index. A lower intake of fruits and vegetables and a

higher intake of SSB are among the dietary factors with the most evidence supporting the presence of a moderation effect with PGS in relation to higher body weight [18]. Nevertheless, there is still conflicting results regarding those foods. Some studies reported evidence that PGS moderated the association of the consumption of SSB [19-21] and fruits and vegetables [21-23] with BMI, while others did not [22, 24, 25]. There is a lack of studies examining the possible role of PGSs in moderating the relation between food intake and BMI in children specifically. One study [26] observed higher associations of pizza and milk consumption with and BMI z-score in children with higher PGS values, and another [25] found a statistically significant interaction between the consumption of fibers (but not fruits and vegetables) and a PGS in relation to BMI.

Similarly, inconsistent evidence has been observed when examining macronutrient and energy intake. BMI-PGSs were found to moderate the association between consumption of lipids, proteins, carbohydrates, and total kilocalories and adiposity-related measures in some studies [27, 28]. In contrast, other studies did not observe such results [24, 29]. Overall, studies on the potential moderating role of genetic susceptibility to obesity, as measured by PGSs of the effect of food, macronutrient and energy intakes, have thus far focused almost exclusively on adult populations, used cross-sectional designs, and only involved PGSs aggregating the effect of a small subset of BMI-related variants.

The main objective of the present study was to estimate the association between dietary intakes and BMI (mean and change with age) displayed in school-aged children and examine the extent to which PGSs for BMI moderate those associations. Linear mixed models applied to repeated anthropometric measures obtained from a longitudinal study of children from Quebec (Canada) were used to test the hypothesis that food intakes (juice and fruit drinks, sweets and snack foods, meats, and fruits and vegetables), macronutrient intakes (proteins, lipids, carbohydrates) and energy intake are associated with change in BMI throughout childhood and that genetic susceptibility moderates those associations.

Methods

Study design and participants

From October 1997 to July 1998, investigators recruited a total of 2120 children to the Quebec Longitudinal Study of Child Development (QLSCD) which is designed to be representative of the regional birth distribution. Data collection occurred every year or every other year by questionnaires or interviews administered to the person most knowledgeable about the child (usually the mother), questionnaires administered to teachers, questionnaires administered to the child, or via direct measurements by trained assistants. Data collection covered a broad range of characteristics, but the attributes most related to the present study included family characteristics (e.g., socioeconomic and demographic characteristics), dietary habits (e.g., food frequency consumption, and eating behaviours), anthropometric data (self-reported or directly measured), and genetic information. A cohort profile published previously provides comprehensive details related to the QLSCD [30].

The study sample included 1793 participants with food intake data, and 1513 with macronutrient and energy intake data, at least one BMI measure, and covariate data. The subsample for analyses including genetic data consisted of 706 (food intake) and 646 (macronutrient and energy intake) participants with genetic data that passed quality control. All participants included in the macronutrient and energy intake analyses overlapped with those included in the food intake analyses. A flow chart depicting the study sample is provided in Supplementary Figure 6-1. This study was conducted according to the guidelines laid down in the Declaration of Helsinki. The data collections from the QLSCD were approved by the Ethics commitees of the Institut de la Statistique du Québec and the CHU Sainte-Justine Research Center. The secondary analysis of the data from the QLSCD was approved by the University of Ottawa Research Ethics Board (ethics file number: H-01-23-8018). Informed consent was obtained from parents before each wave of the data collection until age 10. At this point, informed consent was obtained directly from the participants themselves.

Anthropometric measurements

Anthropometric data were collected every one or two years throughout the QLSCD, either reported by a parent, reported by participants themselves, or directly measured. In this study, we focused on directly measured anthropometric data, which were obtained at ages 4, 6, 7, 8, 10, 12, and 13 years.

Trained technicians measured height and weight following a standardized protocol. Height and weight, computed as the mean of the two closest values from two to three measures, were used to calculate final individual BMI values.

Food, macronutrient and energy intake

The parent most knowledgeable about the participant was administered (interviewer-administered or self-administered) a short food frequency questionnaire multiple times during follow-up. We used data from four questionnaires completed at ages 1, 2, 3 and 4 years to obtain an average of the food intake in early childhood. These questionnaires contained 14 food items and asked the number of times per day/week the child ate specific foods. The questionnaire asked: In the past week, on average, how many times a week or how many times a day did the child eat the following foods…? The answers varied from: Never, 1-2, 3-4, 5-6 times per week, and 1, 2+ times per day. Intake values were transformed to the number of times a food was consumed per day as a continuous scale by assuming the middle point for responses categorized as a range (e.g., 3-4 times per week transformed to 0.5 times per day) and assuming the minimum point for open-ended responses (e.g., 2+ times per day transformed to 2 times per day). After grouping certain food items, we included in our analyses the consumption of sweets and snack food (including pastry, sweets, cookies, and potato chips), juice and fruit drinks, meats (including poultry, meat, and fish), and fruits and vegetables (including potatoes). The individual items were summed to obtain a consumption per day for a specific food group. Food intake for each of the four groups was then averaged over each of the time points available per participant to obtain one intake measure per participant, per food group.

The nutrition investigation was based on 24-hour multi-pass dietary recall interviews [31] conducted at home by trained dietitians when participants were 4 years old. The mother was asked what type and quantities of food the child had eaten in the past 24 hours. A description of the process involved in deriving macronutrient and energy estimates from the dietary recall interview is available in an previous publication [32]. In this study, the total amount of energy and intake of carbohydrates, proteins and lipids are of interest. Macronutrient and energy intake values represent the consumption reported as number of units per day where a unit represents 100g for proteins, lipids, and carbohydrates, and 1000 for kilocalories.

Genotyping and polygenic score construction

Investigators collected biological material, including blood samples used to derive genetic data, when participants were 10 years old from a subsample of 992 participants. The genetic quality control applied to the data and the PGS construction process are described in more detail in two previous publications [33, 34]. A total of 816 participants had genetic data available and passed genetic quality control for 8,465,216 variants. The adult-derived PGS and child-derived PGS were constructed using publicly available summary statistics from genome-wide association studies of adults [35] and children [36] with the PRS-CS method [37]. Both scores were standardized and adjusted for population stratification with 613,732 genetic variants for the adult-derived PGS, and 689,789 genetic variants for the child-derived PGS.

Statistical analysis

Baseline sociodemographic characteristics of the study participants are summarized as proportions for categorical variables and means (standard deviation) for continuous variables. Continuous variables were deemed normally distributed by visually inspecting histograms. Baseline sociodemographic characteristics were also compared between the study participants with genetic data and the QLSCD participants excluded from the study. Study participants' characteristics were compared to those of QLSCD non-participants using a Chi-square test, and one-way ANOVA, depending on the variable type.

The association between the child and adult-based PGS and BMI level (intercept) and changes in BMI with age (slope) was derived using linear mixed models to account for the repeated BMI measurements. Both models (equation A) included BMI as the outcome variable, with age, an order two polynomial for age, the PGS, the interaction term between the PGS and age, and five baseline covariates (sex, birth weight, maternal education, household income, and whether the mother was born in Canada) as predictor variables.

The predictive role of food, macronutrient and energy intake in BMI was assessed using linear mixed models (equation B). In addition to the intake variables and their interaction with the first polynomial of age, models also included an order two polynomial for age and the five covariates. The estimated coefficient for the interaction term between the intake variables and age corresponds to its association with BMI slope. The estimated coefficient for the intake variables represents the association with BMI level. We also assessed the association between the intake variables and BMI at specific time points by

setting the intercept at 4, 8 and 13 years of age. We examined whether the associations between dietary intakes and BMI differed as a function of sex by adding an interaction term to components already included in the model: 1) dietary intakes and 2) the interaction between dietary intakes and age. Stratified estimates are presented when necessary.

We assessed whether the association between dietary intakes and BMI differed based on a participant's genetic susceptibility to higher BMI. From the models displayed in equation B, we added the adult or the child-derived PGSs and an interaction term between the PGS and the food intake variables as shown in equation C. We examined if the interaction between dietary intakes and the PGSs changed based on age or sex by adding three-way interactions to dietary intake and PGS.

We further performed two supplementary analyses. First, we added total energy intake as a covariate to the regression models for the macronutrient intake variables to estimate associations independently of the energy component. Second, inverse probability weighting was used to reduce the impact of informative loss to follow up on the results. Individual propensity scores to participation in the current study were estimated on the bases of baseline characteristics (sex, maternal education, birth weight, sufficient income and whether the mother was born in Canada). Weights were estimated for each time point to allow differing weights across time points to account for attrition. To obtain a weighted study sample analog to the initial QLSCD sample, we used the average treatment effect to estimate the propensity scores using the R package "twang" [38]. Statistical significance was determined using the student's t-test.

(A) $$BMI_{ij} = \beta_0 + \beta_1 T_{ij} + \beta_2 T^2{}_{ij} + \beta_3 PGS_i + \beta_4 PGS\ x\ T_i + \beta_{5-9} C_i + S_{0i} + S_{1i} T_{ij} + \varepsilon_{ij}$$

(B) $$BMI_{ij} = \beta_0 + \beta_1 T_{ij} + \beta_2 T^2{}_{ij} + \beta_3 Intake_i + \beta_4 Intake\ x\ T_i + \beta_{5-9} C_i + S_{0i} + S_{1i} T_{ij} + \varepsilon_{ij}$$

(C) $$BMI_{ij} = \beta_0 + \beta_1 T_{ij} + \beta_2 T^2{}_{ij} + \beta_3 Intake_i + \beta_4 Intake\ x\ T_i + \beta_5 PGS_i + \beta_6 PGS\ x\ T_i + \beta_7 Intake\ x\ PGS_i + \beta_{8-12} C_i + S_{0i} + S_{1i} T_{ij} + \varepsilon_{ij}$$

with $\varepsilon_{ij} \sim N(0\ ,\ \sigma^2)$, $S_{0i} \sim N(0\ , \sigma^2_{S_0})$, and $S_{1i} \sim N(0\ , \sigma^2_{S_1})$

where i = 1,…,n individuals and j=1, …, 7 measurements.
T: Age (years).

PGS: Child or adult-derived PGS.

Intake: Dietary variable.

C: Covariates: Sex, birthweight, maternal education, sufficient income, mother born in Canada.

Results

A description of the sociodemographic characteristics of participants included in the main analyses is presented in Table 6-1. In both the food intake and macronutrient and energy intake samples, the sex distribution was similar (50.1% and 49.2% girls), more than half of the participants had mothers with post-secondary education (69.3% and 69.3%) and were part of a household with an income below $60,000 (69.7% and 69.5%). Additionally, most of the participants' mothers were aged 20-34 years at delivery (83.8% and 84.1%) and were born in Canada (90.7% and 91.3%). Compared to non-participants, participants included in the genetically informed analysis were mostly composed of girls, of mothers with higher education, and from higher income households (Supplementary Tables 6-1 and 6-2).

The trajectories of BMI throughout childhood are presented in Figure 6-1. Both the child-derived (Figure 6-1A) and the adult-derived (Figure 6-1B) PGSs were strongly associated with BMI slope, meaning that genetic susceptibility predicted a BMI increase with age. An increase in one SD in the child-derived PGS was associated with a mean yearly BMI increase of 0.06 kg/m^2 (95% CI: 0.03; 0.09, $p<0.001$), whereas a one SD increase in the adult-derived PGS yielded a mean yearly BMI increase of 0.11 kg/m^2 (95% CI: 0.08; 0.14, $p<0.001$). These associations did not differ between girls and boys (not shown).

Models described in equation B estimated age-specific associations between dietary intakes in early life and BMI from age 4-13 years, and whether those associations changed with age (BMI slope). Results from this analysis are presented in Table 6-2. Only fruit and vegetable intake from ages 1-4 years was significantly associated with BMI slope. A one time increase per day in intake of fruits and vegetables was associated with a yearly decrease in BMI of 0.03 kg/m^2 (95% CI: -0.06; -0.01, $p = 0.012$) of age. The association between fruit and vegetable intake and BMI level was stronger with increasing child

age from -0.05 kg/m^2 (95% CI: -0.16; 0.06, p = 0.365) at 4 years, to -0.37 kg/m^2 (95% CI: -0.62; -0.11, p = 0.005) at 13 years. The consumption of meats, and sweets and snack food was not associated with BMI slope, although a higher consumption of sweets and snack food was significantly related to lower BMI at 4 years (β: -0.36, 95% CI: -0.61; -0.12, p = 0.004) (Table 6-3). With the exception of juice and fruit drinks (p=0.036), no food intake association with BMI slope varied by sex. Consumption of juice and fruit drinks was significantly associated with positive BMI slope and with BMI level at 13 years in girls only. A one time per day increase in intake of juice and fruit drinks was associated with an increased BMI change with age of 0.08 kg/m^2 (95% CI: 0.02; 0.14, p = 0.008) per year and with an increase in age-specific BMI level of 0.68 kg/m^2 (95% CI: 0.10; 1.27, p = 0.022) at 13 years for girls (Supplementary Table 6-3).

Consumption of higher amounts of proteins, lipids, carbohydrates, and kilocalories at age 4 years was significantly associated with higher BMI from 4 years to 13 years (Table 6-2). Additionally, higher intakes of lipids, carbohydrates, and kilocalories were significantly associated with increased BMI change with age (Table 6-2). The association with the slope of BMI was 0.11 kg/m^2 (95% CI: 0.01; 0.22, p = 0.031), 0.05 kg/m^2 (95% CI: 0.01; 0.08, p = 0.009), and 0.07 kg/m^2 (95% CI: 0.02; 0.12, p = 0.007) per daily intake increases of 100g of lipids, 100g of carbohydrates, and 1000 kilocalories, respectively. The associations between macronutrient and energy intake and BMI did not differ by sex (all p>0.05, not shown). Additional analyses revealed that the association of consumption of lipids and carbohydrates with BMI still increased with age after taking into account total energy consumption (Supplementary Table 6-4). However, the age-specific associations of all three macronutrients (proteins, lipids, and carbohydrates) with BMI were attenuated and were no longer statistically significant when accounting for total energy consumption, except for carbohydrates at 13 years (Supplementary Table 6-4).

None of the interactions between food, macronutrient, or energy intake and the child-derived PGS in relation to BMI were statistically significant. There were significant interactions between the adult-derived PGS and the consumption of meats, proteins, lipids, and kilocalories in relation to BMI (Table 6-3). A one SD increase of the PGS moderated those associations so that such consumption yielded 0.46 kg/m^2 (95% CI: 0.08; 0.85, p = 0.018), 0.54 kg/m^2 (95% CI: 0.03; 1.06, p = 0.038), 0.63 kg/m^2 (95% CI: 0.12; 1.13, p = 0.015), and 0.32 kg/m^2 (95% CI: 0.06; 0.58, p = 0.017), respectively, across childhood. The interactions did not differ by sex (p>0.05, not shown) except for juice and fruit drinks and the adult-derived PGS (p = 0.031), with a significant (negative) interaction with the PGS in girls

only (Supplementary Table 6-5). A one SD increase of the PGS diminished the association between juice and fruit drinks intake and BMI by 0.35 kg/m^2 (95% CI: -0.69; -0.01, p = 0.044) in girls. Therefore, girls at lower genetic risk were more susceptible to the effect of juice and fruit drink consumption on BMI.

Table 6-4 displays the association between increased daily consumption of meat (one more serving per day), protein (additional 100g per day), lipid (additional 100g per day), and kilocalorie (1000 additional kilocalories per day) and BMI, by age (4, 8, and 13 years), reported separately by selected values of the adult-derived PGS (mean score -1 SD, mean score, mean score +1 SD). Higher protein, lipid, and kilocalorie intakes were associated with increased BMI at the mean PGS value and for the mean +1 SD value at all ages. The association between increased frequency of meat intake and higher BMI was stronger at older ages and with higher values of the adult-derived PGS, without statistically significant associations. The positive interactions between the adult-derived PGS and the consumption of proteins and lipids in relation to increases in BMI level were still retained once total energy intake was included in the models (Supplementary Table 6-6). However, age-specific associations between protein and lipid intakes and BMI by PGS values were attenuated and not statistically significant after adjusting for total energy intake, except for protein intake in those of high genetic risk at 8 years (Supplementary Table 6-6).

Finally, applying inverse probability weighting to attenuate the impact of informative loss to follow up did not change the overall results, with only minimal effects on the estimated coefficients (Supplementary Tables 6-7 to 6-9).

Discussion

The present study estimated the association between specific food, macronutrient and energy intakes assessed in early childhood and BMI measured later in childhood and determined if genetic susceptibility to obesity, measured through genetic scores, moderated these associations. Consumption of fruits and vegetables (age 1-4 years), and of proteins, lipids, carbohydrates, and total energy (age 4 years) were associated with BMI from 4 years to 13 years. Most notably, our findings revealed (1) an increase with age of the associations of fruits and vegetables, lipids, carbohydrates, proteins, and energy intake with BMI, (2) that a low genetic susceptibility to obesity protects from an increase in

BMI associated with higher intakes of proteins, lipids, and total energy, and (3) that taking into account total energy intake dissipates most of the macronutrients contribution to mean BMI, but not the related slope of BMI and interaction with genetic susceptibility.

Two recent-meta-analyses of 40 [9] and 20 [11] studies, including a large proportion of prospective cohorts supported the association of a higher intake of SSB with increased BMI and risk of obesity in childhood. Our study builds on these findings by showing that in childhood, the association between juice and fruit drink intake and BMI increases with age, but only in girls. A similar association was found exclusively in girls [39] and women [40] before, but our findings should be interpreted with caution as the aforementioned meta-analyses did not investigate differences by sex. Importantly, consumption of juice and fruit drinks in our study, which was measured up to age 4 years, did not include carbonated beverages.

Our observation that a higher consumption of snacks is associated with a lower BMI at early age aligns with previous evidence. Higher intake of sweets and snacks tends to lead to lower BMI with several studies reporting a decrease in body weight related to higher intake of sweets [11]. Although the associations of juice and fruit drinks and snacks in opposite directions seem counter-intuitive, a possible explanation for this observation stems from the context of the consumption of specific types of food. Sweets and snacks are often eaten between meals, which could lead to a decrease in appetite at subsequent meals [11]. Furthermore, the association between higher intake of fruits and vegetables and lower BMI grew stronger with age. The consumption of fruits and vegetables is encouraged in children to decrease the risk of obesity [7], However, studies found only weak to no evidence for an association between a higher consumption of fruits and vegetables and a reduction in BMI in children [11, 41].Thus, our study adds robust evidence in favour of higher fruits and vegetables consumption recommendations in childhood.

Our study showed a clear association of higher intakes of proteins, lipids, carbohydrates, and total energy at age 4 years with higher BMI throughout childhood. It is generally accepted that energy intakes play an important part in the etiology of the accumulation of body fat, but there is still uncertainty about exactly what extent macronutrients contribute to the buildup of body fat in the body [42]. Although there is evidence that carbohydrate and lipid intake affect weight change in adults [43, 44], the current literature concerning how the composition of macronutrients in diets influence body weight in childhood is limited [45]. Our study adds to the current literature by showing that the association between the consumption of macronutrients (lipids, carbohydrates, and proteins) and BMI

is seemingly present only when not accounting for energy intake. This aligns with a previous meta-analysis that reported that the carbohydrate composition of a diet is not significantly associated with the risk of obesity [44], suggesting that most of the attributed risk of increased consumption of carbohydrates is related to the increase in energy intake. Energy intake may attenuate the associations between macronutrients and BMI, but the associations with the slope of BMI remained unchanged. As the explanation behind the conservation of the associations with BMI slope, but not the associations with age-specific BMI, remains elusive, future research will be needed to investigate this question in more detail. Finally, the absence of significant association of protein intake at age 4 years with BMI slope in conjunction with strong associations from 4 to 13 years also aligns with the literature. Higher consumption of protein in the first years of life (0-24 months) was associated with higher BMI later in childhood in two recent meta-analyses of prospective studies [46, 47].

Not everyone exposed to the same modern obesogenic environment characterized by easily accessible unhealthy eating options and a sedentary lifestyle develops obesity. This likely stems from the complex interplay between environmental and genetic factors [48]. Identifying interactions between genetic susceptibility to obesity and dietary factors could be important in revealing individuals who are more likely to benefit from specific dietary interventions or prevention methods. We found that the adult-derived PGS modified the association between consumption of proteins, lipids, and total energy and BMI. Previous investigations on the interaction between PGS for obesity and macronutrients are mixed: some studies observed significant interactions with proteins [27], lipids [22, 27, 28], and energy intake [27], and others did not [24, 29]. Our results offer valuable insights considering that only one of the aforementioned studies was prospective [22] and all examined dietary intake and BMI data in adulthood. Interestingly, the previously mentioned studies used different methods to adjust energy intake. One adjusted total macronutrient intakes with energy consumption [27], two [24, 28] examined proportions of the diet allotted to specific macronutrients, and one did not take into account energy intake [29]. Our study aligns with the only other study [27] that found interactions with total protein and lipid intake after adjustment for total energy intake. Our observation that genetic susceptibility also modified the association between the consumption of meat and BMI could stem from the interaction with protein intake.

Consumption of SSB is the type of food that has been the most observed to interact with a PGS related to BMI [19, 20, 49]. PGSs for BMI have also been shown to modify the association of fruit and vegetable intake [21-23, 50] or a more broadly healthy eating diet with BMI. The consumption of juice

and fruit drinks, sweets and snack foods, and fruits and vegetables was not found to interact with genetic susceptibility in our study. This could be explained by differences across studies in how the consumption of food was measured; for example, our beverage variable incorporated juice and fruit drinks but not carbonated sweetened beverages, as noted. Overall, we found that only the PGS derived from adult data, but not from children, modified the relationship between dietary intakes and BMI. This is consistent with the observation that an adult-derived PGS may be able to capture higher BMI variation compared to child-derived PGS as early as 3 years of age [51].

Generally, interaction studies provide clarity to what sub-group of a population or in what circumstances an exposure influences a trait [48]. Our results suggest that protein, lipid and energy intake affect body weight, especially in genetically susceptible children. Considering obesity interventions could be better tailored to individuals for better results [52], genetic susceptibility may indicate which individuals to target. Still, the use of PGS in a clinical setting to identify individuals at higher risk of obesity is not common. PGS predicting BMI cannot identify individuals at higher risk of obesity with appropriate accuracy [53]. However, the presence of an interaction between dietary intakes and a PGS in relation to BMI implies that it may be possible to reduce the effect of genetic factors on body weight by acting on the dietary habits of children. Thus, understanding which foods and macronutrients could have even more impact on genetically susceptible children remains an important undertaking.

The use of a longitudinal design is an important strength of the study that allowed us to portray the influence of dietary intake on body weight throughout childhood in more detail, as a cross-sectional design would have missed crucial patterns of associations. We also used PGSs that were derived using a method incorporating hundreds of thousands of genetic variants, capturing a more diverse array of biological mechanisms related to obesity compared to methods limited to a few dozen variants. Additionally, completing the analyses adjusting for total energy intake revealed important distinctions that are often omitted when simply using one of the two possibilities. Although self-report measures of food intake were used with food frequency questionnaires, we leveraged multiple time points to obtain a more robust estimate. In addition, we used 24h dietary recall interviews to estimate macronutrient and energy intake, which produce good estimations of population dietary intakes and are a cornerstone of national nutrition surveys in Canada [54]. We also note some limitations. First, despite adjusting regression models with potential confounders and opting for a weighted approach to mitigate attrition bias, unmeasured factors that could not be included as covariates (e.g., level of physical activity) or in

the weighting approach could still bias the resulting associations. Second, although BMI is commonly used to represent fat accumulation in the body, it remains an imperfect measure that hardly differentiates between lean and fat mass and is blind to the location of fat mass, which can have a different impact on health outcomes. Third, although we tested if sex modified the associations we studied, some of those tests were three-way interactions (e.g., PGS, energy intake, sex) which usually require large sample sizes to obtain sufficient power. Therefore, larger studies may be needed to explore these questions more specifically. Finally, considering we used a PGS based on summary statistics from populations of European ancestry, our results may not be generalizable to populations of non-European ancestry.

Overall, using a longitudinal design and incorporating genetic factors in our study provide a robust approach to investigating the association between dietary intakes and BMI change in childhood. We observed that lower consumption of fruits, and vegetables as well as a higher consumption of macronutrients (proteins, lipids, carbohydrates), and total energy was associated with higher BMI at different points throughout childhood. We also note that the association between some dietary intakes and BMI increased with age (fruits and vegetables, lipids, carbohydrates, and total energy), or increased with genetic susceptibility (proteins, lipids, and total energy). Those results highlight that the effectiveness of obesity prevention efforts that target dietary intake may depend on individual genetic susceptibility to higher body weight.

Tables and Figures

Table 6-1. Characteristics of study participants with food intake and macronutrient and energy intake data available.

Characteristic	*Food intake N=1793*		*Macronutrient intake N=1513*	
Categorical variables	%	n	%	n
Sex				
Girl	50.1	898	49.2 (745)	745
Boy	49.9	895	50.8 (768)	768
Preterm birth				
Yes	4.9	87	4.8 (73)	73
No	95.1	1706	95.2 (1440)	1440
Maternal age at birth				
≤ 20 years	2.6	46	2.3 (35)	35
20-34 years	83.8	1502	84.1 (1272)	1272
≥ 35 years	13.7	245	13.6 (206)	206
Maternal education				
Secondary education or less	43.2	775	42.3 (640)	640
Post-sec. education	56.8	1018	57.7 (873)	873
Household income				
< 30 000$	28.2	505	27.3 (413)	413
30 000 – <60 000$	41.5	744	42.2 (638)	638
60 000 – <80 000$	15.7	281	16.0 (242)	242
≥ 80 000$	14.7	263	14.5 (220)	220
Born in Canada	90.7	1626	91.3 (1381)	1381
Continuous variables	Mean	SD	Mean	SD
Birth weight, kg	3.40	0.50	3.40	0.50
BMI at 4 years, kg/m^2	15.76	1.81	15.76	1.81
BMI at 8 years, kg/m^2	16.74	2.59	16.74	2.59
BMI at 13 years, kg/m^2	20.96	4.16	20.10	4.02
Juice and fruit drinks, times/day	1.39	0.49	-	-
Sweets and snacks, times/day	0.66	0.34	-	-
Meats, times/day	1.03	0.30	-	-
Fruits and vegetables, times/day	2.74	0.81	-	-
Proteins, 100g per day	-	-	0.60	0.23

Lipids, 100g per day	-	-	0.58	0.22
Carbohydrates, 100g per day	-	-	2.19	0.68
Total energy, 1000 kcals per day	-	-	1.61	0.46

N, sample size. SD, standard deviation. Kcals, kilocalories.

Table 6-2. Age-specific association between food, macronutrient and energy intake, and BMI at 4, 8, and 13 years, and association with change in BMI with age.

Model	(A) BMI 4 years		(B) BMI 8 years		(C) BMI 13 years		(D) BMI slope	
	β (95% CI)	**p-value**	**β (95% CI)**	**p-value**	**β (95% CI)**	**p-value**	**β (95% CI)**	**p-value**
Food intake								
Juice and fruit drinks	-0.03 (-0.20, 0.14)	0.760	0.13 (-0.11, 0.36)	0.303	0.35 (-0.08, 0.78)	0.108	0.04 (-0.00, 0.09)	0.073
Sweets and snacks	-0.36 (-0.61, -0.12)	**0.004**	-0.33 (-0.67, 0.02)	0.062	-0.27 (-0.88, 0.34)	0.391	0.01 (-0.05, 0.08)	0.745
Meats	-0.07 (-0.34, 0.21)	0.638	0.09 (-0.29, 0.48)	0.630	0.20 (-0.49, 0.90)	0.567	0.03 (-0.04, 0.10)	0.416
Fruits and vegetables	-0.05 (-0.16, 0.06)	0.365	-0.18 (-0.33, -0.03)	**0.016**	-0.37 (-0.62, -0.11)	**0.005**	-0.03 (-0.06, -0.01)	**0.012**
Macronutrient intake								
Proteins	0.81 (0.44, 1.18)	**<0.001**	0.98 (0.43, 1.52)	**<0.001**	1.19 (0.22, 2.17)	**0.016**	0.04 (-0.06, 0.14)	0.410
Lipids	0.76 (0.37, 1.15)	**<0.001**	1.23 (0.67, 1.80)	**<0.001**	1.79 (0.79, 2.79)	**<0.001**	0.11 (0.01, 0.22)	**0.031**
Carbohydrates	0.31 (0.18, 0.44)	**<0.001**	0.49 (0.30, 0.68)	**<0.001**	0.72 (0.39, 1.05)	**<0.001**	0.05 (0.01, 0.08)	**0.009**
Total energy	0.51 (0.32, 0.70)	**<0.001**	0.79 (0.52, 1.06)	**<0.001**	1.13 (0.65, 1.61)	**<0.001**	0.07 (0.02, 0.12)	**0.007**

BMI, body mass index. CI, confidence interval. Estimated association between food, macronutrient and energy intakes and (A) BMI at 4 years, (B) BMI at 8 years, (C) BMI at 13 years, and (D) BMI slope (intake x age interaction) with 95% CI using linear mixed models. All coefficients estimated based on the model described in equation B, which includes the following covariates: sex, birthweight, maternal education, sufficient income, and whether the mother was born in Canada. Significance threshold (**bold**) set at 0.05.

Table 6-3. Coefficient for the interaction of food, macronutrient and energy intake with PGSs.

Model	Child-derived PGS		Adult-derived PGS	
	β (95% CI)	**p-value**	**β (95% CI)**	**p-value**
Juice and fruit drinks	0.08 (-0.14, 0.31)	0.482	-0.14 (-0.38, 0.11)	0.285
Sweets and snacks	0.03 (-0.30, 0.37)	0.851	-0.13 (-0.49, 0.23)	0.474
Meats	0.07 (-0.33, 0.47)	0.741	0.46 (0.08, 0.85)	**0.018**
Fruits and vegetables	-0.01 (-0.17, 0.14)	0.866	0.07 (-0.09, 0.23)	0.392
Proteins	0.05 (-0.50, 0.59)	0.866	0.54 (0.03, 1.06)	**0.038**
Lipids	0.53 (-0.07, 1.12)	0.083	0.63 (0.12, 1.13)	**0.015**
Carbohydrates	0.12 (-0.06, 0.31)	0.197	0.16 (-0.03, 0.36)	0.107
Total energy	0.19 (-0.09, 0.47)	0.178	0.32 (0.06, 0.58)	**0.017**

PGS, polygenic score. CI, confidence interval. Estimated interaction of food, macronutrient and energy intakes with the child-derived and adult-derived PGS in relation to BMI level with 95% CI using linear mixed models. All coefficients estimated based on the model described in equation C, which includes the following covariates: sex, birthweight, maternal education, sufficient income, and whether the mother was born in Canada. Significance threshold (**bold**) set at 0.05.

Table 6-4. Age-specific association of food, macronutrient and energy intake with BMI by PGS value.

Model	PGS: mean score -1 SD		PGS: mean score		PGS: mean score +1 SD	
	β (95% CI)	**p-value**	**β (95% CI)**	**p-value**	**β (95% CI)**	**p-value**
4 years						
Proteins	0.47 (-0.27, 1.21)	0.210	1.01 (0.47, 1.55)	**<0.001**	1.55 (0.80, 2.30)	**<0.001**
Lipids	0.28 (-0.47, 1.02)	0.468	0.90 (0.35, 1.45)	**0.001**	1.52 (0.77, 2.28)	**<0.001**
Meats	-0.49 (-1.05, 0.07)	0.087	-0.04 (-0.45, 0.38)	0.862	0.42 (-0.15, 0.99)	0.151
Total energy	0.27 (-0.10, 0.64)	0.154	0.58 (0.31, 0.86)	**<0.001**	0.90 (0.52, 1.28)	**<0.001**
8 years						
Proteins	0.73 (-0.16, 1.61)	0.107	1.25 (0.53, 1.98)	**0.001**	1.78 (0.89, 2.68)	**<0.001**
Lipids	0.65 (-0.25, 1.55)	0.155	1.26 (0.52, 2.00)	**0.001**	1.87 (0.96, 2.77)	**<0.001**
Meats	-0.27 (-0.94, 0.40)	0.427	0.16 (-0.39, 0.71)	0.577	0.58 (-0.09, 1.26)	0.092
Total energy	0.66 (0.22, 1.10)	**0.003**	0.97 (0.61, 1.33)	**<0.001**	1.28 (0.83, 1.73)	**<0.001**
13 years						
Proteins	1.05 (-0.34, 2.43)	0.138	1.60 (0.31, 2.88)	**0.015**	2.14 (0.75, 3.54)	**0.003**
Lipids	1.09 (-0.32, 2.50)	0.130	1.72 (0.40, 3.04)	**0.011**	2.34 (0.93, 3.76)	**0.001**
Meats	-0.10 (-1.14, 0.95)	0.855	0.37 (-0.60, 1.35)	0.453	0.85 (-0.21, 1.90)	0.116
Total energy	1.13 (0.44, 1.82)	**0.001**	1.45 (0.80, 2.09)	**<0.001**	1.76 (1.07, 2.46)	**<0.001**

PGS, polygenic score. CI, confidence interval. Estimated association of food, macronutrient and energy intakes with BMI according to age and the adult-derived PGS with 95% CI using linear mixed models. All coefficients estimated based on the model described in equation C, which includes the following covariates: sex, birthweight, maternal education, sufficient income, and whether the mother was born in Canada. Significance threshold (**bold**) set at 0.05.

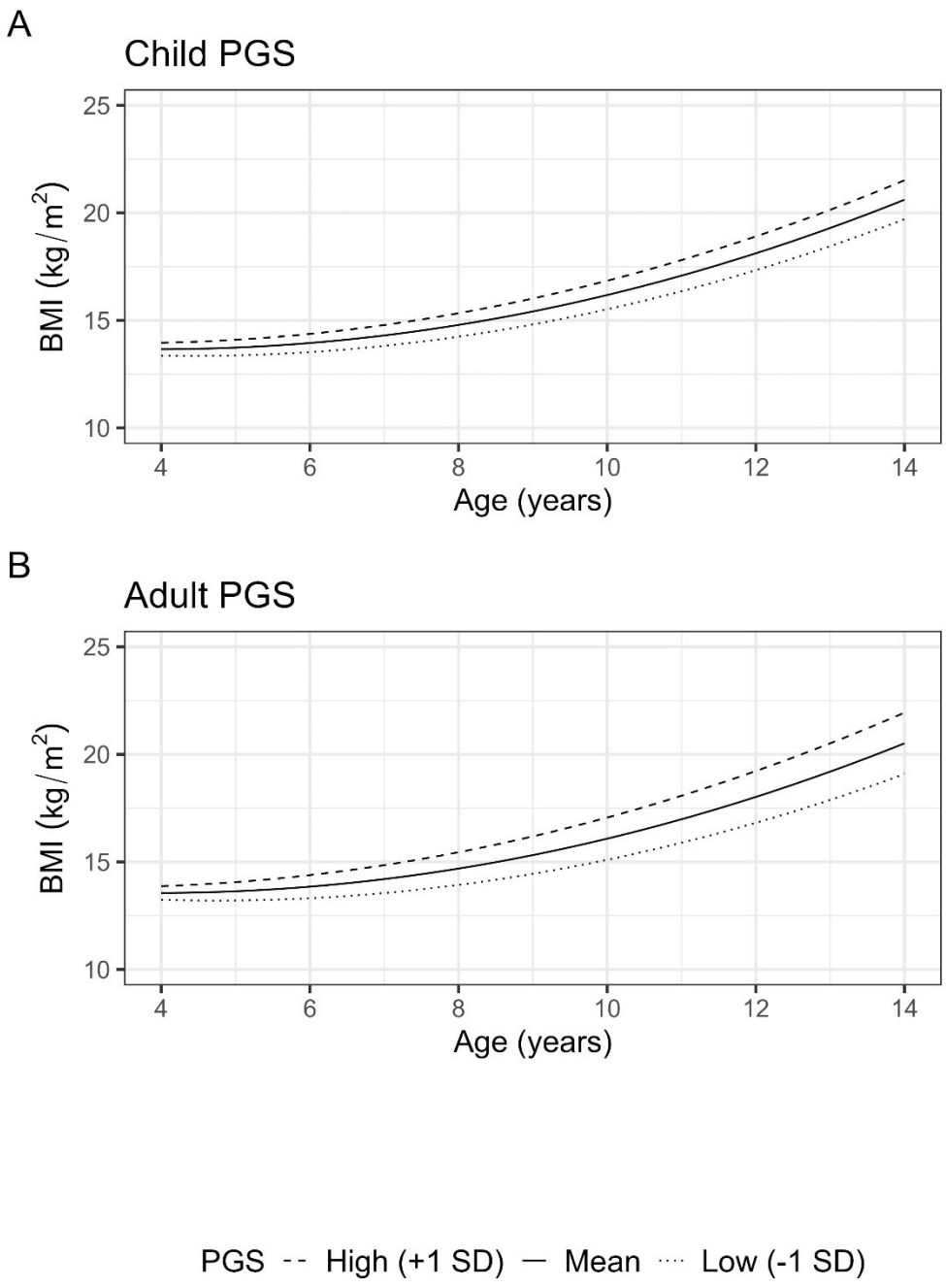


Figure 6-1. Predicted trajectories of BMI throughout childhood by (A) the child-derived PGS, and (B) the adult-derived PGS. **Predicted value of BMI by age based on the value of PGS of -1 SD, mean, +1 SD, based on equation A.**

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

AUTHOR CONTRIBUTIONS

DG, LD and MB contributed to the conception and design of the study. LD and MB were involved in the data collection from QLSCD participants. DG performed the statistical analyses with from CG. DG was responsible for the first draft of the manuscript. All authors participated in discussions concerning the development of the study and were involved in the revisions to the manuscript and approved the submitted version.

ACKNOWLEDGEMENTS

We also acknowledge the contribution of Till Andlauer, Stéphane Paquin, Geneviève Morneau-Vaillancourt, Isabelle Ouellet-Morin and Michel Boivin, who were involved in the quality control of the genetic data of the QLSCD participants that are used in the research. We are grateful to the QLSCD participants and their families who took part in the various data collection rounds over the years.

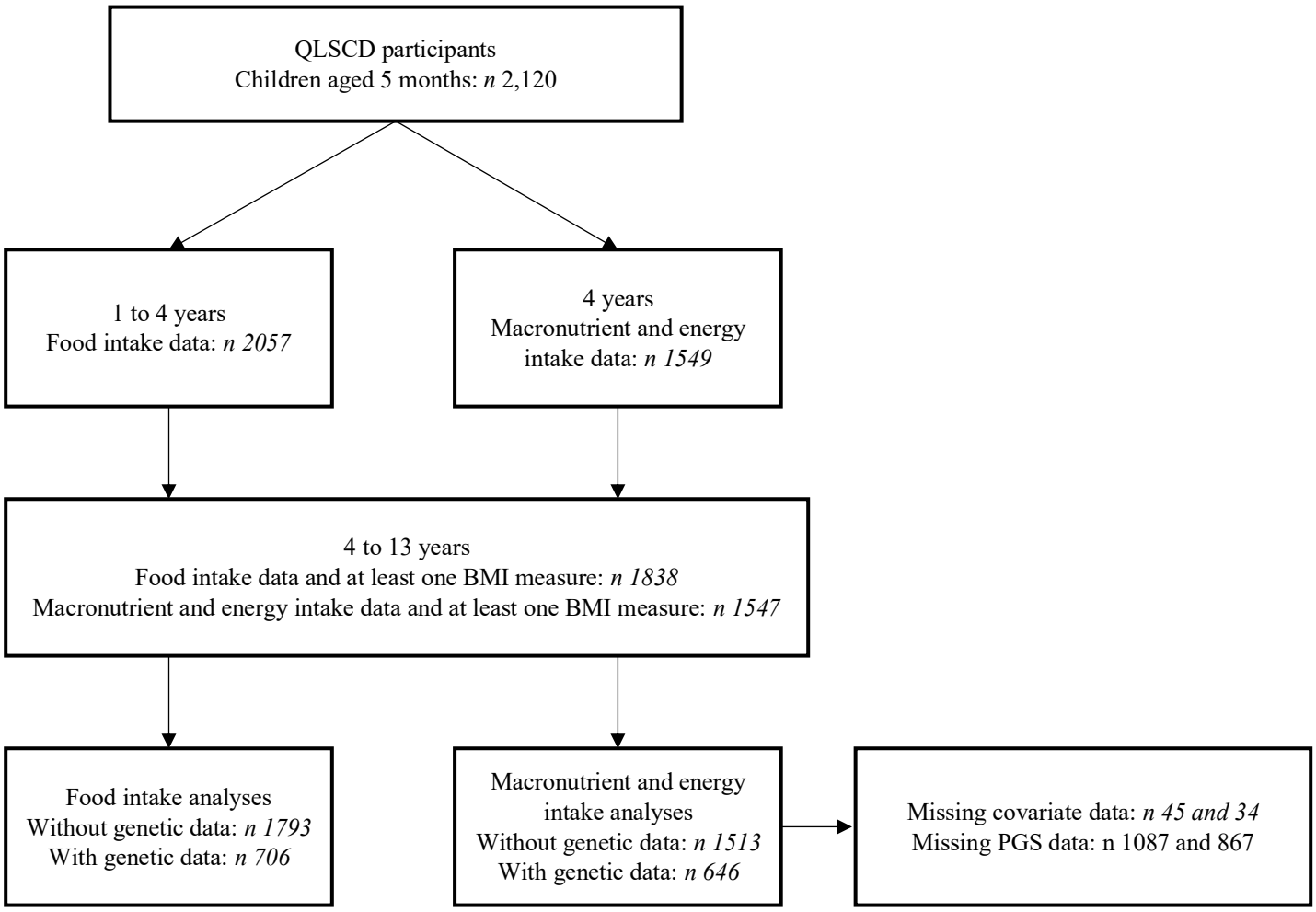
FUNDING

This work was supported by a CIHR operating grant (#165964). The funders were not involved in the design or conduct of the study; collection, management, analysis, or interpretation of the data; or preparation, review, or approval of the manuscript. The analyses were performed using data from the Quebec Longitudinal Study of Child Development (QLSCD), conducted by Sante Quebec, a division of the Institut de la Statistique du Quebec (ISQ) and funded by the Ministry of Health and Social Services of Quebec.

DATA AVAILABILITY STATEMENT

The data used to support the findings of this study was obtained from the Québec Longitudinal Study of Child Development, conducted by Santé Québec, a division of the Institut de la Statistique du Québec and may be released upon application to the Institut de la Statistique du Québec, through the Zone de recherche at: https://statistique.quebec.ca/fr/institut/services-recherche#/accueil

Supplementary materials



Supplementary Figure 6-1. Flow chart depicting QLSCD study sample.

Supplementary Table 6-1. Comparisons of characteristics of study participants with food intake and genetic data and excluded QLSCD participants.

Characteristic	*Study sample N=706*	*Excluded QLSCD N=1414*	*p-value*	*SMD*
Categorical variables, % (n)				
Sex			**<0.001**	0.182
Female	55.1 (389)	46.0 (651)		
Male	44.9 (317)	54.0 (763)		
Preterm birth			0.738	0.020
Yes	5.0 (35)	4.5 (64)		
No	95.0 (671)	95.5 (1350)		
Maternal age at birth			0.092	0.106
≤ 20 years	1.7 (12)	3.3 (47)		
20-34 years	84.3 (595)	83.5 (1180)		
≥ 35 years	14.0 (99)	13.2 (186)		
Maternal education			**0.003**	0.140
Secondary education or less	39.8 (281)	46.7 (662)		
Post-sec. education	60.2 (425)	53.3 (752)		
Household income			**0.002**	0.181
< 30 000$	25.1 (177)	32.3 (445)		
30 000 – <60 000$	41.2 (291)	40.1 (552)		
60 000 – <80 000$	17.1 (121)	15.0 (206)		
≥ 80 000$	16.6 (117)	12.6 (173)		
Born in Canada	97.7 (690)	83.2 (1177)	**<0.001**	0.510
Continuous variables, mean ± SD				
Birth weight, kg	3.42 ± 0.49	3.40 ± 0.50	0.354	0.043
Juice and fruit drinks, times/day	1.38 ± 0.48	1.40 ± 0.50	0.375	0.041
Sweets and snacks, times/day	0.66 ± 0.35	0.66 ± 0.34	0.917	0.005
Meats, times/day	1.02 ± 0.29	1.04 ± 0.32	0.176	0.064
Fruits and vegetables, times/day	2.83 ± 0.77	2.67 ± 0.83	**<0.001**	0.198

SD, standard deviation. SMD, standardized mean difference. Study participants included those (N = 706) who had complete anthropometric, genetic, food intake, and covariate data available, and were thus part of the analyses with genetic data. Chi-square test used for categorical variables, ANOVA test used for normally distributed continuous variables, and Wilcoxon ran sum test used for non-normally distributed continuous variables. Significance threshold (**bold**) set at 0.05.

Supplementary Table 6-2. Comparisons of characteristics of study participants with macronutrient and energy intake and genetic data and excluded QLSCD participants.

Characteristic	*Study sample N=646*	*Excluded QLSCD N=1474*	*p-value*	SMD
Categorical variables, % (n)				
Sex			**<0.001**	0.170
Female	55.0 (355)	46.5 (685)		
Male	45.0 (291)	53.5 (789)		
Preterm birth			1.00	0.002
Yes	4.6 (30)	4.7 (69)		
No	95.4 (616)	95.3 (1405)		
Maternal age at birth			0.122	0.102
≤ 20 years	1.7 (11)	3.3 (48)		
20-34 years	84.2 (544)	83.6 (1231)		
≥ 35 years	14.1 (91)	13.2 (194)		
Maternal education			**0.001**	0.166
Secondary education or less	42.2 (250)	46.9 (693)		
Post-sec. education	61.3 (396)	53.1 (781)		
Household income			**0.001**	0.197
< 30 000$	24.3 (157)	32.4 (465)		
30 000 – <60 000$	41.8 (270)	39.9 (573)		
60 000 – <80 000$	17.0 (110)	15.1 (217)		
≥ 80 000$	16.9 (109)	12.6 (181)		
Born in Canada	98.0 (633)	83.7 (1234)	**<0.001**	0.511
Continuous variables, mean ± SD				
Birth weight, kg	3.43 ± 0.47	3.39 ± 0.51	0.096	0.080
Proteins, 100g per day	0.60 ± 0.23	0.60 ± 0.23	0.603	0.027
Lipids, 100g per day	0.58 ± 0.23	0.57 ± 0.22	0.470	0.037
Carbohydrates, 100g per day	2.20 ± 0.68	2.18 ± 0.67	0.560	0.030
Total energy, 1000 calories per day	1.61 ± 0.46	1.60 ± 0.46	0.634	0.025

SD, standard deviation. SMD, standardized mean difference. Study participants included those (N = 646) who had complete anthropometric, genetic, macronutrient and energy intake, and covariate data available, and were thus part of the analyses with genetic data. Chi-square test used for categorical variables, and ANOVA test used for normally distributed continuous variables. Significance threshold (**bold**) set at 0.05.

Supplementary Table 6-3. Age-specific association between juice and fruit drinks intake and BMI at 4, 8, and 13 years, and association with change in BMI with age, by sex.

Model	(A) BMI 4 years		(B) BMI 8 years		(C) BMI 13 years		(D) BMI slope	
	β (95% CI)	p-value	β (95% CI)	p-value	β (95% CI)	p-value	β (95% CI)	p-value
Girls								
Juice and fruit drinks	-0.06 (-0.30, 0.18)	0.632	0.25 (-0.08, 0.59)	0.142	0.68 (0.10, 1.27)	**0.022**	0.08 (0.02, 0.14)	**0.008**
Boys								
Juice and fruit drinks	-0.01 (-0.25, 0.24)	0.950	0.01 (-0.34, 0.35)	0.976	0.04 (-0.59, 0.66)	0.908	0.01 (-0.06, 0.07)	0.882

BMI, body-mass index. CI, confidence interval. Estimated association between juice and fruit drink intakes and (A) BMI at 4 years, (B) BMI at 8 years, (C) BMI at 13 years, and (D) BMI slope (intake x age interaction) with 95% CI using linear mixed models. All coefficients estimated based on the model described in equation B, which includes the following covariates: sex, birthweight, maternal education, sufficient income, and whether the mother was born in Canada. Significance threshold (**bold**) set at 0.05.

Supplementary Table 6-4. Age-specific association between macronutrient intake and BMI at 4, 8, and 13 years, and association with change in BMI with age, adjusted for total energy intake.

Model	(A) BMI 4 years		(B) BMI 8 years		(C) BMI 13 years		(D) BMI slope	
	β (95% CI)	p-value	β (95% CI)	p-value	β (95% CI)	p-value	β (95% CI)	p-value
Macronutrient intake								
Proteins	0.17 (-0.35, 0.70)	0.514	0.35 (-0.31, 1.00)	0.301	0.55 (-0.49, 1.59)	0.297	0.04 (-0.06, 0.14)	0.413
Lipids	-0.42 (-1.12, 0.29)	0.247	0.02 (-0.80, 0.84)	0.959	0.61 (-0.55, 1.77)	0.303	0.11 (0.01, 0.22)	**0.031**
Carbohydrates	0.03 (-0.22, 0.28)	0.831	0.20 (-0.08, 0.49)	0.166	0.44 (0.04, 0.83)	**0.029**	0.05 (0.01, 0.08)	**0.009**

BMI, body-mass index. CI, confidence interval. Estimated association between macronutrient intakes and (A) BMI at 4 years, (B) BMI at 8 years, (C) BMI at 13 years, and (D) BMI slope (intake x age interaction) with 95% CI using linear mixed models. All coefficients estimated based on the model described in equation B, which includes the following covariates: sex, birthweight, maternal education, sufficient income, whether the mother was born in Canada, and total energy intake. Significance threshold (**bold**) set at 0.05.

Supplementary Table 6-5. Age-specific association of juice and fruit drinks intake with BMI by PGS value, by sex.

Model	PGS: mean score -1 SD		PGS: mean score		PGS: mean score +1 SD		Drink x PGS	
	β (95% CI)	p-value	β (95% CI)	p-value	β (95% CI)	p-value	β (95% CI)	p-value
Girls							-0.35 (-0.69, -0.01)	**0.044**
4 years	0.19 (-0.28, 0.66)	0.418	-0.15 (-0.49, 0.19)	0.381	-0.50 (-1.00, -0.00)	**0.049**		
8 years	0.64 (0.09, 1.19)	**0.023**	0.29 (-0.15, 0.74)	0.197	-0.05 (-0.63, 0.52)	0.862		
13 years	1.20 (0.40, 2.01)	**0.003**	0.85 (0.11, 1.59)	**0.025**	0.49 (-0.33, 1.32)	0.242		
Boys							0.14 (-0.22, 0.50)	0.455
4 years	-0.01 (-0.53, 0.50)	0.957	0.12 (-0.24, 0.49)	0.503	0.26 (-0.25, 0.77)	0.314		
8 years	-0.08 (-0.69, 0.54)	0.805	0.06 (-0.43, 0.55)	0.803	0.20 (-0.41, 0.81)	0.517		
13 years	-0.13 (-1.13, 0.87)	0.797	0.01 (-0.93, 0.94)	0.989	0.14 (-0.86, 1.14)	0.777		

PGS, polygenic score. CI, confidence interval. Estimated association of juice and fruit drinks intake with BMI according to age and the adult-derived PGS with 95% CI using linear mixed models. All coefficients estimated based on the model described in equation C, which includes the following covariates: sex, birthweight, maternal education, sufficient income, and whether the mother was born in Canada. Significance threshold (**bold**) set at 0.05.

Supplementary Table 6-6. Age-specific association of protein and lipid intake with BMI by PGS value, adjusted for total energy intake.

Model	PGS: mean score -1 SD		PGS: mean score		PGS: mean score +1 SD	
	β (95% CI)	p-value	β (95% CI)	p-value	β (95% CI)	p-value
Nutrient x PGS						
Proteins	0.56 (0.05, 1.07)	**0.032**	-	-	-	-
Lipids	0.62 (0.12, 1.12)	**0.016**	-	-	-	-
4 years						
Proteins	-0.26 (-1.16, 0.65)	0.579	0.30 (-0.46, 1.05)	0.441	0.85 (-0.06, 1.76)	0.068
Lipids	-1.06 (-2.18, 0.06)	0.063	-0.44 (-1.45, 0.56)	0.385	0.17 (-0.96, 1.30)	0.764
8 years						
Proteins	-0.02 (-1.05, 1.00)	0.965	0.52 (-0.37, 1.41)	0.254	1.06 (0.03, 2.09)	**0.043**
Lipids	-0.70 (-1.92, 0.52)	0.263	-0.10 (-1.21, 1.02)	0.867	0.51 (-0.72, 1.74)	0.420
13 years						
Proteins	0.33 (-1.15, 1.81)	0.663	0.89 (-0.50, 2.28)	0.208	1.45 (-0.03, 2.93)	0.054
Lipids	-0.25 (-1.89, 1.38)	0.762	0.37 (-1.19, 1.93)	0.643	0.99 (-0.65, 2.63)	0.237

PGS, polygenic score. CI, confidence interval. Estimated association of protein and lipid intakes with BMI according to age and the adult-derived PGS with 95% CI using linear mixed models. All coefficients estimated based on the model described in equation C, which includes the following covariates: sex, birthweight, maternal education, sufficient income, whether the mother was born in Canada, and total energy intake. Significance threshold (**bold**) set at 0.05.

Supplementary Table 6-7. Age-specific association between food, macronutrient and energy intake and BMI at 4, 8, and 13 years, and association with change in BMI with age, with IPW.

Model	(A) BMI 4 years		(B) BMI 8 years		(C) BMI 13 years		(D) BMI slope	
	β (95% CI)	p-value	β (95% CI)	p-value	β (95% CI)	p-value	β (95% CI)	p-value
Food intake								
Juice and fruit drinks	-0.03 (-0.21, 0.14)	0.725	0.12 (-0.12, 0.36)	0.314	0.34 (-0.08, 0.77)	0.115	0.04 (-0.00, 0.09)	0.077
Sweets and snacks	-0.36 (-0.61, -0.11)	**0.005**	-0.33 (-0.67, 0.02)	0.062	-0.28 (-0.89, 0.34)	0.375	0.01 (-0.06, 0.08)	0.776
Meats	-0.07 (-0.36, 0.21)	0.601	0.09 (-0.29, 0.48)	0.638	0.21 (-0.49, 0.90)	0.561	0.03 (-0.04, 0.11)	0.397
Fruits and vegetables	-0.06 (-0.17, 0.05)	0.323	-0.18 (-0.33, -0.04)	**0.015**	-0.36 (-0.62, -0.11)	**0.006**	-0.03 (-0.06, -0.01)	**0.015**
Macrconutrient intake								
Proteins	0.78 (0.40, 1.17)	**<0.001**	0.97 (0.43, 1.52)	**<0.001**	1.20 (0.22, 2.18)	**0.016**	0.05 (-0.06, 0.15)	0.375
Lipids	0.76 (0.36, 1.16)	**<0.001**	1.24 (0.67, 1.80)	**<0.001**	1.78 (0.78, 2.79)	**0.001**	0.11 (0.01, 0.22)	**0.034**
Carbohydrates	0.31 (0.17, 0.44)	**<0.001**	0.49 (0.30, 0.67)	**<0.001**	0.71 (0.38, 1.04)	**<0.001**	0.05 (0.01, 0.08)	**0.011**
Total energy	0.51 (0.31, 0.70)	**<0.001**	0.79 (0.52, 1.06)	**<0.001**	1.12 (0.64, 1.61)	**<0.001**	0.07 (0.02, 0.12)	**0.008**

BMI, body-mass index. CI, confidence interval. Estimated association between food, macronutrient and energy intakes and (A) BMI at 4 years, (B) BMI at 8 years, (C) BMI at 13 years, and (D) BMI slope (intake x age interaction) with 95% CI using linear mixed models. All coefficients estimated based on the model described in equation B, which includes the following covariates: sex, birthweight, maternal education, sufficient income, and whether the mother was born in Canada. Significance threshold (**bold**) set at 0.05.

Supplementary Table 6-8. Coefficient for the interaction of food, macronutrient and energy intake with PGSs, with IPW.

Model	Child-derived PGS		Adult-derived PGS	
	β (95% CI)	p-value	β (95% CI)	p-value
Juice and fruit drinks	0.08 (-0.14, 0.31)	0.466	-0.13 (-0.38, 0.12)	0.301
Sweets and snacks	0.05 (-0.29, 0.38)	0.788	-0.13 (-0.49, 0.23)	0.478
Meats	0.03 (-0.36, 0.43)	0.867	0.46 (0.08, 0.85)	**0.018**
Fruits and vegetables	-0.03 (-0.18, 0.13)	0.719	0.07 (-0.08, 0.23)	0.355
Proteins	0.03 (-0.52, 0.59)	0.902	0.59 (0.07, 1.11)	**0.025**
Lipids	0.57 (-0.04, 1.18)	0.066	0.63 (0.12, 1.14)	**0.016**
Carbohydrates	0.13 (-0.06, 0.32)	0.174	0.17 (-0.03, 0.36)	0.101
Total energy	0.20 (-0.08, 0.49)	0.155	0.32 (0.06, 0.59)	**0.015**

PGS, polygenic score. CI, confidence interval. Estimated interaction between food, macronutrient and energy intakes and the child-derived and adult-derived PGS in relation to BMI level with 95% CI using linear mixed models. All coefficients estimated based on the model described in equation C, which includes the following covariates: sex, birthweight, maternal education, sufficient income, and whether the mother was born in Canada. Significance threshold (**bold**) set at 0.05.

Supplementary Table 6-9. Age-specific association of food, macronutrient and energy intake with BMI by PGS value, with IPW.

Model	PGS: mean score -1 SD		PGS: mean score		PGS: mean score +1 SD	
	β (95% CI)	p-value	β (95% CI)	p-value	β (95% CI)	p-value
4 years						
Proteins	0.43 (-0.31, 1.18)	0.255	1.02 (0.47, 1.58)	**<0.001**	1.61 (0.85, 2.38)	**<0.001**
Lipids	0.28 (-0.47, 1.03)	0.463	0.91 (0.34, 1.47)	**0.002**	1.53 (0.76, 2.30)	**<0.001**
Meats	-0.48 (-1.06, 0.10)	0.107	-0.03 (-0.46, 0.40)	0.890	0.42 (-0.16, 0.99)	0.157
Total energy	0.26 (-0.12, 0.64)	0.179	0.58 (0.30, 0.86)	**<0.001**	0.90 (0.51, 1.29)	**<0.001**
8 years						
Proteins	0.66 (-0.22, 1.55)	0.143	1.25 (0.53, 1.98)	**0.001**	1.85 (0.95, 2.75)	**<0.001**
Lipids	0.65 (-0.25, 1.55)	0.155	1.25 (0.51, 2.00)	**0.001**	1.86 (0.95, 2.77)	**<0.001**
Meats	-0.26 (-0.93, 0.42)	0.456	0.15 (-0.40, 0.70)	0.585	0.56 (-0.12, 1.24)	0.104
Total energy	0.65 (0.21, 1.09)	**0.004**	0.96 (0.60, 1.32)	**<0.001**	1.28 (0.83, 1.73)	**<0.001**
13 years						
Proteins	0.97 (-0.42, 2.36)	0.171	1.57 (0.27, 2.86)	**0.018**	2.16 (0.77, 3.56)	**0.002**
Lipids	1.06 (-0.35, 2.48)	0.141	1.69 (0.37, 3.01)	**0.012**	2.32 (0.90, 3.74)	**0.001**
Meats	-0.11 (-1.15, 0.94)	0.843	0.36 (-0.62, 1.33)	0.475	0.82 (-0.24, 1.88)	0.129
Total energy	1.11 (0.42, 1.80)	**0.002**	1.43 (0.79, 2.08)	**<0.001**	1.76 (1.07, 2.46)	**<0.001**

PGS, polygenic score. CI, confidence interval. Estimated association of food, macronutrient and energy intakes with BMI according to age and the adult-derived PGS with 95% CI using linear mixed models. All coefficients estimated based on the model described in equation C, which includes the following covariates: sex, birthweight, maternal education, sufficient income, and whether the mother was born in Canada. Significance threshold (**bold**) set at 0.05.

References

1. Rao, D.P., et al., *Childhood overweight and obesity trends in Canada.* Health Promot Chronic Dis Prev Can, 2016. **36**(9): p. 194-8 DOI: 10.24095/hpcdp.36.9.03.
2. Bancej, C., et al., *Evidence Brief--Trends and projections of obesity among Canadians.* Health Promot Chronic Dis Prev Can, 2015. **35**(7): p. 109-12 DOI: 10.24095/hpcdp.35.7.02.
3. Kansra, A.R., S. Lakkunarajah, and M.S. Jay, *Childhood and Adolescent Obesity: A Review.* Frontiers in pediatrics, 2021. **8**: p. 581461-581461 DOI: 10.3389/fped.2020.581461.
4. Simmonds, M., et al., *Predicting adult obesity from childhood obesity: a systematic review and meta-analysis.* Obesity Reviews, 2016. **17**(2): p. 95-107 DOI: https://doi.org/10.1111/obr.12334.
5. Reilly, J.J. and J. Kelly, *Long-term impact of overweight and obesity in childhood and adolescence on morbidity and premature mortality in adulthood: systematic review.* International Journal of Obesity, 2011. **35**(7): p. 891-898 DOI: 10.1038/ijo.2010.222.
6. Weihrauch-Blüher, S., P. Schwarz, and J.-H. Klusmann, *Childhood obesity: increased risk for cardiometabolic disease and cancer in adulthood.* Metabolism, 2019. **92**: p. 147-152 DOI: https://doi.org/10.1016/j.metabol.2018.12.001.
7. Styne, D.M., et al., *Pediatric Obesity-Assessment, Treatment, and Prevention: An Endocrine Society Clinical Practice Guideline.* J Clin Endocrinol Metab, 2017. **102**(3): p. 709-757 DOI: 10.1210/jc.2016-2573.
8. Ambrosini, G.L., *Childhood dietary patterns and later obesity: a review of the evidence.* Proc Nutr Soc, 2014. **73**(1): p. 137-46 DOI: 10.1017/s0029665113003765.
9. Nguyen, M., et al., *Sugar-sweetened beverage consumption and weight gain in children and adults: a systematic review and meta-analysis of prospective cohort studies and randomized controlled trials.* Am J Clin Nutr, 2023. **117**(1): p. 160-174 DOI: 10.1016/j.ajcnut.2022.11.008.
10. Hu, F.B., *Resolved: there is sufficient scientific evidence that decreasing sugar-sweetened beverage consumption will reduce the prevalence of obesity and obesity-related diseases.* Obes Rev, 2013. **14**(8): p. 606-19 DOI: 10.1111/obr.12040.
11. Poorolajal, J., et al., *Behavioral factors influencing childhood obesity: a systematic review and meta-analysis.* Obesity Research & Clinical Practice, 2020. **14**(2): p. 109-118 DOI: https://doi.org/10.1016/j.orcp.2020.03.002.

12. Saeed, S., A. Bonnefond, and P. Froguel, *Obesity: exploring its connection to brain function through genetic and genomic perspectives.* Molecular Psychiatry, 2024 DOI: 10.1038/s41380-024-02737-9.
13. Ndiaye, F.K., et al., *The expression of genes in top obesity-associated loci is enriched in insula and substantia nigra brain regions involved in addiction and reward.* International Journal of Obesity, 2020. **44**(2): p. 539-543 DOI: 10.1038/s41366-019-0428-7.
14. Gagnon, E., et al., *Genetic control of body weight by the human brain proteome.* iScience, 2023. **26**(4) DOI: 10.1016/j.isci.2023.106376.
15. Silventoinen, K., et al., *Genetic and environmental effects on body mass index from infancy to the onset of adulthood: an individual-based pooled analysis of 45 twin cohorts participating in the COllaborative project of Development of Anthropometrical measures in Twins (CODATwins) study.* The American journal of clinical nutrition, 2016. **104**(2): p. 371-379 DOI: 10.3945/ajcn.116.130252.
16. Bouchard, C., *The study of human variability became a passion.* European Journal of Clinical Nutrition, 2022. **76**(5): p. 631-636 DOI: 10.1038/s41430-021-00871-z.
17. Heianza, Y. and L. Qi, *Gene-Diet Interaction and Precision Nutrition in Obesity.* International journal of molecular sciences, 2017. **18**(4): p. 787 DOI: 10.3390/ijms18040787.
18. Han, H.Y., et al., *Interactions between Polygenic Risk of Obesity and Dietary Factors on Anthropometric Outcomes: A Systematic Review and Meta-Analysis of Observational Studies.* The Journal of Nutrition, 2024 DOI: https://doi.org/10.1016/j.tjnut.2024.10.014.
19. Brunkwall, L., et al., *Sugar-sweetened beverage consumption and genetic predisposition to obesity in 2 Swedish cohorts.* The American journal of clinical nutrition, 2016. **104**(3): p. 809-815 DOI: 10.3945/ajcn.115.126052.
20. Qi, Q., et al., *Sugar-Sweetened Beverages and Genetic Risk of Obesity.* 2012. **367**(15): p. 1387-1396 DOI: 10.1056/NEJMoa1203039.
21. Ding, M., et al., *Diet quality and genetic association with body mass index: results from 3 observational studies.* The American journal of clinical nutrition, 2018. **108**(6): p. 1291-1300 DOI: 10.1093/ajcn/nqy203.
22. Wang, T., et al., *Improving adherence to healthy dietary patterns, genetic risk, and long term weight gain: gene-diet interaction analysis in two prospective cohort studies.* BMJ (Clinical research ed.), 2018. **360**: p. j5644-j5644 DOI: 10.1136/bmj.j5644.

23. Wang, T., et al., *Improving fruit and vegetable intake attenuates the genetic association with long-term weight gain.* Am J Clin Nutr, 2019. **110**(3): p. 759-768 DOI: 10.1093/ajcn/nqz136.
24. Tyrrell, J., et al., *Gene–obesogenic environment interactions in the UK Biobank study.* International Journal of Epidemiology, 2017. **46**(2): p. 559-575 DOI: 10.1093/ije/dyw337 %J International Journal of Epidemiology.
25. Hüls, A., et al., *Polygenic risk for obesity and its interaction with lifestyle and sociodemographic factors in European children and adolescents.* International Journal of Obesity, 2021. **45**(6): p. 1321-1330 DOI: 10.1038/s41366-021-00795-5.
26. Viljakainen, H., et al., *Interaction between genetic susceptibility to obesity and food intake on BMI in Finnish school-aged children.* Scientific Reports, 2023. **13**(1): p. 15265 DOI: 10.1038/s41598-023-42430-5.
27. Goni, L., et al., *A genetic risk tool for obesity predisposition assessment and personalized nutrition implementation based on macronutrient intake.* Genes Nutr, 2015. **10**(1): p. 445 DOI: 10.1007/s12263-014-0445-z.
28. Casas-Agustench, P., et al., *Saturated fat intake modulates the association between an obesity genetic risk score and body mass index in two US populations.* Journal of the Academy of Nutrition and Dietetics, 2014. **114**(12): p. 1954-1966 DOI: 10.1016/j.jand.2014.03.014.
29. Lee, W.J., et al., *Analysis of the Interaction between Polygenic Risk Score and Calorie Intake in Obesity in the Korean Population.* Lifestyle Genomics, 2021. **14**(1): p. 20-29 DOI: 10.1159/000511333.
30. Orri, M., et al., *Cohort Profile: Quebec Longitudinal Study of Child Development (QLSCD).* Social Psychiatry and Psychiatric Epidemiology, 2020 DOI: 10.1007/s00127-020-01972-z.
31. Desrosiers, H., *Enquête de nutrition auprès des enfants québécois de 4 ans*. 2008: < bound method Organization. get_name_with_acronym of< Organization ….
32. Dubois, L., et al., *Demographic and socio-economic factors related to food intake and adherence to nutritional recommendations in a cohort of pre-school children.* Public Health Nutrition, 2011. **14**(6): p. 1096-1104 DOI: 10.1017/S1368980010003769.
33. Perret, L.C., et al., *Polygenic risk score and peer victimisation independently predict depressive symptoms in adolescence: results from the Quebec Longitudinal Study of Children Development.* Journal of Child Psychology and Psychiatry, 2022. **n/a**(n/a) DOI: https://doi.org/10.1111/jcpp.13706.

34. Goulet, D., et al., *Mediation of genetic susceptibility to obesity through eating behaviours in children.* Pediatric Obesity, 2024. **n/a**(n/a): p. e13180 DOI: https://doi.org/10.1111/ijpo.13180.
35. Yengo, L., et al., *Meta-analysis of genome-wide association studies for height and body mass index in ∼700000 individuals of European ancestry.* Human molecular genetics, 2018. **27**(20): p. 3641-3649 DOI: 10.1093/hmg/ddy271.
36. Vogelezang, S., et al., *Novel loci for childhood body mass index and shared heritability with adult cardiometabolic traits.* PLoS genetics, 2020. **16**(10): p. e1008718-e1008718 DOI: 10.1371/journal.pgen.1008718.
37. Ge, T., et al., *Polygenic prediction via Bayesian regression and continuous shrinkage priors.* Nature communications, 2019. **10**(1): p. 1776-1776 DOI: 10.1038/s41467-019-09718-5.
38. Cefalu, M., et al., *twang: Toolkit for Weighting and Analysis of Nonequivalent Groups*. 2021, https://CRAN.R-project.org/package=twang.
39. Libuda, L., et al., *Pattern of beverage consumption and long-term association with body-weight status in German adolescents--results from the DONALD study.* Br J Nutr, 2008. **99**(6): p. 1370-9 DOI: 10.1017/s0007114507862362.
40. Nissinen, K., et al., *Sweets and sugar-sweetened soft drink intake in childhood in relation to adult BMI and overweight. The Cardiovascular Risk in Young Finns Study.* Public Health Nutr, 2009. **12**(11): p. 2018-26 DOI: 10.1017/s1368980009005849.
41. Ledoux, T.A., M.D. Hingle, and T. Baranowski, *Relationship of fruit and vegetable intake with adiposity: a systematic review.* Obesity Reviews, 2011. **12**(5): p. e143-e150 DOI: https://doi.org/10.1111/j.1467-789X.2010.00786.x.
42. Martinez, J.A., et al., *Personalized weight loss strategies—the role of macronutrient distribution.* Nature Reviews Endocrinology, 2014. **10**(12): p. 749-760 DOI: 10.1038/nrendo.2014.175.
43. Willems, A.E.M., et al., *Effects of macronutrient intake in obesity: a meta-analysis of low-carbohydrate and low-fat diets on markers of the metabolic syndrome.* Nutr Rev, 2021. **79**(4): p. 429-444 DOI: 10.1093/nutrit/nuaa044.
44. Sartorius, K., et al., *Does high-carbohydrate intake lead to increased risk of obesity? A systematic review and meta-analysis.* BMJ Open, 2018. **8**(2): p. e018449 DOI: 10.1136/bmjopen-2017-018449.
45. Jakobsen, D.D., L. Brader, and J.M. Bruun, *Effects of foods, beverages and macronutrients on BMI z-score and body composition in children and adolescents: a systematic review and meta-*

analysis of randomized controlled trials. European Journal of Nutrition, 2023. **62**(1): p. 1-15 DOI: 10.1007/s00394-022-02966-0.

46. Stokes, A., et al., *Protein Intake from Birth to 2 Years and Obesity Outcomes in Later Childhood and Adolescence: A Systematic Review of Prospective Cohort Studies.* Advances in Nutrition, 2021. **12**(5): p. 1863-1876 DOI: https://doi.org/10.1093/advances/nmab034.
47. Arnesen, E.K., et al., *Protein intake in children and growth and risk of overweight or obesity: A systematic review and meta-analysis.* Food Nutr Res, 2022. **66** DOI: 10.29219/fnr.v66.8242.
48. Pérusse, L., et al., *Understanding Gene-Lifestyle Interaction in Obesity: The Role of Mediation versus Moderation.* Lifestyle Genomics, 2022. **15**(2): p. 67-76 DOI: 10.1159/000523813.
49. Olsen, N.J., et al., *Interactions between genetic variants associated with adiposity traits and soft drinks in relation to longitudinal changes in body weight and waist circumference.* The American Journal of Clinical Nutrition, 2016. **104**(3): p. 816-826 DOI: 10.3945/ajcn.115.122820.
50. Heianza, Y., et al., *Healthful plant-based dietary patterns, genetic risk of obesity, and cardiovascular risk in the UK biobank study.* Clin Nutr, 2021. **40**(7): p. 4694-4701 DOI: 10.1016/j.clnu.2021.06.018.
51. Helgeland, Ø., et al., *Characterization of the genetic architecture of infant and early childhood body mass index.* Nature Metabolism, 2022 DOI: 10.1038/s42255-022-00549-1.
52. Smith, J.D., E. Fu, and M.A. Kobayashi, *Prevention and Management of Childhood Obesity and Its Psychological and Health Comorbidities.* Annu Rev Clin Psychol, 2020. **16**: p. 351-378 DOI: 10.1146/annurev-clinpsy-100219-060201.
53. Loos, R.J.F. and G.S.H. Yeo, *The genetics of obesity: from discovery to biology.* Nature Reviews Genetics, 2022. **23**(2): p. 120-133 DOI: 10.1038/s41576-021-00414-z.
54. Brassard, D., et al., *Differences in Population-Based Dietary Intake Estimates Obtained From an Interviewer-Administered and a Self-Administered Web-Based 24-h Recall.* Front Nutr, 2020. **7**: p. 137 DOI: 10.3389/fnut.2020.00137.

CHAPTER 7 - DISCUSSION

7.1 Summary of main findings

Using data from a birth cohort study that followed children born in 1997-1998, this thesis leveraged PGS to study the interplay between dietary habits and genetic susceptibility in relation to body weight. We first constructed two PGS derived from adult and children GWAS summary statistics to compare their association with BMI z-score and risk of obesity. Then, we used those two PGS to examine if eating behaviours (fussy eating and over-eating) assessed in early childhood mediate the association between adult-and child-derived PGS and BMI z-score. Finally, we estimated the association between preschool dietary (food, macronutrient, and energy) intakes and childhood BMI (mean and change with age) and examined if the adult- and child-derived PGS can modify those associations.

7.1.1 Manuscript 1

The findings from the first manuscript (see Chapter 4) suggest that both PGS are good indicators of genetic susceptibility to obesity. However, the adult-derived PGS had a stronger predictive ability across levels of analyses, compared to the child-derived PGS. Although both scores identified marked differences in BMI z-scores and proportion of obesity between children in the first PGS quintile compared to those in the fifth quintile, the adult-derived score identified a higher gradient, especially starting at 8 years of age. For instance, the difference in the proportion of obesity (1^{st} vs. 5^{th} quintile) increased drastically from 4 to 6 years (6.8%-9.7%) to 8-13 years (15.4%-18.2%) for the adult-derived PGS, whereas the difference remained more stable from 4 to 13 years for the child-derived PGS (9.2%-13.6%). The adult-derived PGS also consistently captured a higher proportion of variance of BMI z-score from 4 to 13 years (3.9% to 10.4%), compared to its child-derived counterpart (3.5% to 6.9%). Nonetheless, the child-derived PGS had a stronger association with the risk of obesity from 4 to 7 years (OR_{child} 1.75-2.33 vs. OR_{adult} 1.74-2.06), highlighting its utility in specific circumstances.

The results presented in this thesis are significant in 3 ways. First, the longitudinal design used in the study allowed to describe in detail through time the contribution of common variants to BMI variation and the risk of obesity in childhood. Notably, our results align with previous observations [16] suggesting that genetic variants influencing adult BMI already emerge in childhood. Second, from a methodological perspective, our results support the current use of adult-derived PGS to represent genetic susceptibility to higher BMI in childhood in most circumstances. Our results align with two previous investigations [128, 129] that showed that adult-derived PGS have better performance compared to child-derived PGS. An important caveat is the inequality in terms of GWAS sample size, which can produce naturally more accurate adult-derived PGS, compared to a child-derived PGS, which renders direct comparison more difficult. Third, although the child-derived PGS performed worse in most metrics, the score still yielded a higher proportion of variance explained compared to the few child-derived PGS available in the literature, which explained 3.6% of childhood BMI variation at most [21]. This is likely due to our use of a PGS construction method allowing for the inclusion of thousands of variants using statistical shrinkage, compared to other scores [21, 97, 128] that limited the inclusion to a subset of statistically significant variants.

7.1.2 Manuscript 2

In Chapter 5, employing linear regression, we observed that both PGS were significantly associated with fussy eating (β_{child}: -0.14, 95% CI: -0.24;-0.04, p=0.007; $\beta_{adult:}$ -0.11, 95% CI: -0.21;-0.01, p=0.030) and over-eating (β_{child}: 0.13, 95% CI: 0.07;0.20, p<0.001; β_{adult}: 0.15, 95% CI: 0.08;0.22, p<0.001). Using longitudinal growth curve mediation analysis, we confirmed that over-eating, but not fussy eating, mediated the association between both the adult-derived and the child-derived PGS on BMI z-scores from 6 to 13 years of age, with the proportion mediated decreasing as the children got older. Indeed, the proportion mediated by over-eating decreased from 18.0% to 11.4% for the child-derived PGS, and from 14.2% to 6.2% for the adult-derived PGS. Further analysis revealed that the decreasing level of mediation was due to the negative association of over-eating with the slope of BMI z-scores (i.e., the effect of over-eating on BMI z-scores decreased with time). The structural nature of the analysis also revealed a positive direct association between the adult-derived, but not the child-derived, PGS and the slope of BMI z-scores (β = 0.012 kg/m^2 per year and one SD of the PGS), explaining the higher rate of decrease in mediation for the adult-derived PGS. The absence of

mediation for fussy eating was explained by an association close to the null between fussy eating and BMI z-score (intercept and slope).

Our results add important evidence in support of the behavioural theory, where studies confirming the mediation of individual genetic susceptibility to obesity (measured using PGS) through appetite-related behaviours were few in children specifically. The results presented in this manuscript also complement those obtained in manuscript 1 in two ways. First, although the adult-derived PGS explain a higher proportion of BMI variance, the child-derived PGS may be better suited to capture appetite-related behaviours underpinning the genetic susceptibility to obesity in childhood. Second, the observed increase over time of the effect of the adult-derived PGS on BMI z-score supports the conclusion of the first manuscript on the emergence of adult BMI-related variants in childhood and adolescence.

7.1.3 Manuscript 3

Leveraging linear mixed models to consider the repeated measures of BMI, we showed in Chapter 6 that a lower consumption of fruits and vegetables, and a higher consumption of juice and fruit drinks (in girls), proteins, lipids, carbohydrates, and total energy at 4 years is associated with higher BMI from 4 to 13 years. We also found that the associations with BMI increased with age (kg/m^2 per year) for fruits and vegetables (β: -0.03, 95%CI: -0.06;-0.01), lipids (β: 0.11, 95%CI: 0.01;0.22), carbohydrates (β: 0.05, 95%CI: 0.01;0.08), and total energy (β: 0.07, 95%CI: 0.02;0.12). This is consistent with the general energy balance theory that posits that the accumulation of adiposity is related to increased food intake in response to changes in the food environment. More specifically, the energy balance theory highlights the predominant role of the central nervous system in aggregating both internal (e.g., hormones or signals of energy status) and external (e.g., availability and marketing of energy-dense foods) stimuli to regulate food intake and body weight [206]. Contrary to some theoretical models arguing that diets rich in carbohydrates are the primary driver of the increase in the prevalence of obesity [207], we observed that the associations of all three macronutrients with BMI were mitigated when adjusting for total energy intake. This evidence supports previous findings that the primary metabolic driver of obesity is an excess of calorie intake due to several dietary factors beyond diet composition [206].

Similarly, we also observed that the associations between proteins (β: 0.54, 95%CI: 0.03;1.06), lipids (β: 0.63, 95%CI: 0.12;1.13), and total energy ((β: 0.32, 95%CI: 0.06;0.58) intake and BMI were stronger in children with higher values of the adult-derived PGS. Combined with the results from the second manuscript discussed earlier about the mediating role of over-eating in genetic susceptibility to obesity, this also aligns with, and presents evidence for, the behavioural susceptibility theory. Individuals at high genetic risk of obesity would be more vulnerable to the effects of increased food intake, especially since a large part of the genetic susceptibility is translated through appetite regulation. Showcasing that individuals at higher genetic risk are more susceptible to the effects of the diet on BMI is also noteworthy, as this emphasizes the need for obesity prevention efforts creating a healthier food environment that could mitigate the inherent differential genetic susceptibility in children.

7.2 Contributions to public health and clinical care

The results obtained from this thesis benefit public health efforts and clinical care for obesity in three ways. First, although it is not possible to modify an individual's genotype, mediation and interaction analyses can help identify modifiable factors translating or attenuating the effects of genetic susceptibility on body weight. Second, observational evidence gathered from this thesis can help refine theoretical models crucial to design more effective childhood obesity prevention strategies. Third, our results added information regarding the use of PGS as an emerging tool for stratifying individuals based on their obesity risk.

7.2.1 Identify targets for obesity prevention and management

The identification of optimal targets for obesity management and prevention is an incentive for both the mediation and interaction analyses conducted throughout this thesis. Indeed, it may not always be possible, or optimal, to act on important disease risk factors. Mediation analyses can help pinpoint targets for interventions that play an intermediary role in the development of a disease [23]. Similarly,

although it is not feasible to intervene directly on some exposures, it may be possible to act on their moderators [23]. Our results from the first manuscript confirmed that individual differences in genetic susceptibility are a major determinant of variability in BMI starting early in life. Considering it is not possible to directly change an individual's genotype, an alternative solution is to target interventions on behaviours translating genetic susceptibility, or on factors accentuating or attenuating the effect of genetic factors on actual phenotypic variance. However, there is a concerning lack of evidence in children, specifically regarding the interplay between BMI-PGS and dietary factors for those two aspects. We identified over-eating as an important intermediate in the association between PGS and BMI z-score as early as four years of age. Our results also suggest the presence of an interaction between specific dietary intakes and the adult-derived PGS in relation to childhood BMI. Although, our study showed the inverse relationship, the presence of an interaction also implies that a diet characterized by lower energy, lipid and protein intake may attenuate the effect of a PGS on BMI. Incidentally, the results gathered in the second and third manuscripts add crucial evidence in childhood, revealing that appetite-related behaviours and dietary intakes may be good targets for intervention and prevention.

Those two contributions to knowledge fall in line with the observation that most obesity prevention efforts and interventions aimed at children follow an individual-driven behavioural approach. However, this type of approach has only yielded minimal long-term efficacy in reducing the burden of obesity in children [208]. Rather, approaches with a focus on the family as a complete unit have found more success in the treatment of obesity in young children [37]. This is coherent with the biopsychosocial model of health first developed by George L. Engel in 1977 [209]. Applied to childhood obesity, the model emphasizes that children's diet and appetitive traits mediate the effect of biological and genetic factors on body weight, and that parental feeding styles and practices are crucial in shaping their children's behaviours [210]. The longitudinal design used throughout the three manuscripts allowed us to provide evidence that 1) the effect of a PGS derived from adult GWAS summary statistics on BMI increases through time, accompanied by a decrease in the relative proportion mediated through over-eating, and 2) the protective effect of the consumption of fruits and vegetables, and the nefarious effect of a higher caloric load on BMI increases through childhood. Practically speaking, this suggests that parental feeding styles may have the most impact on their child's eating behaviours in early childhood, highlighting the need to act on the whole family as early as possible. Furthermore, as children age, it may be necessary to act on more distal environments (e.g., food marketing) since they become progressively more independent in their food choices.

7.2.2 Reinforcing the conceptual framework of obesity

The results presented in this thesis can also help reinforce or refine conceptual frameworks for the development of obesity. In turn, public health and clinical care can benefit from this evidence, since accurate theoretical models of the development of obesity are essential in tailoring obesity prevention efforts and interventions to the mechanisms likely to have the most impact. While there is evidence that the increasingly unhealthy food environment is a primary driver responsible for the rise of obesity in the past decades [11], a complete theoretical model should integrate genetic factors and their interplay with environmental drivers of obesity in the pathogenesis of obesity. As a reminder, the behavioural susceptibility theory postulates that appetite-related behaviours translate a large part of the genetic component of obesity [131]. According to that theory, individual differences in body weight come from varying levels of genetic susceptibility to the exposure of an obesogenic environment.

The theory hinges on four central hypotheses that are expanded upon in a recent publication [18]. We were able to support two of them with evidence from our manuscripts. The first states that "appetite […] mediates genetic influence on weight in early life" (p.2) [18]. Mediation analyses can help better understand specific causal pathways [23]. For example, although there is evidence that body weight and eating behaviours have a small to moderate genetic correlation [18] suggesting they share a common genetic architecture, one could ask whether the genetic variants affect body weight and eating behaviours independently or if eating behaviours are in the pathway between the effect of genetic variants and variations in body weight. The central nervous system, and more specifically the leptin-melanocortin biological pathway, is an important facet of the regulation of appetite [15]. MC4R neurons within the hypothalamus can promote or discourage food intake based on downstream signals from the concentration leptin hormone, which circulates in function with fat mass and changes in energy intake [15]. Although there is convincing evidence that eating behaviours and BMI share a common genetic architecture [211], there is a lack of longitudinal studies examining if eating behaviours mediate the association between BMI-PGS and adiposity-related traits, especially in children. We were able to address those limitations and incidentally support the behavioural susceptibility theory by showing that the association between two PGS and BMI z-score from 6 to 13 years was mediated through over-eating expressed from 2 to 6 years of age. The confirmation of mediation for both the adult- and the child-derived summary statistics is important because the few

studies [158-160] addressing this question in childhood used PGS derived from either exclusively adult summary statistics, or a combination of adult and child summary statistics. Our results are also complementary with prior investigations [212, 213] showing that higher values of a BMI-PGS were associated with lower volume and thickness of brain regions related to eating behaviours and reward regulation.

A second hypothesis posits that there is "[…] gene-environment interaction in weight development in early life" (p.2) [18]. There is already convincing evidence from twin studies supporting the presence of an interaction between genetic and environmental factors in relation to body weight at the population level. Notably, these investigations showed that children in environments classified as more obesogenic (e.g., parents with lower educational attainment [145], or living in high obesity risk homes [22]) have higher genetic variation in BMI. At the individual level, environmental exposure could be reflected by individual differences in dietary intakes. We were able to observe that the effect of higher total energy, protein, and lipid intake on BMI is increased in children with a higher value of a BMI-PGS. Although similar results had previously been obtained in adults, there was a lack of such observations in samples composed of children. Although our results focused on proximal factors (food, macronutrient, and energy intakes), it is likely that individual differences in dietary intake stem from more distal environmental exposures. In fact, our results complement previous results of over 300,000 adults showcasing that the association between home proximity to fast-food chains was higher in individuals with a higher value of a BMI-PGS [214]. In general, our results highlight the interplay between genetic factors and dietary habits in the development of obesity and thus support measures taken to build a healthier food environment. Indeed, a crucial part of obesity prevention efforts is aimed at changing environmental factors driving the overall increase in obesity prevalence apparent in the past few decades. Policy interventions targeting more distal systemic drivers of obesity (e.g., government policies promoting overconsumption) and environmental drivers related to food supply promoting higher energy intake can be leveraged to offer a healthier environment [11].

According to a recent evaluation of the current food environment in Canada based on policies and actions taken by different levels of the government, the food environment does not incite the consumption of a healthy diet [215]. More specifically, the evaluation highlighted that although the restrictions applied to food marketing to children in Quebec were recognized as an international benchmark, this aspect of policy remains an area of concern for the rest of Canada [215]. This is concerning considering we have added rare evidence at the individual level showcasing the interplay

between genetic susceptibility and dietary habits already starts in childhood. Additionally, the report highlights that intake targets of specific foods and macronutrients, especially fruits and vegetables are fundamental aspects in providing a healthier food environment in Canada [215]. Our results support this conclusion, since we showed that a higher consumption of fruits and vegetables is associated with progressively lower BMI throughout childhood.

Overall, the theoretical frameworks for obesity supported from evidence gathered through the present thesis and the literature support that a multidimensional approach is necessary to address the epidemic of obesity. Indeed, public health measures aimed at limiting the burden of obesity must include 1) health promotion and individual interventions applied to the behavioural patterns (e.g., over-eating) responsible for translating genetic susceptibility to obesity, and 2) policy-led solutions to target systemic and environmental drivers of obesity that are essential in accentuating or attenuating the inherent differences in susceptibility between individuals [11].

7.2.2 Usefulness of PGS for clinical care

In terms of clinical utility, the aspiration is for PGS to be implemented when obesity is not yet present as a tool to stratify individuals based on their obesity risk [216]. Due to the relatively low cost of genotyping procedures for PGS [217], it may be possible to implement such a tool in the general clinical care practice to inform personalized prevention and follow-up plans. However, current PGS (including the adult-derived and child-derived BMI-PGS developed for this thesis) only have weak predictive ability [15], suggesting that they may not have a meaningful clinical role on their own. The adult-derived PGS used in this thesis revealed associations with BMI z-scores (3.9%-10.4% variance explained from 4 to 13 years of age) and risk of obesity (0.65-0.74 AUC). This PGS compared favourably with others constructed using pruning and threshold methods that account for 4.2% to 6.6% of BMI variation in children [123, 124, 127]. However, the association between a PGS and a phenotype or high relative risk of disease does not necessarily translate to high discriminatory ability of a PGS, which is a clinically more useful measure of prediction [218]. For instance, just over 25% of children in the top quintile of the adult-derived PGS were afflicted with obesity at 13 years of age, meaning that, although at the highest risk of obesity, the majority of children at the higher end of the PGS distribution would still not be categorized as such. Nonetheless, most individual risk factors are not meant to be

used alone, but rather in combination with assessment tools to offer meaningful predictions. For example, risk prediction models for cardiovascular disease with good prediction usually combine multiple behavioural, biological, familial, or environmental risk factors, which on their own would not be able to properly predict disease risk [101, 219]. There is also evidence that integrating PGS with other genetic and non-genetic determinants could enhance the screening of breast cancer [220].

Alternative clinical utilities can also be considered, including sharing genetic risk information to induce behavioural changes in patients, or helping the decision process of treatment options using genetic information. Although primary care providers may be receptive to the introduction of PGS in their regular practice, additional evidence for their clinical utility is needed [221], especially considering evidence suggests providing individuals with genetic risk information is only slightly effective [222] or not effective [107, 223] in provoking behaviour change, with concerns for demotivation and increased anxiety [223]. More specifically, providing genetic risk information to a sample of 190 parents of 3- to 7-year-old children does not improve parental feeding [224]. Another potential clinical use of PGS includes the help of treatment decision making by determining if individuals have different reactions to treatments depending on their genetic background. For example, a previous investigation suggests that individuals at higher genetic risk of cardiovascular disease may respond more favourably to treatment by statins to limit the risk of coronary events [101, 225]. Regarding obesity treatment more specifically, GLP-1 receptor agonists are pharmacological treatments for type-2 diabetes that can also have a substantive effect on weight loss. There is emerging evidence that individuals carrying specific alleles may benefit more from these treatments compared to non-carriers [226]. Albeit beyond the scope of this study, this evidence highlights the potential role of genetic testing for the treatment of obesity.

7.3 Thesis strengths and limitations

7.3.1 Strengths

The PGS developed for the purpose of the research undertaken in this thesis were computed using Bayesian shrinkage techniques, which allowed for the inclusion of hundreds of thousands of genetic variants across the genome accounting for LD. Those types of scores are recognized as being able to

explain a larger proportion of phenotypic variance compared to pruning and thresholding methods [119]. Obesity is a complex disease where a wide array of biological processes is implicated in its development. Therefore, using those types of PGS allowed to capture as much as those processes as possible. In addition, using PGS derived from GWAs summary statistics from both adults and children allowed us to document the differences in prediction between the scores. Despite the overlap in genetic variants predicting BMI in children and adults, the unique contributions of PGS can differ, as new genetic variants can emerge in early childhood [17]. Using two different PGS sets this research apart from other investigations, as most studies measuring genetic susceptibility to obesity in childhood using a PGS only incorporate a single tool.

The use of longitudinal designs also allowed to give additional context to our observations. Most previous studies investigated the mediating and moderating role of dietary habits in genetic susceptibility to obesity measured through a PGS using cross-sectional designs [25, 134]. This aspect is fundamental because genetic variation of BMI increases drastically, starting at around 4 years of age [13], suggesting that the underlying biological processes behind genetic susceptibility to obesity could change with time. Therefore, the repeated measures of BMI added robustness to our results by adding a time dimension to our observations.

7.3.2 Limitations

Residual confounding is one of the limits of observational studies regarding causal inference. Confounding occurs when a variable affects both the exposure and the outcome of the relationship of interest, without being in the causal pathway [227]. The associations between PGS and BMI assessed throughout the thesis are unlikely to be affected by such a problem considering genetic variation is established at conception and thus unlikely to be affected by other factors. Population stratification, the presence of different allelic frequencies between different genetic ancestries, is still established as a possible process that could produce confounding in genetic associations [228]. More specifically, the genetic diversity of populations in the province of Quebec decreases in West-East fashion due to founder effects [229]. Considering regional differences in the prevalence of childhood obesity [230], there is the possibility of population stratification to bias the genetic associations with BMI described in the thesis. Nonetheless, the risk of possible confounding due to population stratification was

mitigated in our studies by the presence of over 97% of children in analyses involving genetic associations having mothers born in Canada, and by adjusting both our PGS by the first 10 principal components of genetic ancestry. However, residual confounding is possible in the associations of eating behaviours or dietary intakes with BMI if there are unobserved confounders, if confounders were failed to be identified, or if some variables were not accurately measured. For example, physical activity is related to both eating behaviours and BMI in children [231].

There is a risk of selection bias when the process of selecting the study participants or their probability of remaining in the study can induce systematic changes in the relationship of interest compared to the alternative where all the study population could take part in the study [232]. More specifically, the genetic analyses presented in the studies from this thesis are susceptible to selection bias from two main sources: (1) the longitudinal designs used in our studies that can introduce attrition even when the exposure of interest is genetic susceptibility, and (2) the selection of participants that accepted DNA collection for genetic analyses [233]. Practically, selective participation in a study can lead to collider bias when the exposure and the outcome of interest are related, either directly or indirectly to the participation or continued participation in the study. For example, BMI-PGS can be associated with depressive symptoms [234], which are known to affect retention in longitudinal studies [235]. Considering individuals with higher BMI are also likelier to drop out over time [235], this could bias estimates of the association between BMI-PGS and BMI. Additionally, study samples from our analyses differed from the complete QLSCD sample in terms of baseline characteristics (e.g., household income, maternal education, or biological sex). The analyses using propensity weights to simulate a study sample with similar characteristics, as the complete QLSCD sample showed similar results compared to unadjusted analyses, indicating a low risk of selection bias. However, we cannot rule out that unobserved variables not included in the weighting approach may still be present.

To attenuate the risk of misclassification of BMI, we exclusively used anthropometric directly measured by trained technicians instead of parent- or self-reported measurements. Still, BMI as a measure of excess adiposity has some limitations, such as not differentiating between fat and lean mass, and not being informative of the location of fat mass, two factors that can influence the health impact of increased body weight. We chose BMI to represent adiposity because the most comprehensive GWAS summary statistics available for obesity-related measures are for BMI. Additionally, there is the possibility of misclassification due to the hardship of measuring dietary intakes and eating behaviours. Although both assessments were based on validated questionnaires, the

answers were reported by a parent, which introduces the possibility of recall bias. Crucially, parents of children of higher BMI are more likely to under-report the dietary intakes of their child, compared to parents of children of lower BMI [236]. This potential differential misclassification is likely to bias the associations between dietary intakes and BMI toward the null for most dietary factors.

It is also important to acknowledge the limits of using PGS to represent genetic susceptibility to obesity. While these scores are increasingly used in observational studies, PGS can only represent the genetic susceptibility conferred by common variants [219]. More specifically, the effect of rare variants on BMI, the non-additive effects of genes, epigenetic effects, and gene-environment interactions [76, 77], which are all important components of the effect of genes on individual variations of BMI, are not captured by GWAS (and PGS). Furthermore, current GWAS are closing on the theoretical upper bound of the variance in adult BMI that common variants can explain [216]. On the other hand, there is a wider gap to fill in the child BMI variance explained by common variants due to the lower sample size in GWAS for child BMI. Consequently, it is difficult to make equal comparisons between PGS derived from adult and child BMI GWAS summary statistics.

Although the distribution of the complete QLSCD sample was regionally representative in the province of Quebec, the study samples used in our analyses were generally of higher socioeconomic status. More specifically, the higher household income, higher maternal education, and higher proportion of girls in our study samples compared to the initial QLSCD sample, are due to attrition and the incorporation of genetic data, which was not available for all participants. Although adjusting the analyses for this aspect with propensity weighting, it remains unclear if the results are generalizable to a population of lower socioeconomic status. Additionally, the generalization of our results may be limited by the particular nature of the genetic structure of Quebec. The province is known for its founder population, consisting of ~8500 initial settlers originating from France from the foundation of Quebec City in 1608 to the British conquest in 1759 [237]. This contributed to genetic stratification correlated with geographical location, and rare mutations concentrated in specific regions [229]. Despite those specificities, the overall genetic diversity of this population is similar to that observed in other European populations [238]. It is established that PGS are hardly transferable between ancestries [88]. Incidentally, the results obtained from genetic associations in the province of Quebec should be generalizable to populations of European ancestry, but not populations of other ancestry.

7.4 Suggestions for future research

Genetic susceptibility to obesity implicates hundreds of variants with various biological underpinnings. We established that over-eating mediates, in part, the genetic susceptibility to higher BMI measured by PGS, and that the consumption of macronutrients and energy interact with genetic susceptibility to influence BMI. However, it is likely that both these processes are not independent of each other. Therefore, future research could be enhanced by considering both these aspects in conjunction, either by using statistical methods uniting mediation and interaction analyses [23], or by refining a PGS to include variants especially involved in brain regions involved in appetite regulation. Such approach used in one previous study found that the interaction between dietary habits and a PGS in relation to BMI was stronger when the variants included in the PGS were selected based on their role in the central nervous system, compared to an approach without discrimination in the function of variants [163].

Finally, we added evidence supporting that the association between higher energy and macronutrient intake and increased BMI in childhood is accentuated by a high genetic risk of obesity. Beyond the relevance for prevention measures aimed at encouraging healthier dietary habits at the individual level, those results suggest that wide-ranging policies targeting the food environment may be able to contribute to reducing the inherent inequalities in obesity risk conferred by varying levels of genetic susceptibility. However, since our research has been conducted using individual level dietary factors, the transferability to more distal food environment factors remains uncertain. Therefore, further research may be directed toward the interaction between genetic susceptibility to obesity and food environmental factors, such as exposure to food marketing, the accessibility to fast food restaurants, or the presence of policies lowering access to energy-dense foods (e.g., higher taxes on specific products).

7.5 Conclusions

Evidence from the comparison of BMI-PGS derived from child and adult GWAS summary statistics supports the current use of adult-derived PGS in target populations of children in the literature but suggests that child-derived PGS may be better suited in investigations relating to obesity risk in children younger than 8 years of age. Despite the observation that the prevalence of obesity in children plateaued in recent years, obesity remains a major public health concern affecting more than 25% of children in Canada [8, 239]. In addition to providing valuable methodological information, the research conducted throughout this thesis managed to add crucial evidence supporting the theoretical framework behind the genetic susceptibility to obesity, which is relevant to obesity prevention in childhood. Theoretical frameworks are essential to guide public health measures aimed at reducing the burden of obesity in children. We found that over-eating mediates the association between genetic risk measured using PGS and BMI, and that lipid, protein and energy intake interact with genetic susceptibility to affect BMI throughout childhood. Those results support the behavioural susceptibility theory that frames obesity as a complex genetic disease, where the effect of individual genes affects adiposity through appetite regulation [18]. In this framework individual variations in genetically defined susceptibility to higher BMI are influenced by the overall food environment. Therefore, our results support the underlying hypotheses behind the behavioural susceptibility theory and highlight the need to ameliorate the current food environment in Canada that does not facilitate the adoption of healthy dietary habits.

BIBLIOGRAPHY

1. World Health, O., *Obesity : Preventing and managing the global epidemic : Report of a WHO consultation.* WHO Tech Rep Ser, 2000. **894**: p. 1-53, https://cir.nii.ac.jp/crid/1571135650391747072.
2. Institute for Health Metrics and Evaluation (IHME), *GBD 2021 Cause and Risk Summary: High body mass index - Level 2 risk*. 2024, IHME, University of Washington: Seattle, USA, https://www.healthdata.org/research-analysis/diseases-injuries-risks/factsheets/2021-high-body-mass-index-level-2-risk.
3. Phelps, N.H., et al., *Worldwide trends in underweight and obesity from 1990 to 2022: a pooled analysis of 3663 population-representative studies with 222 million children, adolescents, and adults.* The Lancet, 2024. **403**(10431): p. 1027-1050 DOI: 10.1016/S0140-6736(23)02750-2.
4. N. C. D. Risk Factor Collaboration, *Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128·9 million children, adolescents, and adults.* Lancet (London, England), 2017. **390**(10113): p. 2627-2642 DOI: 10.1016/S0140-6736(17)32129-3.
5. Heindel, J.J. and L.N. Vandenberg, *Developmental origins of health and disease: a paradigm for understanding disease cause and prevention.* Current opinion in pediatrics, 2015. **27**(2): p. 248-253 DOI: 10.1097/MOP.0000000000000191.
6. Olds, T., et al., *Evidence that the prevalence of childhood overweight is plateauing: data from nine countries.* Int J Pediatr Obes, 2011. **6**(5-6): p. 342-60 DOI: 10.3109/17477166.2011.605895.
7. Rokholm, B., J.L. Baker, and T.I.A. Sørensen, *The levelling off of the obesity epidemic since the year 1999–a review of evidence and perspectives.* Obesity reviews, 2010. **11**(12): p. 835-846.
8. Statistics Canada, *Table 13-10-0373-01 Overweight and obesity based on measured body mass index, by age group and sex.* 2020 DOI: https://doi.org/10.25318/1310037301-eng.
9. Styne, D.M., et al., *Pediatric Obesity-Assessment, Treatment, and Prevention: An Endocrine Society Clinical Practice Guideline.* J Clin Endocrinol Metab, 2017. **102**(3): p. 709-757 DOI: 10.1210/jc.2016-2573.
10. Simmonds, M., et al., *Predicting adult obesity from childhood obesity: a systematic review and meta-analysis.* Obesity Reviews, 2016. **17**(2): p. 95-107 DOI: https://doi.org/10.1111/obr.12334.
11. Swinburn, B.A., et al., *The global obesity pandemic: shaped by global drivers and local environments.* Lancet, 2011. **378**(9793): p. 804-14 DOI: 10.1016/s0140-6736(11)60813-1.
12. Silventoinen, K., et al., *Differences in genetic and environmental variation in adult BMI by sex, age, time period, and region: an individual-based pooled analysis of 40 twin cohorts.* The American Journal of Clinical Nutrition, 2017. **106**(2): p. 457-466 DOI: https://doi.org/10.3945/ajcn.117.153643.
13. Silventoinen, K., et al., *Genetic and environmental effects on body mass index from infancy to the onset of adulthood: an individual-based pooled analysis of 45 twin cohorts participating in the COllaborative project of Development of Anthropometrical measures in Twins (CODATwins) study.* The American journal of clinical nutrition, 2016. **104**(2): p. 371-379 DOI: 10.3945/ajcn.116.130252.

14. Yengo, L., et al., *Meta-analysis of genome-wide association studies for height and body mass index in ~700000 individuals of European ancestry.* Human molecular genetics, 2018. **27**(20): p. 3641-3649 DOI: 10.1093/hmg/ddy271.
15. Loos, R.J.F. and G.S.H. Yeo, *The genetics of obesity: from discovery to biology.* Nature Reviews Genetics, 2022. **23**(2): p. 120-133 DOI: 10.1038/s41576-021-00414-z.
16. Khera, A.V., et al., *Polygenic Prediction of Weight and Obesity Trajectories from Birth to Adulthood.* Cell, 2019. **177**(3): p. 587-596.e9 DOI: 10.1016/j.cell.2019.03.028.
17. Silventoinen, K., et al., *Changing genetic architecture of body mass index from infancy to early adulthood: an individual based pooled analysis of 25 twin cohorts.* Int J Obes (Lond), 2022. **46**(10): p. 1901-1909 DOI: 10.1038/s41366-022-01202-3.
18. Llewellyn, C.H., et al., *Behavioural susceptibility theory: the role of appetite in genetic susceptibility to obesity in early life.* Philosophical Transactions of the Royal Society B: Biological Sciences, 2023. **378**(1885): p. 20220223 DOI: 10.1098/rstb.2022.0223.
19. Locke, A.E., et al., *Genetic studies of body mass index yield new insights for obesity biology.* Nature, 2015. **518**(7538): p. 197-206 DOI: 10.1038/nature14177.
20. Ndiaye, F.K., et al., *The expression of genes in top obesity-associated loci is enriched in insula and substantia nigra brain regions involved in addiction and reward.* International Journal of Obesity, 2020. **44**(2): p. 539-543 DOI: 10.1038/s41366-019-0428-7.
21. Vogelezang, S., et al., *Novel loci for childhood body mass index and shared heritability with adult cardiometabolic traits.* PLoS genetics, 2020. **16**(10): p. e1008718-e1008718 DOI: 10.1371/journal.pgen.1008718.
22. Schrempft, S., et al., *Variation in the Heritability of Child Body Mass Index by Obesogenic Home Environment.* JAMA Pediatr, 2018. **172**(12): p. 1153-1160 DOI: 10.1001/jamapediatrics.2018.1508.
23. VanderWeele, T.J., *Explanation in causal inference : methods for mediation and interaction*. 2015, New York: Oxford University Press.
24. Pérusse, L., et al., *Understanding Gene-Lifestyle Interaction in Obesity: The Role of Mediation versus Moderation.* Lifestyle Genomics, 2022. **15**(2): p. 67-76 DOI: 10.1159/000523813.
25. Silventoinen, K. and H. Konttinen, *Obesity and eating behavior from the perspective of twin and genetic research.* Neuroscience & Biobehavioral Reviews, 2020. **109**: p. 150-165 DOI: https://doi.org/10.1016/j.neubiorev.2019.12.012.
26. Han, H.Y., et al., *Interactions between Polygenic Risk of Obesity and Dietary Factors on Anthropometric Outcomes: A Systematic Review and Meta-Analysis of Observational Studies.* The Journal of Nutrition, 2024 DOI: https://doi.org/10.1016/j.tjnut.2024.10.014.
27. National Research Council (US) Committee on Diet and Health, *Calories: Total Macronutrient Intake, Energy Expenditure, and Net Energy Stores.*, in *Diet and Health: Implications for Reducing Chronic Disease Risk*. 1989, National Academies Press (US): Washington (DC).
28. National Academies of Sciences, E., and Medicine,, *Factors Affecting Energy Expenditure and Requirements*, in *Dietary Reference Intakes for Energy*. 2023, National Academies Press (US): Washington (DC).
29. Oussaada, S.M., et al., *The pathogenesis of obesity.* Metabolism, 2019. **92**: p. 26-36 DOI: 10.1016/j.metabol.2018.12.012.
30. Saeed, S., A. Bonnefond, and P. Froguel, *Obesity: exploring its connection to brain function through genetic and genomic perspectives.* Molecular Psychiatry, 2024 DOI: 10.1038/s41380-024-02737-9.
31. Koliaki, C., S. Liatis, and A. Kokkinos, *Obesity and cardiovascular disease: revisiting an old relationship.* Metabolism - Clinical and Experimental, 2019. **92**: p. 98-107 DOI: 10.1016/j.metabol.2018.10.011.

32. Bastien, M., et al., *Overview of epidemiology and contribution of obesity to cardiovascular disease.* Prog Cardiovasc Dis, 2014. **56**(4): p. 369-81 DOI: 10.1016/j.pcad.2013.10.016.
33. Després, J.P., et al., *Regional distribution of body fat, plasma lipoproteins, and cardiovascular disease.* Arteriosclerosis, 1990. **10**(4): p. 497-511 DOI: 10.1161/01.atv.10.4.497.
34. Liu, J., et al., *Impact of Abdominal Visceral and Subcutaneous Adipose Tissue on Cardiometabolic Risk Factors: The Jackson Heart Study.* The Journal of Clinical Endocrinology & Metabolism, 2010. **95**(12): p. 5419-5426 DOI: 10.1210/jc.2010-1378.
35. Dagenais, G.R., et al., *Prognostic impact of body weight and abdominal obesity in women and men with cardiovascular disease.* American Heart Journal, 2005. **149**(1): p. 54-60 DOI: https://doi.org/10.1016/j.ahj.2004.07.009.
36. de Koning, L., et al., *Waist circumference and waist-to-hip ratio as predictors of cardiovascular events: meta-regression analysis of prospective studies.* European Heart Journal, 2007. **28**(7): p. 850-856 DOI: 10.1093/eurheartj/ehm026.
37. Kansra, A.R., S. Lakkunarajah, and M.S. Jay, *Childhood and Adolescent Obesity: A Review.* Frontiers in pediatrics, 2021. **8**: p. 581461-581461 DOI: 10.3389/fped.2020.581461.
38. Adab, P., M. Pallan, and P.H. Whincup, *Is BMI the best measure of obesity?* BMJ, 2018. **360**: p. k1274 DOI: 10.1136/bmj.k1274.
39. Bell, J.A., et al., *Associations of Body Mass and Fat Indexes With Cardiometabolic Traits.* Journal of the American College of Cardiology, 2018. **72**(24): p. 3142-3154 DOI: https://doi.org/10.1016/j.jacc.2018.09.066.
40. Douketis, J.D., et al., *Canadian guidelines for body weight classification in adults: application in clinical practice to screen for overweight and obesity and to assess disease risk.* Cmaj, 2005. **172**(8): p. 995-8 DOI: 10.1503/cmaj.045170.
41. Haththotuwa, R.N., C.N. Wijeyaratne, and U. Senarath, *Chapter 1 - Worldwide epidemic of obesity*, in *Obesity and Obstetrics (Second Edition)*, T.A. Mahmood, S. Arulkumaran, and F.A. Chervenak, Editors. 2020, Elsevier. p. 3-8.
42. WHO, *WHO Child Growth Standards based on length/height, weight and age.* Acta Paediatr Suppl, 2006. **450**: p. 76-85 DOI: 10.1111/j.1651-2227.2006.tb02378.x.
43. de Onis, M., et al., *Development of a WHO growth reference for school-aged children and adolescents.* Bulletin of the World Health Organization, 2007. **85**(9): p. 660-667 DOI: 10.2471/blt.07.043497.
44. Di Angelantonio, E., et al., *Body-mass index and all-cause mortality: individual-participant-data meta-analysis of 239 prospective studies in four continents.* The Lancet, 2016. **388**(10046): p. 776-786 DOI: 10.1016/S0140-6736(16)30175-1.
45. Bhaskaran, K., et al., *Association of BMI with overall and cause-specific mortality: a population-based cohort study of 3·6 million adults in the UK.* The lancet. Diabetes & endocrinology, 2018. **6**(12): p. 944-953 DOI: 10.1016/S2213-8587(18)30288-2.
46. Banack, H.R. and A. Stokes, *The 'obesity paradox' may not be a paradox at all.* International Journal of Obesity, 2017. **41**(8): p. 1162-1163 DOI: 10.1038/ijo.2017.99.
47. Li, J., et al., *Association of BMI, comorbidities and all-cause mortality by using a baseline mortality risk model.* PLOS ONE, 2021. **16**(7): p. e0253696 DOI: 10.1371/journal.pone.0253696.
48. Brauer, M., et al., *Global burden and strength of evidence for 88 risk factors in 204 countries and 811 subnational locations, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021.* The Lancet, 2024. **403**(10440): p. 2162-2203 DOI: 10.1016/S0140-6736(24)00933-4.

49. Kivimäki, M., et al., *Body-mass index and risk of obesity-related complex multimorbidity: an observational multicohort study.* Lancet Diabetes Endocrinol, 2022. **10**(4): p. 253-263 DOI: 10.1016/s2213-8587(22)00033-x.
50. Vazquez, G., et al., *Comparison of Body Mass Index, Waist Circumference, and Waist/Hip Ratio in Predicting Incident Diabetes: A Meta-Analysis.* Epidemiologic Reviews, 2007. **29**(1): p. 115-128 DOI: 10.1093/epirev/mxm008.
51. Abdullah, A., et al., *The magnitude of association between overweight and obesity and the risk of diabetes: a meta-analysis of prospective cohort studies.* Diabetes Res Clin Pract, 2010. **89**(3): p. 309-19 DOI: 10.1016/j.diabres.2010.04.012.
52. Bell, J.A., M. Kivimaki, and M. Hamer, *Metabolically healthy obesity and risk of incident type 2 diabetes: a meta-analysis of prospective cohort studies.* Obes Rev, 2014. **15**(6): p. 504-15 DOI: 10.1111/obr.12157.
53. Vucenik, I. and J.P. Stains, *Obesity and cancer risk: evidence, mechanisms, and recommendations.* Annals of the New York Academy of Sciences, 2012. **1271**(1): p. 37-43 DOI: https://doi.org/10.1111/j.1749-6632.2012.06750.x.
54. Polednak, A.P., *Estimating the number of U.S. incident cancers attributable to obesity and the impact on temporal trends in incidence rates for obesity-related cancers.* Cancer Detect Prev, 2008. **32**(3): p. 190-9 DOI: 10.1016/j.cdp.2008.08.004.
55. Hruby, A. and F.B. Hu, *The Epidemiology of Obesity: A Big Picture.* Pharmacoeconomics, 2015. **33**(7): p. 673-89 DOI: 10.1007/s40273-014-0243-x.
56. Wilkins, K., et al., *Blood pressure in Canadian adults.* Health reports, 2010. **21**(1): p. 37.
57. Marinou, K., et al., *Obesity and cardiovascular disease: From pathophysiology to risk stratification.* International Journal of Cardiology, 2010. **138**(1): p. 3-8 DOI: https://doi.org/10.1016/j.ijcard.2009.03.135.
58. Poirier, P., et al., *Obesity and cardiovascular disease: pathophysiology, evaluation, and effect of weight loss: an update of the 1997 American Heart Association Scientific Statement on Obesity and Heart Disease from the Obesity Committee of the Council on Nutrition, Physical Activity, and Metabolism.* Circulation, 2006. **113**(6): p. 898-918.
59. Wormser, D., et al., *Emerging Risk Factors Collaboration Separate and combined associations of body-mass index and abdominal adiposity with cardiovascular disease: collaborative analysis of 58 prospective studies.* Lancet, 2011. **377**(9771): p. 1085-95.
60. Wilson, P.W., et al., *Overweight and obesity as determinants of cardiovascular risk: the Framingham experience.* Archives of internal medicine, 2002. **162**(16): p. 1867-1872.
61. Flint, A.J., et al., *Excess weight and the risk of incident coronary heart disease among men and women.* Obesity (Silver Spring), 2010. **18**(2): p. 377-83 DOI: 10.1038/oby.2009.223.
62. Petrilli, C.M., et al., *Factors associated with hospital admission and critical illness among 5279 people with coronavirus disease 2019 in New York City: prospective cohort study.* Bmj, 2020. **369**: p. m1966 DOI: 10.1136/bmj.m1966.
63. Cummings, M.J., et al., *Epidemiology, clinical course, and outcomes of critically ill adults with COVID-19 in New York City: a prospective cohort study.* Lancet, 2020. **395**(10239): p. 1763-1770 DOI: 10.1016/s0140-6736(20)31189-2.
64. Reilly, J.J. and J. Kelly, *Long-term impact of overweight and obesity in childhood and adolescence on morbidity and premature mortality in adulthood: systematic review.* International Journal of Obesity, 2011. **35**(7): p. 891-898 DOI: 10.1038/ijo.2010.222.
65. Twig, G., et al., *Body-Mass Index in 2.3 Million Adolescents and Cardiovascular Death in Adulthood.* New England Journal of Medicine, 2016. **374**(25): p. 2430-2440 DOI: doi:10.1056/NEJMoa1503840.

66. Park, M.H., et al., *The impact of childhood obesity on morbidity and mortality in adulthood: a systematic review.* Obesity Reviews, 2012. **13**(11): p. 985-1000 DOI: https://doi.org/10.1111/j.1467-789X.2012.01015.x.
67. Weihrauch-Blüher, S., P. Schwarz, and J.-H. Klusmann, *Childhood obesity: increased risk for cardiometabolic disease and cancer in adulthood.* Metabolism, 2019. **92**: p. 147-152 DOI: https://doi.org/10.1016/j.metabol.2018.12.001.
68. National Cancer Institute. *NCI Dictionary of Genetics Terms*. [cited 2024; Available from: https://www.cancer.gov/publications/dictionaries/genetics-dictionary/.
69. Barry, C.S., et al., *How to estimate heritability: a guide for genetic epidemiologists.* Int J Epidemiol, 2023. **52**(2): p. 624-632 DOI: 10.1093/ije/dyac224.
70. Young, A.I., *Solving the missing heritability problem.* PLOS Genetics, 2019. **15**(6): p. e1008222 DOI: 10.1371/journal.pgen.1008222.
71. Bulik-Sullivan, B.K., et al., *LD Score regression distinguishes confounding from polygenicity in genome-wide association studies.* Nat Genet, 2015. **47**(3): p. 291-5 DOI: 10.1038/ng.3211.
72. Yang, J., et al., *GCTA: a tool for genome-wide complex trait analysis.* American journal of human genetics, 2011. **88**(1): p. 76-82 DOI: 10.1016/j.ajhg.2010.11.011.
73. Huang, J., et al., *Genomics and phenomics of body mass index reveals a complex disease network.* Nature Communications, 2022. **13**(1): p. 7973 DOI: 10.1038/s41467-022-35553-2.
74. Yang, J., et al., *Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index.* Nature Genetics, 2015. **47**(10): p. 1114-1120 DOI: 10.1038/ng.3390.
75. Yang, J., et al., *Genome-wide genetic homogeneity between sexes and populations for human height and body mass index.* Human Molecular Genetics, 2015. **24**(25): p. 7445-7449 DOI: 10.1093/hmg/ddv443.
76. Yang, J., et al., *Concepts, estimation and interpretation of SNP-based heritability.* Nature Genetics, 2017. **49**(9): p. 1304-1310 DOI: 10.1038/ng.3941.
77. Eichler, E.E., et al., *Missing heritability and strategies for finding the underlying causes of complex disease.* Nat Rev Genet, 2010. **11**(6): p. 446-50 DOI: 10.1038/nrg2809.
78. Visscher, P.M., et al., *10 Years of GWAS Discovery: Biology, Function, and Translation.* The American Journal of Human Genetics, 2017. **101**(1): p. 5-22 DOI: 10.1016/j.ajhg.2017.06.005.
79. Frayling, T.M., et al., *A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity.* Science, 2007. **316**(5826): p. 889-94 DOI: 10.1126/science.1141634.
80. Willer, C.J., et al., *Six new loci associated with body mass index highlight a neuronal influence on body weight regulation.* Nat Genet, 2009. **41**(1): p. 25-34 DOI: 10.1038/ng.287.
81. Speliotes, E.K., et al., *Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index.* Nat Genet, 2010. **42**(11): p. 937-48 DOI: 10.1038/ng.686.
82. Berndt, S.I., et al., *Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture.* Nat Genet, 2013. **45**(5): p. 501-12 DOI: 10.1038/ng.2606.
83. Buniello, A., et al., *The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019.* Nucleic acids research, 2019. **47**(D1): p. D1005-D1012 DOI: 10.1093/nar/gky1120.
84. Kilpeläinen, T.O., et al., *Genetic variation near IRS1 associates with reduced adiposity and an impaired metabolic profile.* Nat Genet, 2011. **43**(8): p. 753-60 DOI: 10.1038/ng.866.
85. Lu, Y., et al., *New loci for body fat percentage reveal link between adiposity and cardiometabolic disease risk.* Nat Commun, 2016. **7**: p. 10495 DOI: 10.1038/ncomms10495.

86. Pulit, S.L., et al., *Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry.* Human Molecular Genetics, 2019. **28**(1): p. 166-174 DOI: 10.1093/hmg/ddy327.
87. Chu, A.Y., et al., *Multiethnic genome-wide meta-analysis of ectopic fat depots identifies loci associated with adipocyte development and differentiation.* Nat Genet, 2017. **49**(1): p. 125-130 DOI: 10.1038/ng.3738.
88. Martin, A.R., et al., *Clinical use of current polygenic risk scores may exacerbate health disparities.* Nat Genet, 2019. **51**(4): p. 584-591 DOI: 10.1038/s41588-019-0379-x.
89. Albuquerque, D., et al., *Current review of genetics of human obesity: from molecular mechanisms to an evolutionary perspective.* Molecular Genetics and Genomics, 2015. **290**(4): p. 1191-1221 DOI: 10.1007/s00438-015-1015-9.
90. Bradfield, J.P., et al., *A genome-wide association meta-analysis identifies new childhood obesity loci.* Nature genetics, 2012. **44**(5): p. 526-531 DOI: 10.1038/ng.2247.
91. Bradfield, J.P., et al., *A trans-ancestral meta-analysis of genome-wide association studies reveals loci associated with childhood obesity.* Human molecular genetics, 2019. **28**(19): p. 3327-3338 DOI: 10.1093/hmg/ddz161.
92. Graff, M., et al., *Genome-wide analysis of BMI in adolescents and young adults reveals additional insight into the effects of genetic loci over the life course.* Human molecular genetics, 2013. **22**(17): p. 3597-3607 DOI: 10.1093/hmg/ddt205.
93. Melén, E., et al., *Genome-wide association study of body mass index in 23 000 individuals with and without asthma.* Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology, 2013. **43**(4): p. 463-474 DOI: 10.1111/cea.12054.
94. Namjou, B., et al., *EMR-linked GWAS study: investigation of variation landscape of loci for body mass index in children.* Frontiers in genetics, 2013. **4**: p. 268-268 DOI: 10.3389/fgene.2013.00268.
95. Stergiakouli, E., et al., *Genome-wide association study of height-adjusted BMI in childhood identifies functional variant in ADCY3.* Obesity (Silver Spring, Md.), 2014. **22**(10): p. 2252-2259 DOI: 10.1002/oby.20840.
96. Warrington, N.M., et al., *A genome-wide association study of body mass index across early life and childhood.* International journal of epidemiology, 2015. **44**(2): p. 700-712 DOI: 10.1093/ije/dyv077.
97. Felix, J.F., et al., *Genome-wide association analysis identifies three new susceptibility loci for childhood body mass index.* Human molecular genetics, 2016. **25**(2): p. 389-403 DOI: 10.1093/hmg/ddv472.
98. Hao, L., et al., *Development of a clinical polygenic risk score assay and reporting workflow.* Nature Medicine, 2022 DOI: 10.1038/s41591-022-01767-6.
99. Lambert, S.A., et al., *The Polygenic Score Catalog as an open database for reproducibility and systematic evaluation.* Nature Genetics, 2021. **53**(4): p. 420-425 DOI: 10.1038/s41588-021-00783-5.
100. Dudbridge, F., *Polygenic Mendelian Randomization.* Cold Spring Harb Perspect Med, 2021. **11**(2) DOI: 10.1101/cshperspect.a039586.
101. Lewis, C.M. and E. Vassos, *Polygenic risk scores: from research tools to clinical instruments.* Genome Medicine, 2020. **12**(1): p. 44 DOI: 10.1186/s13073-020-00742-5.
102. Tyrrell, J., et al., *Gene–obesogenic environment interactions in the UK Biobank study.* International Journal of Epidemiology, 2017. **46**(2): p. 559-575 DOI: 10.1093/ije/dyw337 %J International Journal of Epidemiology.
103. Solovieff, N., et al., *Pleiotropy in complex traits: challenges and strategies.* Nat Rev Genet, 2013. **14**(7): p. 483-95 DOI: 10.1038/nrg3461.

104. Choi, S.W., T.S.-H. Mak, and P.F. O'Reilly, *Tutorial: a guide to performing polygenic risk score analyses.* Nature Protocols, 2020. **15**(9): p. 2759-2772 DOI: 10.1038/s41596-020-0353-1.
105. Bycroft, C., et al., *The UK Biobank resource with deep phenotyping and genomic data.* Nature, 2018. **562**(7726): p. 203-209 DOI: 10.1038/s41586-018-0579-z.
106. Wand, H., et al., *Improving reporting standards for polygenic scores in risk prediction studies.* Nature, 2021. **591**(7849): p. 211-219 DOI: 10.1038/s41586-021-03243-6.
107. Hollands, G.J., et al., *The impact of communicating genetic risks of disease on risk-reducing health behaviour: systematic review with meta-analysis.* Bmj, 2016. **352**: p. i1102 DOI: 10.1136/bmj.i1102.
108. Li, S., et al., *Cumulative effects and predictive value of common obesity-susceptibility variants identified by genome-wide association studies.* Am J Clin Nutr, 2010. **91**(1): p. 184-90 DOI: 10.3945/ajcn.2009.28403.
109. Loos, R.J.F. and A.C.J.W. Janssens, *Predicting Polygenic Obesity Using Genetic Information.* Cell Metabolism, 2017. **25**(3): p. 535-543 DOI: 10.1016/j.cmet.2017.02.013.
110. Richardson, T.G., et al., *Use of genetic variation to separate the effects of early and later life adiposity on disease risk: mendelian randomisation study.* BMJ (Clinical research ed.), 2020. **369**: p. m1203-m1203 DOI: 10.1136/bmj.m1203.
111. Truong, B., et al., *Integrative polygenic risk score improves the prediction accuracy of complex traits and diseases.* Cell Genomics, 2024. **4**(4) DOI: 10.1016/j.xgen.2024.100523.
112. Tanigawa, Y., et al., *Significant sparse polygenic risk scores across 813 traits in UK Biobank.* PLOS Genetics, 2022. **18**(3): p. e1010105 DOI: 10.1371/journal.pgen.1010105.
113. Tanigawa, Y. and M. Kellis, *Power of inclusion: Enhancing polygenic prediction with admixed individuals.* The American Journal of Human Genetics, 2023. **110**(11): p. 1888-1902 DOI: 10.1016/j.ajhg.2023.09.013.
114. Raben, T.G., et al., *Biobank-scale methods and projections for sparse polygenic prediction from machine learning.* Scientific Reports, 2023. **13**(1): p. 11662 DOI: 10.1038/s41598-023-37580-5.
115. Monti, R., et al., *Evaluation of polygenic scoring methods in five biobanks shows larger variation between biobanks than methods and finds benefits of ensemble learning.* The American Journal of Human Genetics, 2024. **111**(7): p. 1431-1447 DOI: 10.1016/j.ajhg.2024.06.003.
116. Jung, H., et al., *Integration of risk factor polygenic risk score with disease polygenic risk score for disease prediction.* Communications Biology, 2024. **7**(1): p. 180 DOI: 10.1038/s42003-024-05874-7.
117. Vilhjálmsson, B.J., et al., *Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores.* American journal of human genetics, 2015. **97**(4): p. 576-592 DOI: 10.1016/j.ajhg.2015.09.001.
118. Mak, T.S.H., et al., *Polygenic scores via penalized regression on summary statistics.* Genetic Epidemiology, 2017. **41**(6): p. 469-480 DOI: https://doi.org/10.1002/gepi.22050.
119. Ge, T., et al., *Polygenic prediction via Bayesian regression and continuous shrinkage priors.* Nature communications, 2019. **10**(1): p. 1776-1776 DOI: 10.1038/s41467-019-09718-5.
120. Lloyd-Jones, L.R., et al., *Improved polygenic prediction by Bayesian multiple regression on summary statistics.* Nature Communications, 2019. **10**(1): p. 5086 DOI: 10.1038/s41467-019-12653-0.
121. Ma, Y., et al., *ExPRSweb: An online repository with polygenic risk scores for common health-related exposures.* The American Journal of Human Genetics, 2022. **109**(10): p. 1742-1760 DOI: 10.1016/j.ajhg.2022.09.001.

122. Borisevich, D., et al., *Non-linear interaction between physical activity and polygenic risk score of body mass index in Danish and Russian populations.* PloS one, 2021. **16**(10): p. e0258748-e0258748 DOI: 10.1371/journal.pone.0258748.
123. Coleman, J.R.I., et al., *Individual and shared effects of social environment and polygenic risk scores on adolescent body mass index.* Scientific reports, 2018. **8**(1): p. 6344-6344 DOI: 10.1038/s41598-018-24774-5.
124. Xie, T., et al., *Genetic Risk Scores for Complex Disease Traits in Youth.* Circulation: Genomic and Precision Medicine, 2020. **13**(4): p. e002775 DOI: 10.1161/CIRCGEN.119.002775.
125. Hüls, A., et al., *Polygenic risk for obesity and its interaction with lifestyle and sociodemographic factors in European children and adolescents.* International Journal of Obesity, 2021. **45**(6): p. 1321-1330 DOI: 10.1038/s41366-021-00795-5.
126. Belsky, D.W., et al., *Polygenic risk, rapid childhood growth, and the development of obesity: evidence from a 4-decade longitudinal study.* Arch Pediatr Adolesc Med, 2012. **166**(6): p. 515-21 DOI: 10.1001/archpediatrics.2012.131.
127. Sanz-de-Galdeano, A., A. Terskaya, and A. Upegui, *Association of a genetic risk score with BMI along the life-cycle: Evidence from several US cohorts.* PLoS One, 2020. **15**(9): p. e0239067 DOI: 10.1371/journal.pone.0239067.
128. Lange, K., et al., *Can adult polygenic scores improve prediction of body mass index in childhood?* International Journal of Obesity, 2022. **46**(7): p. 1375-1383 DOI: 10.1038/s41366-022-01130-2.
129. Helgeland, Ø., et al., *Characterization of the genetic architecture of infant and early childhood body mass index.* Nature Metabolism, 2022 DOI: 10.1038/s42255-022-00549-1.
130. Carnell, S. and J. Wardle, *Appetite and adiposity in children: evidence for a behavioral susceptibility theory of obesity.* The American Journal of Clinical Nutrition, 2008. **88**(1): p. 22-29 DOI: 10.1093/ajcn/88.1.22.
131. Llewellyn, C.H. and A. Fildes, *Behavioural Susceptibility Theory: Professor Jane Wardle and the Role of Appetite in Genetic Risk of Obesity.* Curr Obes Rep, 2017. **6**(1): p. 38-45 DOI: 10.1007/s13679-017-0247-x.
132. Keskitalo, K., et al., *The Three-Factor Eating Questionnaire, body mass index, and responses to sweet and salty fatty foods: a twin study of genetic and environmental associations.* Am J Clin Nutr, 2008. **88**(2): p. 263-71 DOI: 10.1093/ajcn/88.2.263.
133. Sung, J., et al., *Heritability of eating behavior assessed using the DEBQ (Dutch Eating Behavior Questionnaire) and weight-related traits: the Healthy Twin Study.* Obesity (Silver Spring), 2010. **18**(5): p. 1000-5 DOI: 10.1038/oby.2009.389.
134. Herle, M., et al., *The Role of the Environment in Overweight and Eating Behavior Variability: Insights from a Multivariate Twin Study.* Twin Research and Human Genetics, 2020. **23**(6): p. 338-344 DOI: 10.1017/thg.2020.90.
135. Llewellyn, C.H., et al., *Inherited behavioral susceptibility to adiposity in infancy: a multivariate genetic analysis of appetite and weight in the Gemini birth cohort123.* The American Journal of Clinical Nutrition, 2012. **95**(3): p. 633-639 DOI: https://doi.org/10.3945/ajcn.111.023671.
136. Gagnon, E., et al., *Genetic control of body weight by the human brain proteome.* iScience, 2023. **26**(4) DOI: 10.1016/j.isci.2023.106376.
137. Bouchard, C., *The study of human variability became a passion.* European Journal of Clinical Nutrition, 2022. **76**(5): p. 631-636 DOI: 10.1038/s41430-021-00871-z.
138. Hemphill, E., et al., *Exploring Obesogenic Food Environments in Edmonton, Canada: The Association between Socioeconomic Factors and Fast-Food Outlet Access.* American Journal of Health Promotion, 2008. **22**(6): p. 426-431 DOI: 10.4278/ajhp.22.6.426.

139. Smoyer-Tomic, K.E., et al., *The association between neighborhood socioeconomic status and exposure to supermarkets and fast food outlets.* Health Place, 2008. **14**(4): p. 740-54 DOI: 10.1016/j.healthplace.2007.12.001.
140. Seale, E., M.d. Groh, and L. Greene-Finestone, *Fast food consumption in adults living in Canada: alternative measurement methods, consumption choices, and correlates.* Applied Physiology, Nutrition, and Metabolism, 2023. **48**(2): p. 163-171 DOI: 10.1139/apnm-2022-0252 %M 36322952.
141. Ravensbergen, L., et al., *"Socioeconomic inequalities in children's accessibility to food retailing: Examining the roles of mobility and time".* Social Science & Medicine, 2016. **153**: p. 81-89 DOI: https://doi.org/10.1016/j.socscimed.2016.01.030.
142. Min, J., D.T. Chiu, and Y. Wang, *Variation in the heritability of body mass index based on diverse twin studies: a systematic review.* Obesity Reviews, 2013. **14**(11): p. 871-882 DOI: https://doi.org/10.1111/obr.12065.
143. Dinescu, D., et al., *Socioeconomic modifiers of genetic and environmental influences on body mass index in adult twins.* Health Psychol, 2016. **35**(2): p. 157-66 DOI: 10.1037/hea0000255.
144. Johnson, W., et al., *Education modifies genetic and environmental influences on BMI.* PLoS One, 2011. **6**(1): p. e16290 DOI: 10.1371/journal.pone.0016290.
145. Silventoinen, K., et al., *Parental Education and Genetics of BMI from Infancy to Old Age: A Pooled Analysis of 29 Twin Cohorts.* Obesity, 2019. **27**(5): p. 855-865 DOI: https://doi.org/10.1002/oby.22451.
146. Park, S., et al., *Interactions between Polygenic Risk Scores, Dietary Pattern, and Menarche Age with the Obesity Risk in a Large Hospital-Based Cohort.* Nutrients, 2021. **13**(11): p. 3772 DOI: 10.3390/nu13113772.
147. Stunkard, A.J. and S. Messick, *The three-factor eating questionnaire to measure dietary restraint, disinhibition and hunger.* Journal of Psychosomatic Research, 1985. **29**(1): p. 71-83 DOI: https://doi.org/10.1016/0022-3999(85)90010-8.
148. Karlsson, J., et al., *Psychometric properties and factor structure of the Three-Factor Eating Questionnaire (TFEQ) in obese men and women. Results from the Swedish Obese Subjects (SOS) study.* International Journal of Obesity, 2000. **24**(12): p. 1715-1725 DOI: 10.1038/sj.ijo.0801442.
149. Anglé, S., et al., *Three factor eating questionnaire-R18 as a measure of cognitive restraint, uncontrolled eating and emotional eating in a sample of young Finnish females.* International Journal of Behavioral Nutrition and Physical Activity, 2009. **6**: p. 1-7.
150. Wardle, J., et al., *Development of the Children's Eating Behaviour Questionnaire.* Journal of Child Psychology and Psychiatry, 2001. **42**(7): p. 963-970 DOI: https://doi.org/10.1111/1469-7610.00792.
151. Ashcroft, J., et al., *Continuity and stability of eating behaviour traits in children.* European Journal of Clinical Nutrition, 2008. **62**(8): p. 985-990 DOI: 10.1038/sj.ejcn.1602855.
152. Jacob, R., et al., *The role of eating behavior traits in mediating genetic susceptibility to obesity.* The American Journal of Clinical Nutrition, 2018. **108**(3): p. 445-452 DOI: 10.1093/ajcn/nqy130 %J The American Journal of Clinical Nutrition.
153. Konttinen, H., et al., *Appetitive traits as behavioural pathways in genetic susceptibility to obesity: a population-based cross-sectional study.* Scientific Reports, 2015. **5**(1): p. 14726 DOI: 10.1038/srep14726.
154. de Lauzon-Guillain, B., et al., *Mediation and modification of genetic susceptibility to obesity by eating behaviors.* The American Journal of Clinical Nutrition, 2017. **106**(4): p. 996-1004 DOI: 10.3945/ajcn.117.157396 %J The American Journal of Clinical Nutrition.

155. Masip, G., et al., *The genetic architecture of the association between eating behaviors and obesity: combining genetic twin modeling and polygenic risk scores.* The American Journal of Clinical Nutrition, 2020. **112**(4): p. 956-966 DOI: 10.1093/ajcn/nqaa181.
156. Brunner, E.J., et al., *Appetite disinhibition rather than hunger explains genetic effects on adult BMI trajectory.* International Journal of Obesity, 2021. **45**(4): p. 758-765 DOI: 10.1038/s41366-020-00735-9.
157. Begum, S., et al., *Mediation and moderation of genetic risk of obesity through eating behaviours in two UK cohorts.* International Journal of Epidemiology, 2023. **52**(6): p. 1926-1938 DOI: 10.1093/ije/dyad092.
158. Llewellyn, C.H., et al., *Satiety Mechanisms in Genetic Risk of Obesity.* JAMA Pediatrics, 2014. **168**(4): p. 338-344 DOI: 10.1001/jamapediatrics.2013.4944 %J JAMA Pediatrics.
159. Steinsbekk, S., et al., *Polygenic Risk, Appetite Traits, and Weight Gain in Middle Childhood: A Longitudinal Study.* JAMA Pediatrics, 2016. **170**(2): p. e154472-e154472 DOI: 10.1001/jamapediatrics.2015.4472 %J JAMA Pediatrics.
160. de Lauzon-Guillain, B., et al., *Association between genetic obesity susceptibility and mother-reported eating behaviour in children up to 5 years.* 2019. **14**(5): p. e12496 DOI: https://doi.org/10.1111/ijpo.12496.
161. Dubois, L., et al., *Eating behaviors, dietary patterns and weight status in emerging adulthood and longitudinal associations with eating behaviors in early childhood.* Int J Behav Nutr Phys Act, 2022. **19**(1): p. 139 DOI: 10.1186/s12966-022-01376-z.
162. Wang, T., et al., *Improving adherence to healthy dietary patterns, genetic risk, and long term weight gain: gene-diet interaction analysis in two prospective cohort studies.* BMJ (Clinical research ed.), 2018. **360**: p. j5644-j5644 DOI: 10.1136/bmj.j5644.
163. Ding, M., et al., *Diet quality and genetic association with body mass index: results from 3 observational studies.* The American journal of clinical nutrition, 2018. **108**(6): p. 1291-1300 DOI: 10.1093/ajcn/nqy203.
164. Chang, X., et al., *Gene-diet interaction effects on BMI levels in the Singapore Chinese population.* Nutrition Journal, 2018. **17**(1): p. 31 DOI: 10.1186/s12937-018-0340-3.
165. Baratali, L., M. Mean, and P. Marques-Vidal, *Impact of dietary and obesity genetic risk scores on weight gain.* The American Journal of Clinical Nutrition, 2021. **114**(2): p. 741-751 DOI: https://doi.org/10.1093/ajcn/nqab069.
166. Livingstone, K.M., et al., *Associations between three diet quality indices, genetic risk and body composition: A prospective cohort study.* Clinical Nutrition, 2022. **41**(9): p. 1942-1949 DOI: 10.1016/j.clnu.2022.07.005.
167. Heianza, Y., et al., *Healthful plant-based dietary patterns, genetic risk of obesity, and cardiovascular risk in the UK biobank study.* Clin Nutr, 2021. **40**(7): p. 4694-4701 DOI: 10.1016/j.clnu.2021.06.018.
168. Han, H.Y., et al. *Diet Quality and Food Prices Modify Associations between Genetic Susceptibility to Obesity and Adiposity Outcomes*. Nutrients, 2020. **12**, DOI: 10.3390/nu12113349.
169. Nettleton, J.A., et al., *Gene × dietary pattern interactions in obesity: analysis of up to 68 317 adults of European ancestry.* Human Molecular Genetics, 2015. **24**(16): p. 4728-4738 DOI: 10.1093/hmg/ddv186 %J Human Molecular Genetics.
170. Seral-Cortes, M., et al., *Interaction Effect of the Mediterranean Diet and an Obesity Genetic Risk Score on Adiposity and Metabolic Syndrome in Adolescents: The HELENA Study.* Nutrients, 2020. **12**(12) DOI: 10.3390/nu12123841.
171. Qi, Q., et al., *Sugar-Sweetened Beverages and Genetic Risk of Obesity.* 2012. **367**(15): p. 1387-1396 DOI: 10.1056/NEJMoa1203039.

172. Brunkwall, L., et al., *Sugar-sweetened beverage consumption and genetic predisposition to obesity in 2 Swedish cohorts.* The American journal of clinical nutrition, 2016. **104**(3): p. 809-815 DOI: 10.3945/ajcn.115.126052.
173. Olsen, N.J., et al., *Interactions between genetic variants associated with adiposity traits and soft drinks in relation to longitudinal changes in body weight and waist circumference.* The American Journal of Clinical Nutrition, 2016. **104**(3): p. 816-826 DOI: 10.3945/ajcn.115.122820.
174. Gholami, F., et al., *Healthy beverages may reduce the genetic risk of abdominal obesity and related metabolic comorbidities: a gene-diet interaction study in Iranian women.* Diabetology & Metabolic Syndrome, 2022. **14**(1): p. 143 DOI: 10.1186/s13098-022-00911-z.
175. Viljakainen, H., et al., *Interaction between genetic susceptibility to obesity and food intake on BMI in Finnish school-aged children.* Scientific Reports, 2023. **13**(1): p. 15265 DOI: 10.1038/s41598-023-42430-5.
176. Wang, T., et al., *Improving fruit and vegetable intake attenuates the genetic association with long-term weight gain.* Am J Clin Nutr, 2019. **110**(3): p. 759-768 DOI: 10.1093/ajcn/nqz136.
177. Qi, Q., et al., *Fried food consumption, genetic risk, and body mass index: gene-diet interaction analysis in three US cohort studies.* BMJ : British Medical Journal, 2014. **348**: p. g1610 DOI: 10.1136/bmj.g1610.
178. Goni, L., et al., *A genetic risk tool for obesity predisposition assessment and personalized nutrition implementation based on macronutrient intake.* Genes Nutr, 2015. **10**(1): p. 445 DOI: 10.1007/s12263-014-0445-z.
179. Lee, W.J., et al., *Analysis of the Interaction between Polygenic Risk Score and Calorie Intake in Obesity in the Korean Population.* Lifestyle Genomics, 2021. **14**(1): p. 20-29 DOI: 10.1159/000511333.
180. Casas-Agustench, P., et al., *Saturated fat intake modulates the association between an obesity genetic risk score and body mass index in two US populations.* Journal of the Academy of Nutrition and Dietetics, 2014. **114**(12): p. 1954-1966 DOI: 10.1016/j.jand.2014.03.014.
181. Freuer, D., C. Meisinger, and J. Linseisen, *Causal relationship between dietary macronutrient composition and anthropometric measures: A bidirectional two-sample Mendelian randomization analysis.* Clinical Nutrition, 2021. **40**(6): p. 4120-4131 DOI: 10.1016/j.clnu.2021.01.047.
182. Jeté, M. and L. Des Groseilliers, *Introduction à l'ÉLDEQ 1998–2002. In: L'enquête: description et méthodologie. Étude longitudinale du développement des enfants du Québec (ÉLDEQ 1998–2002).* Institut de la statistique du Québec, 2000. **1**(1), https://collections.banq.qc.ca/ark:/52327/bs26651.
183. Orri, M., et al., *Cohort Profile: Quebec Longitudinal Study of Child Development (QLSCD).* Social Psychiatry and Psychiatric Epidemiology, 2020 DOI: 10.1007/s00127-020-01972-z.
184. Shields, M., et al., *Bias in self-reported estimates of obesity in Canadian health surveys: an update on correction equations for adults.* Health Rep, 2011. **22**(3): p. 35-45, https://www150.statcan.gc.ca/n1/pub/82-003-x/2011003/article/11533-eng.pdf.
185. Shields, M., et al., *Obesity estimates for children based on parent-reported versus direct measures.* Health Rep, 2011. **22**(3): p. 47-58.
186. de Onis, M. and T. Lobstein, *Defining obesity risk status in the general childhood population: Which cut-offs should we use?* International Journal of Pediatric Obesity, 2010. **5**(6): p. 458-460 DOI: 10.3109/17477161003615583.
187. van Erp, S., D.L. Oberski, and J. Mulder, *Shrinkage priors for Bayesian penalized regression.* Journal of Mathematical Psychology, 2019. **89**: p. 31-50 DOI: https://doi.org/10.1016/j.jmp.2018.12.004.

188. Auton, A., et al., *A global reference for human genetic variation.* Nature, 2015. **526**(7571): p. 68-74 DOI: 10.1038/nature15393.
189. Delaneau, O., J.-F. Zagury, and J. Marchini, *Improved whole-chromosome phasing for disease and population genetic studies.* Nature Methods, 2013. **10**(1): p. 5-6 DOI: 10.1038/nmeth.2307.
190. Howie, B.N., P. Donnelly, and J. Marchini, *A Flexible and Accurate Genotype Imputation Method for the Next Generation of Genome-Wide Association Studies.* PLOS Genetics, 2009. **5**(6): p. e1000529 DOI: 10.1371/journal.pgen.1000529.
191. Peterson, R.E., et al., *Genome-wide Association Studies in Ancestrally Diverse Populations: Opportunities, Methods, Pitfalls, and Recommendations.* Cell, 2019. **179**(3): p. 589-603 DOI: 10.1016/j.cell.2019.08.051.
192. Schultz, L.M., et al., *Stability of polygenic scores across discovery genome-wide association studies.* HGG Adv, 2022. **3**(2): p. 100091 DOI: 10.1016/j.xhgg.2022.100091.
193. Majara, L., et al., *Low and differential polygenic score generalizability among African populations due largely to genetic diversity.* HGG Adv, 2023. **4**(2): p. 100184 DOI: 10.1016/j.xhgg.2023.100184.
194. Khera, A.V., et al., *Whole-Genome Sequencing to Characterize Monogenic and Polygenic Contributions in Patients Hospitalized With Early-Onset Myocardial Infarction.* Circulation, 2019. **139**(13): p. 1593-1602 DOI: 10.1161/circulationaha.118.035658.
195. Golding, et al., *ALSPAC–The Avon Longitudinal Study of Parents and Children.* Paediatric and Perinatal Epidemiology, 2001. **15**(1): p. 74-87 DOI: https://doi.org/10.1046/j.1365-3016.2001.00325.x.
196. Dubois, L., et al., *Genetic and environmental influences on eating behaviors in 2.5- and 9-year-old children: a longitudinal twin study.* The international journal of behavioral nutrition and physical activity, 2013. **10**: p. 134-134 DOI: 10.1186/1479-5868-10-134.
197. Dubois, L. and M. Girard, *Trends in dietary behaviours and practices.* Québec Longitudinal Study of Child Development–From Birth to 29 Months, 2002.
198. Dubois, L., et al., *Preschool children's eating behaviours are related to dietary adequacy and body weight.* European Journal of Clinical Nutrition, 2007. **61**(7): p. 846-855 DOI: 10.1038/sj.ejcn.1602586.
199. Health Canada, *Canadian Nutrient File*. 2001, https://food-nutrition.canada.ca/cnf-fce/?lang=eng.
200. Agricultural Research Service, *The USDA Food and Nutrient Database for Dietary Studies. Release 1.0*, F.S.R. Group, Editor. 2004: Beltsville (MD).
201. Sobel, M.E., *Asymptotic Confidence Intervals for Indirect Effects in Structural Equation Models.* Sociological Methodology, 1982. **13**: p. 290-312 DOI: 10.2307/270723.
202. MacKinnon, D.P., *Introduction to statistical mediation analysis*. Multivariate applications. 2008, New York: Lawrence Erlbaum Associates.
203. VanderWeele, T.J. and S. Vansteelandt, *Mediation Analysis with Multiple Mediators.* Epidemiologic methods, 2014. **2**(1): p. 95-115 DOI: 10.1515/em-2012-0010.
204. VanderWeele, T.J. and E.J. Tchetgen Tchetgen, *Mediation analysis with time varying exposures and mediators.* J R Stat Soc Series B Stat Methodol, 2017. **79**(3): p. 917-938 DOI: 10.1111/rssb.12194.
205. Sullivan, A., et al., *Longitudinal Mediation Analysis with Latent Growth Curves.* 2021 DOI: https://doi.org/10.48550/arXiv.2103.05765.
206. Hall, K.D., et al., *The energy balance model of obesity: beyond calories in, calories out.* Am J Clin Nutr, 2022. **115**(5): p. 1243-1254 DOI: 10.1093/ajcn/nqac031.
207. Ludwig, D.S., et al., *The carbohydrate-insulin model: a physiological perspective on the obesity pandemic.* Am J Clin Nutr, 2021. **114**(6): p. 1873-1885 DOI: 10.1093/ajcn/nqab270.

208. Weihrauch-Blüher, S., et al., *Current Guidelines for Obesity Prevention in Childhood and Adolescence.* Obesity Facts, 2018. **11**(3): p. 263-276 DOI: 10.1159/000486512.
209. Engel, G.L., *The Need for a New Medical Model: A Challenge for Biomedicine.* Science, 1977. **196**(4286): p. 129-136 DOI: doi:10.1126/science.847460.
210. Russell, C.G. and A. Russell, *A biopsychosocial approach to processes and pathways in the development of overweight and obesity in childhood: Insights from developmental theory and research.* Obesity Reviews, 2019. **20**(5): p. 725-749 DOI: https://doi.org/10.1111/obr.12838.
211. Herle, M., et al., *The Role of Eating Behaviours in Genetic Susceptibility to Obesity.* Current Obesity Reports, 2020. **9**(4): p. 512-521 DOI: 10.1007/s13679-020-00402-0.
212. Morys, F., et al., *Neuroanatomical correlates of genetic risk for obesity in children.* Transl Psychiatry, 2023. **13**(1): p. 1 DOI: 10.1038/s41398-022-02301-5.
213. Opel, N., et al., *Prefrontal gray matter volume mediates genetic risks for obesity.* Mol Psychiatry, 2017. **22**(5): p. 703-710 DOI: 10.1038/mp.2017.51.
214. Kate, E.M., et al., *Genetic risk of obesity as a modifier of associations between neighbourhood environment and body mass index: an observational study of 335 046 UK Biobank participants.* BMJ Nutrition, Prevention & Health, 2021. **3**(2): p. null DOI: 10.1136/bmjnph-2020-000107.
215. Vanderlee, L., et al., *Policies to Create Healthier Food Environments in Canada: Experts' Evaluation and Prioritized Actions Using the Healthy Food Environment Policy Index (Food-EPI).* Int J Environ Res Public Health, 2019. **16**(22) DOI: 10.3390/ijerph16224473.
216. Jansen, P.R., et al., *The utility of obesity polygenic risk scores from research to clinical practice: A review.* Obesity Reviews, 2024. **25**(11): p. e13810 DOI: https://doi.org/10.1111/obr.13810.
217. Puckett, E.E., *Variability in total project and per sample genotyping costs under varying study designs including with microsatellites or SNPs to answer conservation genetic questions.* Conservation Genetics Resources, 2017. **9**(2): p. 289-304 DOI: 10.1007/s12686-016-0643-7.
218. Wald, N.J. and R. Old, *The illusion of polygenic disease risk prediction.* Genetics in Medicine, 2019. **21**(8): p. 1705-1707 DOI: 10.1038/s41436-018-0418-5.
219. Wray, N.R., et al., *From Basic Science to Clinical Application of Polygenic Risk Scores: A Primer.* JAMA Psychiatry, 2021. **78**(1): p. 101-109 DOI: 10.1001/jamapsychiatry.2020.3049 %J JAMA Psychiatry.
220. Walker, M.J., et al., *Implementing Multifactorial Risk Assessment with Polygenic Risk Scores for Personalized Breast Cancer Screening in the Population Setting: Challenges and Opportunities.* Cancers (Basel), 2024. **16**(11) DOI: 10.3390/cancers16112116.
221. Lewis, A.C.F., et al., *Patient and provider perspectives on polygenic risk scores: implications for clinical reporting and utilization.* Genome Medicine, 2022. **14**(1): p. 114 DOI: 10.1186/s13073-022-01117-8.
222. Frieser, M.J., S. Wilson, and S. Vrieze, *Behavioral impact of return of genetic test results for complex disease: Systematic review and meta-analysis.* Health Psychol, 2018. **37**(12): p. 1134-1144 DOI: 10.1037/hea0000683.
223. O'Donovan, C.B., et al., *Knowing your genes: does this impact behaviour change?* Proceedings of the Nutrition Society, 2017. **76**(3): p. 182-191 DOI: 10.1017/S0029665116002949.
224. Persky, S., et al., *Investigating the Efficacy of Genetic, Environmental, and Multifactorial Risk Information When Communicating Obesity Risk to Parents of Young Children.* Annals of Behavioral Medicine, 2021. **55**(8): p. 720-733 DOI: 10.1093/abm/kaaa103.
225. Mega, J.L., et al., *Genetic risk, coronary heart disease events, and the clinical benefit of statin therapy: an analysis of primary and secondary prevention trials.* Lancet, 2015. **385**(9984): p. 2264-2271 DOI: 10.1016/s0140-6736(14)61730-x.

226. Dawed, A.Y., et al., *Pharmacogenomics of GLP-1 receptor agonists: a genome-wide analysis of observational data and large randomised controlled trials.* The Lancet Diabetes & Endocrinology, 2023. **11**(1): p. 33-41 DOI: 10.1016/S2213-8587(22)00340-0.
227. Jager, K.J., et al., *Confounding: What it is and how to deal with it.* Kidney International, 2008. **73**(3): p. 256-260 DOI: https://doi.org/10.1038/sj.ki.5002650.
228. Hellwege, J.N., et al., *Population Stratification in Genetic Association Studies.* Current protocols in human genetics, 2017. **95**: p. 1.22.1-1.22.23 DOI: 10.1002/cphg.48.
229. Bherer, C., et al., *Admixed ancestry and stratification of Quebec regional populations.* American Journal of Physical Anthropology, 2011. **144**(3): p. 432-441 DOI: https://doi.org/10.1002/ajpa.21424.
230. Rao, D.P., et al., *Childhood overweight and obesity trends in Canada.* Health Promot Chronic Dis Prev Can, 2016. **36**(9): p. 194-8 DOI: 10.24095/hpcdp.36.9.03.
231. Mattsson, M., et al., *Eating behaviour, physical activity, TV exposure and sleeping habits in five year olds: a latent class analysis.* BMC Pediatr, 2021. **21**(1): p. 180 DOI: 10.1186/s12887-021-02640-0.
232. Porta, M., *Selection Bias*. 2016, Oxford University Press DOI: 10.1093/acref/9780199976720.013.1711.
233. Munafò, M.R., et al., *Collider scope: when selection bias can substantially influence observed associations.* International Journal of Epidemiology, 2018. **47**(1): p. 226-235 DOI: 10.1093/ije/dyx206.
234. Avinun, R. and A.R. Hariri, *A polygenic score for body mass index is associated with depressive symptoms via early life stress: Evidence for gene-environment correlation.* Journal of Psychiatric Research, 2019. **118**: p. 9-13 DOI: https://doi.org/10.1016/j.jpsychires.2019.08.008.
235. Taylor, A.E., et al., *Exploring the association of genetic factors with participation in the Avon Longitudinal Study of Parents and Children.* International Journal of Epidemiology, 2018. **47**(4): p. 1207-1216 DOI: 10.1093/ije/dyy060.
236. Börnhorst, C., et al., *Prevalence and determinants of misreporting among European children in proxy-reported 24 h dietary recalls.* British Journal of Nutrition, 2013. **109**(7): p. 1257-1265 DOI: 10.1017/S0007114512003194.
237. Laberge, A.M., et al., *Population history and its impact on medical genetics in Quebec.* Clinical Genetics, 2005. **68**(4): p. 287-301 DOI: https://doi.org/10.1111/j.1399-0004.2005.00497.x.
238. Moreau, C., H. Vézina, and D. Labuda, *[Founder effects and genetic variability in Quebec].* Med Sci (Paris), 2007. **23**(11): p. 1008-13 DOI: 10.1051/medsci/200723111008.
239. Statistique Canada. *Indice de masse corporelle des enfants et des jeunes, 2012 à 2013*. 2014 2015 [cited 23 novembre 2020; Available from: https://www150.statcan.gc.ca/n1/pub/82-625-x/2014001/article/14105-fra.htm.

APPENDICES

Appendix 1. Copy of the ethics certificate of approval.

10/02/2023

Université d'Ottawa
Bureau d'éthique et d'intégrité de la recherche

University of Ottawa
Office of Research Ethics and Integrity

CERTIFICAT D'APPROBATION ÉTHIQUE | CERTIFICATE OF ETHICS APPROVAL

Numéro du dossier / Ethics File Number	H-01-23-8018
Titre du projet / Project Title	Role of eating habits in the genetic susceptibility to obesity in children and adolescents
Type de projet / Project Type	Thèse de doctorat / Doctoral thesis
Statut du projet / Project Status	Approuvé / Approved
Date d'approbation (jj/mm/aaaa) / Approval Date (dd/mm/yyyy)	10/02/2023
Date d'expiration (jj/mm/aaaa) / Expiry Date (dd/mm/yyyy)	09/02/2024

Équipe de recherche / Research Team

Chercheur / Researcher	Affiliation	Role
Danick GOULET	Département d'épidémiologie et santé publique / Department of Epidemiology and Public Health	Chercheur Principal / Principal Investigator
Lise DUBOIS	Département d'épidémiologie et santé publique / Department of Epidemiology and Public Health	Superviseur / Supervisor
Michel BOIVIN	Université Laval	Co-superviseur / Co-supervisor

Conditions spéciales ou commentaires / Special conditions or comments

550, rue Cumberland, pièce 154 Ottawa (Ontario) K1N 6N5 Canada
550 Cumberland Street, Room 154 Ottawa, Ontario K1N 6N5 Canada

613-562-5387 • 613-562-5338 • ethique@uOttawa.ca / ethics@uOttawa.ca
www.recherche.uottawa.ca/deontologie | www.recherche.uottawa.ca/ethics

10/02/2023

Université d'Ottawa
Bureau d'éthique et d'intégrité de la recherche

University of Ottawa
Office of Research Ethics and Integrity

Le Comité d'éthique de la recherche (CÉR) de l'Université d'Ottawa, opérant conformément à l'*Énoncé de politique des Trois conseils* (2014) et toutes autres lois et tous règlements applicables, a examiné et approuvé la demande d'éthique du projet de recherche ci-nommé.

L'approbation est valide pour la durée indiquée plus haut et est sujette aux conditions énumérées dans la section intitulée "Conditions Spéciales ou Commentaires". Le formulaire « Renouvellement ou Fermeture de Projet » doit être complété quatre semaines avant la date d'échéance indiquée ci-haut afin de demander un renouvellement de cette approbation éthique ou afin de fermer le dossier.

Toutes modifications apportées au projet doivent être approuvées par le CÉR avant leur mise en place, sauf si le participant doit être retiré en raison d'un danger immédiat ou s'il s'agit d'un changement ayant trait à des éléments administratifs ou logistiques du projet. Les chercheurs doivent aviser le CÉR dans les plus brefs délais de tout changement pouvant augmenter le niveau de risque aux participants ou pouvant affecter considérablement le déroulement du projet, rapporter tout événement imprévu ou indésirable et soumettre toute nouvelle information pouvant nuire à la conduite du projet ou à la sécurité des participants.

The University of Ottawa Research Ethics Board, which operates in accordance with the *Tri-Council Policy Statement* (2014) and other applicable laws and regulations, has examined and approved the ethics application for the above-named research project.

Ethics approval is valid for the period indicated above and is subject to the conditions listed in the section entitled "Special Conditions or Comments". The "Renewal/Project Closure" form must be completedfour weeks before the above-referenced expiry date to request a renewal of this ethics approval or closure of the file.

Any changes made to the project must be approved by the REB before being implemented, except when necessary to remove participants from immediate endangerment or when the modification(s) only pertain to administrative or logistical components of the project. Investigators must also promptly alert the REB of any changes that increase the risk to participant(s), any changes that considerably affect the conduct of the project, all unanticipated and harmful events that occur, and new information that may negatively affect the conduct of the project or the safety of the participant(s).

Germain ZONGO
Responsable d'éthique en recherche / Protocol Officer
Pour/For **Daniel LAGAREC** Président(e) du/ Chair of the **Comité d'éthique de la recherche en sciences de la santé et sciences / Health Sciences and Sciences Research Ethics Board**

550, rue Cumberland, pièce 154 550 Cumberland Street, Room 154
Ottawa (Ontario) K1N 6N5 Canada Ottawa, Ontario K1N 6N5 Canada

613-562-5387 • 613-562-5338 • ethique@uOttawa.ca / ethics@uOttawa.ca
www.recherche.uottawa.ca/deontologie | www.recherche.uottawa.ca/ethics

Appendix 2. Characteristics of QLSCD participants with BMI data available at seven waves of data collection.

Characteristic	All N=2120	4 years N=1529	6 years N=1170	7 years N=1482	8 years N=1346	10 years N=1323	12 years N=1349	13 years N=1229
Age, years	-	4.16 ± 0.26	6.24 ± 0.25	7.16 ± 0.25	8.18 ± 0.26	10.17 ± 0.26	12.14 ± 0.26	13.14 ± 0.26
Sex, boys	1080 (50.9)	773 (50.6)	553 (47.3)	704 (47.5)	643 (47.8)	628 (47.5)	649 (48.1)	568 (46.2)
Birth weight, kg	3.40 ± 0.50	3.40 ± 0.50	3.40 ± 0.51	3.41 (0.50)	3.40 ± 0.50	3.41 ± 0.50	3.41 ± 0.49	3.42 ± 0.49
Preterm birth	99 (4.7)	75 (4.9)	58 (5.0)	70 (4.7)	61 (4.5)	57 (4.3)	55 (4.1)	51 (4.1)
Maternal age at birth								
≤ 20 years	59 (2.8)	37 (2.4)	29 (2.5)	45 (3.0)	35 (2.6)	32 (2.4)	35 (2.6)	30 (2.4)
20-34 years	1775 (83.8)	1283 (83.9)	973 (83.2)	1230 (83.0)	1114 (82.8)	1107 (83.7)	1128 (83.6)	1026 (83.5)
≥ 35 years	285 (13.4)	209 (13.7)	168 (14.4)	207 (14.0)	197 (14.6)	184 (13.9)	186 (13.8)	173 (14.1)
Post-sec. education	1177 (55.6)	887 (58.0)	682 (58.3)	866 (58.5)	786 (58.4)	782 (59.2)	785 (58.2)	730 (59.4)
Household income								
< 30 000$	622 (29.9)	409 (27.0)	304 (26.2)	401 (27.4)	353 (26.6)	341 (26.1)	359 (26.9)	313 (25.8)
30 000 – <60 000$	843 (40.5)	636 (42.0)	478 (41.2)	600 (41.0)	552 (41.6)	550 (42.1)	548 (41.1)	505 (41.6)
60 000 – <80 000$	327 (15.7)	246 (16.3)	197 (17.0)	234 (16.0)	220 (16.6)	213 (16.3)	218 (16.4)	201 (16.5)
≥ 80 000$	290 (13.9)	222 (14.7)	182 (15.7)	230 (15.7)	203 (15.3)	203 (15.5)	208 (15.6)	196 (16.1)
Mother born in Canada	1867 (88.1)	1393 (91.1)	1093 (93.4)	1356 (91.5)	1233 (91.6)	1220 (92.2)	1241 (92.0)	1132 (92.1)
Weight category								
Healthy weight	-	1425 (93.2)	937 (80.1)	1183 (79.8)	1003 (74.5)	884 (66.8)	889 (65.9)	805 (65.5)
Overweight	-	72 (4.7)	158 (13.5)	203 (13.7)	220 (16.3)	268 (20.3)	286 (21.2)	271 (22.1)
Obesity	-	32 (2.1)	75 (6.4)	96 (6.5)	123 (9.1)	171 (12.9)	174 (12.9)	153 (12.4)
BMI, kg/m^2	-	15.76 ± 1.80	15.80 ± 2.05	16.06 ± 2.29	16.72 ± 2.58	18.42 ± 3.28	20.07 ± 3.96	20.93 ± 4.14
BMI z-score, SD	-	0.25 ± 1.20	0.17 ± 1.15	0.18 ± 1.18	0.31 ± 1.25	0.57 ± 1.15	0.57 ± 1.19	0.56 ± 1.18

Continuous variables presented as mean ± standard deviation, and categorial variables presented as n (%). Sex, birth weight, preterm birth, maternal age at birth, maternal education, household income, and mother's immigration status were variables measured at the first wave of data collection (5 months old). All other variables (age, BMI, BMI z-scores, and weight category) were gathered at the respective data collection wave.

Appendix 3. Proof of submission email for manuscript 1, Canadian Journal of Physiology and Pharmacology.

Canadian Journal of Physiology and Pharmacology

Submission Confirmation

Print

Thank you for your revision

Submitted to
Canadian Journal of Physiology and Pharmacology

Manuscript ID
cjpp-2024-0221.R1

Title
Polygenic scores of obesity in childhood based on summary statistics from adults vs. children

Authors
Goulet, Danick
Boivin, Michel
Gravel, Christopher
Little, Julian
Ouellet-Morin, Isabelle
Gouin, Jean-Philippe
Dubois, Lise

Date Submitted
11-Dec-2024

Appendix 4: PRS-RS reporting checklist for Manuscript 1.

Adapted from the original checklist from: Wand, H., Lambert, S.A., Tamburro, C. *et al.* Improving reporting standards for polygenic scores in risk prediction studies. *Nature* **591**, 211–219 (2021). https://doi-org.proxy.bib.uottawa.ca/10.1038/s41586-021-03243-6.

Reporting standard		Description	Page No.	Additional notes
Background	Study type	Specify whether the study aims to develop and/or validate a PRS. When externally validating or combining previously published PRSs or integrated risk models, include identifier(s) of original PRS (PMID, PGS Catalog ID).	Introduction p.56	Developing two PGS
	Risk model purpose and predicted outcome	Specify what the risk model is intended to predict and the purpose. This includes intended use (risk prediction, diagnostic, prognostic, or therapeutic modalities), predicted outcome (if a clinical feature or endpoint within a specific disease) and the current models available for that outcome.	Introduction p.56	Prediction: BMI z-score and obesity.
Study population and data	Study design and recruitment	For each of the datasets describe the study design (for example, cohort, case–control, cross-sectional), eligibility criteria, recruitment period and setting (for example, method and years) and follow-up. State whether the data are primary or secondary data. If secondary analysis, include the full reference to the original study.	Materials and methods (Study design and participants) p.56	Secondary analysis of birth-cohort study.
	Participant demographics and clinical characteristics	Include the distribution of demographic information in each dataset (and the combined total if relevant) used to generate a single risk model (whether a single sample set, or the summary of combined samples) including the mean, standard deviation and range. This should at minimum include age, sex and any other characteristics relevant to describe the study population or the performance of the model. Provide demographics stratified by case–control status, if applicable.	Results p.59-60, Table 4-1	-
	Ancestry	Include the ancestral background distribution of each sample population used during PRS development and validation	Materials and methods	GWAS summary statistics from European

		(including those from any GWAS summary statistics that were included), and the data source of this ancestry information (for example, self-report, genotyping). Ancestry information should be reported using the standardized framework developed by the NHGRI-EBI GWAS Catalog with detailed information beyond this when available. When combining samples from multiple studies, aggregate ancestral distribution information is sufficient. The method of ancestry inference should be provided.	(Genotyping and polygenic score construction) p.57, Table 4-1	ancestry samples. Majority of QLSCD target sample had mothers born in Canada.
	Genetic data	Provide the method for acquiring genetic information (for example, sequencing, genotyping) in each sample, including information about genome build and technical assay details. If imputed, specify the imputation panel and give ancestry information. Report any relevant quality control, including imputation quality filters to exclude low-quality imputed SNPs. If parameters were selected from another study, include reference (PMID, GWAS Catalog ID).	Materials and methods (Genotyping and polygenic score construction) p.5-57	-
	Non-genetic variables	Define any non-genetic variables that were included in the risk model, provide variable definitions and measurement (for example, assay, ICD codes, e-phenotyping algorithms, chart review, self-report). Indicate the scale of each variable, for example, dichotomous, continuous, categorical or ordinal. Explicitly state which variables are included in the final model.	Materials and methods (Statistical analysis) p.59	Covariate-adjusted model described.
	Outcome of interest	Define the predicted outcome(s) of interest and report distribution. If the predicted outcome is a clinical feature or end-point within a specific disease, provide the criteria used to define that disease membership. Include details on how information was ascertained (for example, ICD codes, e-phenotyping algorithms, chart review, self-report). Transformation of continuous data into binary, ordinal, or categorical outcomes should be detailed with justification.	Materials and methods (Anthropometric measures) p.58	BMI z-score (continuous) and obesity (category).

		State whether the predicted phenotype of the polygenic score is the same as or different from the predicted outcome of the risk model. Provide justification for differences, if applicable.		
	Missing data	State explicitly how missing data were handled for all variables included in the model. If imputation was used, include detailed of the approach used and any subsequent filtering or post-processing.	-	Complete case analysis.
Risk model development and application	PRS construction and estimation	Describe how genetic data were included in the PRS. Authors should detail criteria used to determine inclusion in the model for all variants. Define how the variants were selected, weighted and combined into a single score. If the PRS was derived from another study include the reference (PMID, PGS Catalog Score ID).	Materials and methods (Genotyping and polygenic score construction) p.57	PRS-CS method.
	Risk model type	Detail statistical methods used to estimate risk, either relative or absolute, from the continuous risk score distribution. Detail whether risk is cumulative or cross-sectional, with appropriate comparison groups if relative risk is presented. Report time until predicted risk (for example, 5 years, 10 years, lifetime). In an absolute risk model, state the time until the predicted event and the prevalence or incidence of the predicted outcome in the general population.	Materials and methods (Statistical analysis) p.58-59	-
	Integrated risk model(s) description and fitting	State the procedure used to develop the risk models that includes non-genetic and/or genetic variables other than the PRS. If the model(s) was selected for optimal performance, describe measures used to assess performance. Explicitly state all variables used in each risk model.	Materials and methods (Statistical analysis) p.59	Covariate-adjusted model described as supplementary analysis.
Risk model evaluation	PRS distribution	Include a general description of the distribution of the PRS. This details the continuous distribution output directly from the risk score calculation.	Results p.60-61, Figures 4-2 and 4-3	PGS were standardized. Description based on quintiles.
	Risk model predictive ability	Describe and report metrics of overall performance (proportion of variance explained; R^2) and estimates of risk (such as odds or hazards ratios from regression models) used	Results p.60-61, Figure 4-1, Figure 4-4	β, proportion of variance explained, and OR reported.

		to evaluate the PRS and/or risk models. Describe the set of genetic and non-genetic variables included in the analysis.		
	Risk model discrimination	Describe and report metrics (such as AUROC, AUPRC, and—for survival models—the C-index) used to assess the discrimination of evaluated risk models and whether any non-genetic variables were included beyond a PRS in this analysis. Evaluation of the potential clinical utility of models requires evaluating tail-based measures, such as proportions of populations and cases that exceed specified clinically relevant risk thresholds and measures of reclassifications (for example, NRI) at such thresholds for comparison of models.	Results p.61, Figure 4-4, Supplementary Figure 4-1	AUROC reported.
	Risk model calibration	Describe and report metrics used to assess the calibration of evaluated risk scores and models. Describe the set of genetic and non-genetic variables included in the analysis.	Results p.61, Figure 4-4, Supplementary Figure 4-2	Briar score and calibration plot reported.
	Subgroup analyses	Subgroup size, demographics and clinical characteristics should be given. Relevant evaluation methods and measures (distribution, predictive ability, discrimination and calibration) should be described for each subgroup analysis.	Results p.62, Supplementary Tables 4-4 and 4-5	Sex subgroup analysis reported.
Limitations and clinical implications	Risk model interpretation	Summarize the risk models in terms of what they predict, how well and in whom. Explicitly mention the incremental performance of the PRS and/or combined risk model in comparison to conventional risk models, as well as the performance of the PRS and risk model alone. Conventional risk models might include demographic (age, sex), disease-specific risk factors and/or family history of disease.	Discussion p.63	Comparison with other PGS predicting BMI and risk of obesity in children.
	Limitations	Outline limitations of the study with relevance to the results, discuss the effects of these limitations on the interpretation of the risk model and any downstream replication efforts needed. Common considerations include: study design restrictions, use of a surrogate outcome, ascertainment biases, the distribution of participant-level traits (ancestry, age, comorbidities), accuracy or specificity of outcome data, and	Discussion p.66	-

		any statistical considerations. Note and discuss the effects of any unknown reporting items from previous sections.		
	Generalizability	Discuss the intended target groups or populations this score may be applied to and explicitly address any issues with generalizability beyond the included populations. Discuss whether the study externally validates the score and/or model, or if the sample is limited with respect to ancestry, age or other variables.	Discussion p.66	Generalizability discussed in the limits section.
	Risk model intended uses	Discuss whether there is an intended clinical use or utility to the risk model. If so, discuss the 'clinic readiness' and next steps with respect to the interpretation, limitations and generalizability of the model. Discuss how the predictive ability of the model compares with current standards of care or other published work (such as existing PRSs) on predicting the outcome of interest.	Discussion p.63-64	Not intended for clinical use. Comparison between adult-derived and child-derived PGS is the intended use.
Data transparency and availability		Information sufficient to calculate the PRS and the risk model(s) on external samples should be made freely available. For genetic variables this would include information about the variants (for example, rsID, chromosomal location, effect allele and the effect weight) that comprise the score; PRSs with this information should be deposited in the PGS Catalog for findability and to promote reuse and comparison with other established scores. Weights for non-genetic variables should also be provided to make the risk model calculable.	Data availability statement	Not submitted to PGS Catalog since the article not published yet.

Notes: ICD - International Classification of Diseases; PMID - PubMed ID; rsID - reference SNP cluster ID; SNP - single-nucleotide polymorphism. PRS (polygenic risk score) is used instead of PGS throughout the checklist. The original acronym was kept.

Appendix 5: Published manuscript 2

Received: 25 June 2024 | Revised: 9 September 2024 | Accepted: 13 September 2024
DOI: 10.1111/ijpo.13180

ORIGINAL RESEARCH

Pediatric OBESITY WILEY

Mediation of genetic susceptibility to obesity through eating behaviours in children

Danick Goulet[1] | Michel Boivin[2] | Christopher A. Gravel[1,3,4] | Julian Little[1] | Beth K. Potter[1] | Lise Dubois[1]

[1]School of Epidemiology and Public Health, University of Ottawa, Ottawa, Ontario, Canada
[2]Université Laval, Quebec, Quebec, Canada
[3]Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Quebec, Canada
[4]Department of Mathematics and Statistics, University of Ottawa, Ottawa, Ontario, Canada

Correspondence
Lise Dubois, School of Epidemiology and Public Health, University of Ottawa, 600 Peter Morand Crescent, Ottawa, ON, Canada.
Email: ldubois@uottawa.ca

Michel Boivin, Université Laval, Quebec, QC, Canada.
Email: michel.boivin@psy.ulaval.ca

Funding information
Canadian Institutes of Health Research, Grant/Award Number: 165964

Summary

Background/Objectives: Few studies have examined the putative mediating role of eating behaviours linking genetic susceptibility and body weight. The goal of this study was to investigate the extent to which two polygenic scores (PGSs) for body mass index (BMI), based on child and adult data, predicted BMI through over-eating and fussy eating across childhood.

Subjects/Methods: The study sample involved 692 participants from a birth cohort study. Height and weight were measured on six occasions between ages 6 and 13 years. Over-eating and fussy eating behaviours were assessed five times between ages 2 and 6 years. Longitudinal growth curve mediation analysis was used to estimate the contributions of the PGSs to BMI *z*-scores mediated by over-eating and fussy eating.

Results: Both PGSs predicted BMI *z*-scores (PGS_{child}: $\beta = 0.26$, 95% CI: 0.19–0.33; PGS_{adult}: $\beta = 0.34$, 95% CI: 0.27–0.41). Over-eating significantly mediated these associations, but this mediation decreased over time from 6 years (PGS_{child}: 18.0%, 95% CI: 3.1–32.9, p-value = 0.018; PGS_{adult}: 14.2%, 95% CI: 2.8–25.5, p-value = 0.014) to 13 years (PGS_{child}: 11.4%, 95% CI: −0.4-23.1, p-value = 0.057; PGS_{adult}: 6.2%, 95% CI: 0.4–12.0, p-value = 0.037). Fussy eating did not show any mediation.

Conclusions: Our results support the view that appetite is key to translating genetic susceptibility into changes in body weight.

KEYWORDS
children's health, eating behaviours, longitudinal analysis, mediation, obesity, polygenic score

Abbreviations: BMI, body mass index; GWAS, Genome-Wide Association Studies; IPW, inverse probability weighting; LGCMA, longitudinal growth curve mediation analysis; PGS, polygenic score; QLSCD, Quebec Longitudinal Study of Children Development.

1 | INTRODUCTION

Since the World Health Organization declared obesity as a global epidemic in 1997,[1] the prevalence of obesity has continued to rise worldwide. Most worrisome is the rise of obesity in childhood. Children with overweight or obesity between the ages of 5 and 11 years accounted for more than a quarter of Canadian children between 2012 and 2013.[2] This is especially problematic because childhood obesity tends to persist into adulthood. Compared to children of a healthy weight, those with obesity are estimated to have a fivefold increase in risk of presenting with obesity in adulthood.[3] Accordingly, understanding the factors underlying obesity in childhood is a public health priority.

A substantial portion of the observed variability in individual body weight can be accounted for by variation in genetic susceptibility to weight gain. According to twin studies, genetic factors are responsible for 40% to 80% of the variance in body mass index (BMI) in populations of European ancestry.[4] This strong genetic underpinning of BMI can already be observed in early childhood. Dozens of genetic variants have been found to be associated with childhood BMI, each with small independent contributions.[5,6] Combining those individual genetic risks results in a significant association with BMI in childhood.[5]

According to the behavioural susceptibility theory, the genes involved in shaping obesity phenotypes exert their influence mainly through neurological processes regulating appetite, which are translated through various eating behaviours.[7] This is consistent with the observation that many genetic variants associated with BMI are involved in brain regions related to behavioural control.[8] Large-scale genome-wide association studies (GWAS) have revealed that variants associated with BMI are predominantly expressed in the central nervous system in regions involved in cognition and emotion regulation (e.g. hippocampus and limbic system) and in areas responsible for appetite regulation (e.g. hypothalamus and pituitary gland).[9,10] Empirical data support the view that appetite has a strong role in obesity and that appetite is characterized by a strong genetic component as shown by restrained, uncontrolled and emotional eating having moderate heritability.[8] However, due to the scarcity of genetically informed studies in childhood, there is limited empirical evidence on the possible role of eating behaviours as mediating genetic susceptibility to weight gain.

Observational studies have leveraged genotype data to examine the degree to which eating behaviours (e.g. satiety responsiveness, appetite, uncontrolled eating, emotional eating, susceptibility to hunger), usually assessed by questionnaires, mediate the association between genetic susceptibility, assessed through polygenic scores (PGSs) and BMI in adults. For instance, uncontrolled eating and emotional eating partly mediated the association between three different PGSs and BMI in French, UK and Finnish cohorts.[11,12] Additionally, disinhibition mediated the association between a PGS and BMI in French Canadian adults[13] and UK adults.[14] Although investigated in all three cohorts, susceptibility to hunger was only identified as a mediator in the French Canadian study.[13] These studies used PGSs comprising <100 variants. Additionally, another study observed that the association between a PGS comprised of close to 1 million variants and BMI was mediated through infrequent and unhealthy eating, emotional and external eating and snacking.[15]

Three studies have examined the role of behavioural traits related to appetite in the genetic susceptibility of BMI in children. These studies used the Children's Eating Behaviour Questionnaire, one of the most commonly used questionnaires assessing behavioural traits related to appetite.[7] The questionnaire measures eight behaviours related to appetite that can be categorized into food approach (food responsiveness, enjoyment of food, emotional over-eating and desire to drink) and food avoidance (satiety responsiveness, food fussiness, emotional undereating and slowness in eating).[16] Satiety responsiveness was found to mediate the association between a 28-variant PGS and BMI in 2258 children from the UK (mean age ~ 10 years old).[17] In contrast, satiety responsiveness did not mediate the effect of the 32-variant PGS on BMI in younger (4- to 8-year-old) Norwegian children.[18] Food responsiveness, emotional over-eating, enjoyment of food and slowness in eating were also not found to be mediators in the Norwegian study.[18] In the third study, high appetite at 2 years mediated the association between a 16-variant PGS and BMI in French Canadian children aged 2 to 5 years, where the proportion mediated decreased from 47% at 2 years to 24% at 5 years.[19] These studies have the same limitations as the mediation analyses performed on adults; they were mostly cross-sectional and used PGS comprising a small number of variants. Cross-sectional designs are problematic in mediation analyses. Ideally, these mediation studies should be longitudinal to ensure the observed eating behaviours precede changes in body weight. More recent pan-genomic approaches to the construction of PGS include hundreds of thousands of variants yielding stronger BMI prediction.[20]

In summary, previous studies suggest that eating behaviours could play a role in translating genetic susceptibility to higher BMI, but the evidence should be improved by examining this mediation (1) in childhood and adolescence for which the evidence is scarce and mixed, (2) using more recent and state-of-the-art PGS that include a large number of variants to improve prediction and (3) through longitudinal designs to assess the directionality of associations. Moreover, the bulk of evidence has been based on PGSs derived from adult anthropometric data, rather than data from children. Considering the genetic susceptibility to body weight variation changes throughout life,[8] the identity and effect of specific genetic variants may also vary over the life course. Using PGSs derived from both adults and children is likely to yield more insights than using a single score. Therefore, the goal of the present study was to examine (1) the prediction of school-aged child BMI by two BMI-PGSs, respectively derived from child and adult GWAS summary statistics, and (2) the extent to which this prediction is mediated through eating behavioural traits (over-eating, fussy eating) assessed in early childhood (i.e. preschool age).

2 | METHODS

2.1 | Study design and participants

We used data from the Quebec Longitudinal Study of Child Development (QLSCD), a longitudinal birth cohort study from Quebec,

Canada. The study recruited 2120 children born in 1997–1998 by a region-based stratified random selection from the Master birth register to ensure the sample was representative of the province's birth population. The children and their families were assessed annually or biannually starting at 5 months, and follow-up is ongoing. A variety of information was gathered on the children (e.g. behaviour, mental/physical health, diet, height, and weight assessments) and their families using questionnaires and interviews with the parents, teachers and children themselves, as well as through direct assessments. More comprehensive details about the QLSCD, including exclusion criteria, attrition at each data collection, main findings or a more exhaustive look at the data gathered, are available in a published cohort profile.[21] Biological samples were collected from a subset of QLSCD participants when the children were 10 years of age, and DNA was extracted. These DNA samples were later genotyped using the Illumina Infinium PsychArray-24 (see below). The present study included extensive BMI and eating behaviour assessments which were collected during follow-up. Informed consent was obtained from the parents at each step of the data collection until age 10, along with the children's permission. Starting at age 10, informed consent was obtained directly from the participants.

2.2 | Anthropometric measures

Height and weight data were longitudinally collected through different means depending on the wave of data collection. Anthropometric data was reported by the parent, directly measured or reported by the participant. In this study, we focus on the data measured directly. Specifically, trained research assistants measured height and weight on six occasions when participants were aged 6, 7, 8, 10, 12 and 13 years, following a standardized protocol, with the use of a measuring tape, ruler and scale. Two to three assessments were performed, and an average of the two closest values was computed as the final measure. BMI z-scores were then calculated using the World Health Organization Growth reference data,[22] which considers both the sex and age of children using the R package 'childsds'.[23] Considering the low prevalence of obesity in our sample, we decided against separating our sample in weight categories and proceeded with the BMI z-score variable as the outcome of interest.

2.3 | Eating behaviours

Two eating behaviour scores (fussy eating and over-eating) were considered as the putative mediators. The scores were based on five eating behavioural traits assessed five times from age 2 to 6 years. The individual items were translated to French from those used in the Avon Longitudinal Study of Parents and Children.[24] Questions were modified when necessary based on recommendations from an expert advisory group and a pretest conducted in a sample of parents that were not part of the QLSCD.[25] For more details, see a previous publication using eating behaviour data from the QLSCD[26] and the QLSCD website (www.jesuisjeseraistat.gouv.qc.ca). The most knowledgeable person about the child (the mother in most cases) answered items including (1) *When [name of the child] is at home with you for the main meal of the day, how often does [name of the child]* (1) *eat different meals…* (2) *…is [name of the child] fussy about food?* (3) *…refuse to eat?* (4) *… over-eat?* (5) *…eat too fast?* Answers were graded on a four-point *Likert scale* from: 'Almost never (1)' to 'Always (4)', for the first question, and 'Never (1 point), Rarely (2 points), Sometimes (3 points), Often (4 points)' for the last four questions. The sum of 'eating different meals', 'fussy about food' and 'refusing to eat' was calculated at each time point for the 'fussy eating' score, and the sum of 'the items 'over-eating' and 'eating too fast' was used to compute the 'over-eating' score. The final two scores were obtained by averaging the sum at each time point with complete data per behaviour available per participant to obtain a mean score between 2 and 6 years. The internal consistency of the eating behaviour scores was satisfactory, with Cronbach's alpha of 0.82 for fussy eating and 0.74 for overeating.

2.4 | Genotyping and polygenic score construction

DNA collection was proposed to a subsample of 1334 QLSCD participants remaining in the study at the 10-year-old data collection timepoint. A total of 992 participants agreed to DNA sampling and genotyping, which was performed using the Illumina Infinium PsychArray-24. Quality control of genetic data was performed as described in an earlier publication.[27] Variants with minor allele frequency below 0.01 and genotyping rate below 0.98 were excluded. Individuals with a call rate below 0.95 and those flagged for sex mismatch were excluded. Additional checks were also implemented for genetic duplicates, cryptic relatives, genetic outliers and heterozygosity deviations. Variants in the remaining individuals were then re-evaluated to exclude non-autosomal variants, those with call rates below 98%, those with a minor allele frequency below 5% and Hardy–Weinberg Equilibrium test p-values below 1×10^{-3}. Imputation was performed using the 1000 genomes project reference data[28] and programs SHAPEIT v2 (r837)[29] and impute2.[30] Finally, additional quality control was performed post-imputation where variants with a minor allele frequency below 1%, a Hardy–Weinberg Equilibrium test p-values below 1×10^{-6} and an INFO metric below 0.8 were removed. After completing all quality control and imputation steps, 816 participants had genetic data, including 8 465 216 variants.

The PGSs were derived from the GWAS summary statistics obtained from prior studies performed on ~700 000 adults[10] and ~60 000 children of European ancestry[5] using the PRS-CS method[20] and the 1000 Genomes reference panel[28] as the external linkage disequilibrium reference panel. The global shrinkage parameter was set to 0.01 since BMI is a polygenic trait. Both PGSs were then regressed on the first 10 principal components of genetic ancestry to account for population stratification, with the subsequent residuals used to obtain adjusted scores, following the method described by Khera et al.[31] (R script in Text S1). The final PGSs included 689 789 (child-based) and 613 732 (adult-based) individual variants. A total of

715 participants had genetic, eating behaviour and anthropometric data available and passed genotypic quality control processes. An additional 23 participants did not have all covariate data available, leaving 692 participants as the sample size for the main mediation analysis. A description of how the participants were parsed from the initial QLSCD sample to the final study sample is available in Figure S1.

2.5 | Statistical analysis

The sociodemographic characteristics and details about birth (including birth weight, presence of preterm birth, maternal age at birth, maternal education, household income and whether the mother was born in Canada) of study participants are presented as proportions for categorical variables, mean (standard deviation) for symmetric continuous variables and median (IQR) for skewed variables. Study participants' characteristics were compared to those of QLSCD non-participants using a chi-square test, one-way analysis of variance (ANOVA) and Wilcoxon rank sum test, depending on the variable type. Linear mixed models were used to estimate the association between PGSs and BMI *z*-scores to account for the repeated anthropometric measurements. Models included a smooth term for age, a random slope for age at measurement and an autoregressive order 1 covariance structure. We also assessed the association between PGSs and the two eating behaviours using linear regression. Statistical significance was determined using the Student *t*-test.

Longitudinal growth curve mediation analysis (LGCMA) was used to obtain the proportion of the variance of BMI *z*-score accounted for by both PGSs that is mediated by fussy eating and over-eating. The general model (Figure 1) included two latent variables (the intercept and the slope of the longitudinal outcome), a mediator regression model (Equation 1), an intercept regression model (Equation 2) and a slope regression (Equation 3):

$$EB = \beta_0 + \beta_1 PGSc_i + \beta_2 PGSa_i + \varepsilon_i. \quad (1)$$

$$\text{Intercept} = \phi_0 + \phi_1 PGSc_i + \phi_2 PGSa_i + \phi_3 EB_i + \varepsilon_i \quad (2)$$

$$\text{Slope} = \gamma_0 + \gamma_1 PGSc_i + \gamma_2 PGSa_i + \gamma_3 EB_i + \varepsilon_i \quad (3)$$

where *EB* denotes eating behaviour, *PGSc* denotes the child-based PGS and *PGSa* denotes the adult-based PGS. The proportion mediated (e.g. for the child-based PGS) can be obtained by solving the following (Equation 4) at each time-point:

$$\%\text{Mediated} = \text{Mediated effect}/\text{Total effect} \quad (4)$$

With: $\text{Mediated effect} = \beta_1 PGS_i \times (\phi_3 EB_i + \gamma_3 EB_i \times t)$.

And: $\text{Total effect} = \text{Mediated effect} + \phi_1 PGS_i + \gamma_1 PGS_i \times t$.

where *t* denotes age in years.

The methodology and Equations 1–4 were adapted and applied to our study from previously described work.[32] We adjusted for potential baseline characteristics that could confound the association between eating behaviours and BMI *z*-scores. We considered variables available in the QLSCD which may affect eating behaviours and BMI based on available literature.[33–38] Those variables include birth weight, maternal BMI, preterm birth (yes/no), being the mother's first child (yes/no), maternal age at birth (20 years and under, 21–34 years, 35 years and over), maternal education (did not complete secondary school, completed secondary school, completed post-secondary diploma, university diploma), household income (less than $30 000, $30 000 to $60 000, $60 000 to 80 000, more than $80 000) and immigration situation (mother born in Canada or not) and were added to the mediator and intercept regression models. We decided against including energy intake as a covariate in the analyses considering it more likely plays the role of intermediate variable between eating behaviours and BMI, rather than a confounder. There is evidence that eating behaviours are associated with dietary intakes (including energy intake)[39–41] in children, which is in turn associated with variations in BMI.[39,42] Adjusting for energy intake would result in mitigating one of the mechanisms by which eating behaviours can influence

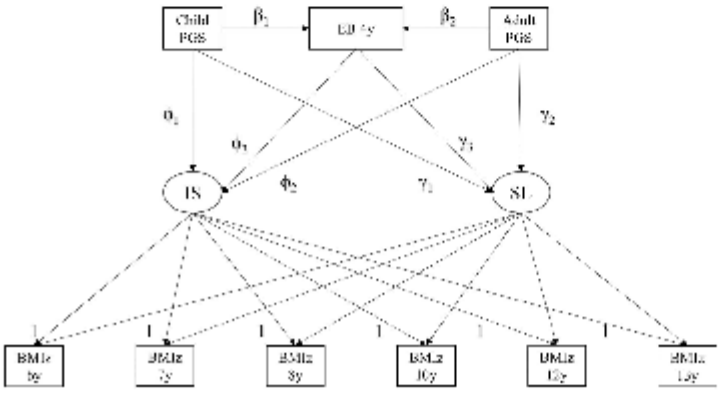


FIGURE 1 Latent growth curve mediation model. Figure depicts three regression models with β, coefficients for the eating behaviour outcome model; ϕ, coefficients for the IS outcome model; and Υ, coefficients for the SL outcome model. EB, eating behaviour; IS, initial state; SL, slope.

BMI. We also assessed whether the mediation differed between girls and boys as supplementary analyses. We compared a constrained model where all the coefficients were restricted to be equal in boys and girls alike to an unconstrained model where the coefficients were allowed to vary between sexes using ANOVA.

In sensitivity analysis, inverse probability weighting (IPW) was used to assess the robustness of the primary analysis to informative loss to follow-up. The R package 'twang'[43] was used to calculate a propensity score per participant based on a model where presence in the current study was predicted by baseline covariates likely to affect attrition (sex, maternal education, maternal BMI, birth weight, preterm birth, maternal age at birth, household income and whether the mother was born in Canada). The average treatment effect was used so that weighted study participants would resemble more accurately the initial QLSCD participants. All statistical analyses were completed using R version 4.3.0.[44]

3 | RESULTS

More than half of the participants were female (54.8%). The majority (86.3%) had mothers aged 20–34 years at time of birth, a household income above $30 000 (74.7%) and almost all had a mother born in Canada (97.7%) (Table 1). When compared to QLSCD participants excluded from the study, the study sample included more females (54.8% vs. 46.3%), had mothers with higher levels of education (60.4% vs. 53.2% with post-secondary education) and lived in households with higher income (33.7% vs. 27.7% ≥$60 000 income) (Table 1).

TABLE 1 Characteristics of study participants.

Characteristic % (n) or mean ± SD	Included, $N = 692$	Excluded, $N = 1428$	*p*-value
Sex, female, % (n)	54.8 (379)	46.3 (661)	**<0.001**
Birth weight, kg mean ± SD	3.42 ± 0.49	3.40 ± 0.50	0.351
Preterm birth, % (n)	5.1 (35)	4.5 (64)	0.631
First child, % (n)	41.8 (289)	45.3 (647)	0.135
Maternal age at birth, % (n)			0.098
≤20 years	1.7 (12)	3.3 (47)	
20–34 years	84.0 (581)	83.7 (1194)	
≥35 years	14.3 (99)	13.0 (186)	
Maternal education, % (n)			**0.009**
<Secondary school diploma	15.6 (110)	19.3 (275)	
Secondary school diploma	24.3 (164)	27.4 (391)	
Post-sec. except university	29.5 (206)	28.4 (405)	
University diploma	30.5 (212)	24.8 (354)	
Household income, % (n)			**0.003**
<$30 000	25.3 (175)	32.2 (447)	
$30 000–<$60 000	41.0 (284)	40.2 (559)	
$60 000–<$80 000	17.2 (119)	15.0 (208)	
≥$80 000	16.5 (114)	12.7 (176)	
Born in Canada, % (n)	97.7 (676)	83.4 (1191)	**<0.001**
Fussy eating med (IQR)	5.60 (4.80, 6.60)	5.67 (4.75, 6.67)	0.745
Over-eating med (IQR)	2.80 (2.33, 3.60)	2.75 (2.20, 3.50)	0.207
Obesity at 6 years, % (n)			0.685
Healthy weight	80.2 (441)	80.0 (496)	
Overweight	14.0 (77)	13.1 (81)	
Obesity	5.8 (32)	6.9 (43)	
Obesity at 13 years, % (n)			0.202
Healthy weight	63.9 (389)	67.1 (416)	
Overweight	22.0 (134)	22.1 (137)	
Obesity	14.1 (86)	10.8 (67)	

Note: Study participants included those ($N = 692$) who had complete anthropometric, genetic, eating behaviour, and covariate data available and were thus part of the main mediation analysis. Chi-square test used for categorical variables, ANOVA test used for normally distributed continuous variables, and Wilcoxon rank sum test used for non-normally distributed continuous variables. Significance threshold (bold) set at 0.05.

The child- and adult-based PGSs were substantially associated with BMI z-scores (Table 2) after accounting for correlations between measures taken in the same children over the range of 6 to 13 years of age. An increase of one standard deviation of the child and adult-based PGS was associated with an increase in BMI of 0.26 standard deviations (95% CI: 0.19–0.33, $p < 0.001$) and 0.34 standard deviations (95% CI: 0.27–0.41, $p < 0.001$), respectively. Both PGSs were inversely associated with fussy eating (child: $\beta = -0.14$, 95% CI: −0.24 to −0.04, $p = 0.007$; adult: $\beta = -0.11$, 95% CI: −0.21 to −0.01, $p = 0.030$) and positively associated with over-eating (child: $\beta = 0.13$, 95% CI: 0.07–0.20, $p < 0.001$; adult: $\beta = 0.15$, 95% CI: 0.08–0.22, $p < 0.001$). Both PGSs were significantly associated with BMI z-scores in boys and girls (Table 2). PGSs were always significantly associated with eating behaviours in boys, but not in girls, although the confidence intervals overlapped (Table 2).

Table 3 shows the proportion of the predictive association between both PGSs and BMI z-scores mediated by fussy eating and

TABLE 2 Association of the PGS with eating behaviours and BMI in children.

	Child PGS		Adult PGS	
Outcome	β (95% CI)	p-value	β (95% CI)	p-value
BMI z-score				
Total	0.26 (0.19, 0.33)	<0.001	0.34 (0.27, 0.41)	<0.001
Girls	0.22 (0.12, 0.32)	<0.001	0.25 (0.16, 0.34)	<0.001
Boys	0.31 (0.20, 0.42)	<0.001	0.45 (0.34, 0.56)	<0.001
Fussy eating				
Total	−0.14 (−0.24, −0.04)	**0.007**	−0.11 (−0.21, −0.01)	**0.030**
Girls	−0.07 (−0.20, 0.07)	0.340	−0.02 (−0.16, 0.11)	0.750
Boys	−0.22 (−0.37, −0.07)	**0.003**	−0.23 (−0.38, −0.07)	**0.004**
Overeating				
Total	0.13 (0.07, 0.20)	<0.001	0.15 (0.08, 0.22)	<0.001
Girls	0.07 (−0.02, 0.17)	0.113	0.10 (0.01, 0.19)	**0.033**
Boys	0.20 (0.10, 0.31)	<0.001	0.22 (0.12, 0.32)	<0.001

Note: Estimated effect (β) of increase in one standard deviation of the child- and adult-based PGS and BMI z-score (linear mixed model), fussy eating (linear regression) and over-eating (linear regression) with 95% confidence interval. BMI z-score model includes age, age-squared and sex as covariates with random intercept and slope for age. Eating behaviour models include sex as a covariate. Significance threshold (bold) set at 0.05.

TABLE 3 Proportion of the effect of PGS on BMI z-score mediated by eating behaviours.

	Child PGS		Adult PGS	
Model	% (95% CI)	p-value	% (95% CI)	p-value
Fussy eating	$N = 692$			
6 years	1.9 (−2.1, 5.8)	0.360	1.3 (−1.6, 4.3)	0.369
7 years	1.8 (−2.1, 5.6)	0.362	1.2 (−1.4, 3.9)	0.370
8 years	1.7 (−2.1, 5.5)	0.372	1.1 (−1.4, 3.6)	0.379
10 years	1.6 (−2.2, 5.3)	0.421	0.9 (−1.3, 3.1)	0.425
12 years	1.4 (−2.7, 5.5)	0.504	0.7 (−1.4, 2.9)	0.505
13 years	1.3 (−3.0, 5.7)	0.554	0.7 (−1.5, 2.8)	0.554
Over-eating	$N = 692$			
6 years	18.0 (3.1, 32.9)	**0.018**	14.2 (2.8, 25.5)	**0.014**
7 years	17.1 (3.0, 31.3)	**0.018**	12.7 (2.6, 22.9)	**0.014**
8 years	16.2 (2.7, 29.8)	**0.019**	11.4 (2.3, 20.5)	**0.015**
10 years	14.4 (1.8, 26.9)	**0.025**	9.1 (1.5, 16.6)	**0.018**
12 years	12.4 (0.4, 24.3)	**0.042**	7.1 (0.8, 13.3)	**0.028**
13 years	11.4 (−0.4, 23.1)	0.057	6.2 (0.4, 12.0)	**0.037**

Note: Estimated proportion mediated by eating behaviours (fussy eating, over-eating) in the association between PGS and BMI z-score with 95% CI using latent growth curve mediation analysis. Models are all adjusted for birth weight, preterm birth, maternal BMI, maternal age, maternal education, household income, whether the mother was born in Canada and whether the participant is the first child as covariates in both the eating behaviour and BMI z-score (effect on the initial state) outcome models. Significance threshold (bold) set at 0.05.

over-eating from 6 to 13 years of age. The mediation by over-eating was statistically significant at all but one time point and was higher for the child-based PGS at all six time points compared to the adult-based PGS. Additionally, the mediation decreased over time for both PGSs. Thus, the strongest mediation was observed at age 6 (child-based PGS: 18.0%, 95% CI: 3.1–32.9, *p*-value = 0.018; adult-based PGS: 14.2%, 95% CI: 2.8–25.5, *p*-value = 0.014) and the lowest at age 13 (child-based PGS: 11.4%, 95% CI: −0.4 to 23.1, *p*-value = 0.057; adult-based PGS: 6.2%, 95% CI: 0.4–12.0, *p*-value = 0.037) adjusting for covariates. Fussy eating did not mediate the association between both PGSs and BMI *z*-scores.

Figures 2 and 3 describe in more detail the paths from the two LGCMA models, one for each behaviour. Each growth model estimated the intercept (IS), analog to stability, and the slope (SL) of BMI from ages 6 to 13. Each model also estimated the putative direct and indirect contributions of each PGS, the latter through over-eating and fussy eating, respectively. The two models revealed unique direct and indirect (i.e. mediated) contributions of both the child-based PGS and adult-based PGS to BMI stability and slope. In the case of fussy eating (Figure 2), direct associations were the rule and mediation pathways were not conclusive. For instance, both the child-based PGS and the adult-based PGS were directly associated with the BMI intercept, but not with the BMI slope. The absence of mediation by fussy eating is indicated by the association between fussy eating and the BMI intercept being close to the null. Figure 3 illustrates that both the child-based PGS and the adult-based PGS were associated with over-eating, which was in turn associated with the BMI intercept. Over-eating was inversely associated with the BMI slope, which

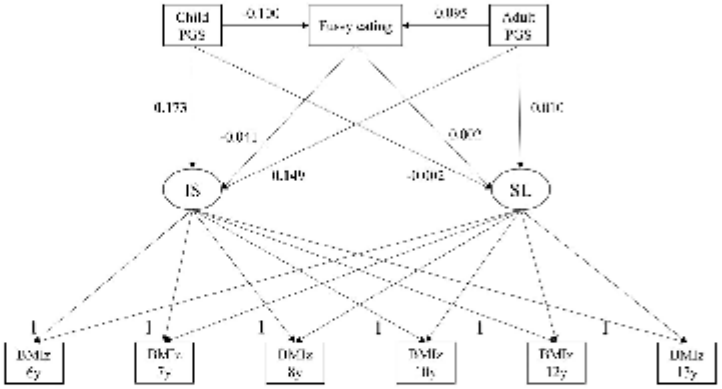


FIGURE 2 Latent growth curve mediation model, for fussy eating. The figure depicts four direct associations between PGSs and BMI *z*-scores: from the child PGS to the initial state of BMI *z*-scores (1) and the slope of BMI *z*-scores (2) and from the adult PGS to the initial state (3) and the slope (4). Four indirect associations between PGSs and BMI *z*-scores are presented: from the child PGS through fussy eating to the initial state (1) and the slope (2) and from the adult PGS through fussy eating to the initial state (3) and the slope (4). Paths in bold are significant at $p < 0.05$. IS, initial state; SL, slope.

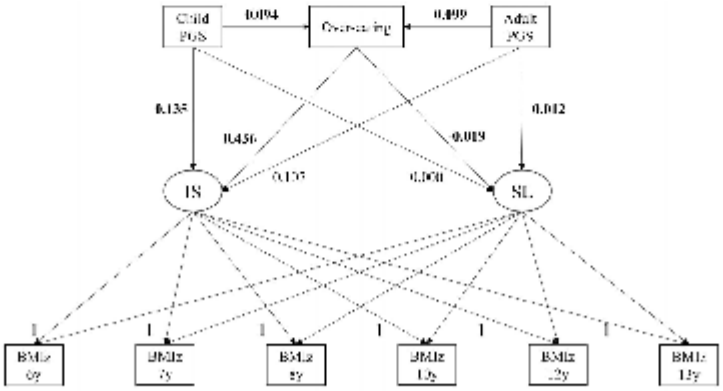


FIGURE 3 Latent growth curve mediation model, for over-eating. The figure depicts four direct associations between PGSs and BMI *z*-scores: from the child PGS to the initial state of BMI *z*-scores (1) and the slope of BMI *z*-scores (2) and from the adult PGS to the initial state (3) and the slope (4). Four indirect associations between PGSs and BMI *z*-scores are presented: from the child PGS through over-eating to the initial state (1) and the slope (2) and from the adult PGS through over-eating to the initial state (3) and the slope (4). Paths in bold are significant at $p < 0.05$. IS, initial state; SL, slope.

explains the decreasing mediation observed for both PGSs through time. The association between the adult-based PGS and BMI slope (β: 0.012, i.e. increase over time in the direct effect on BMI) was higher compared to the child-based PGS (β: 0.000), which illustrates the higher decrease in mediation observed through time for the adult-based PGS compared to the child-based PGS. In additional analysis, we assessed whether the proportion mediated by over-eating differed between boys and girls. The models did not differ by sex (Δ Chi-square: 10.50, Δ DF: 8, *p*-value: 0.232).

To assess the robustness of our findings, we performed IPW to ensure the study sample resembled the initial sample in terms of baseline characteristics. As seen in Table S1, IPW did not change the overall results, with the estimated proportion of effect mediated by over-eating decreasing with time for both PGSs. In terms of magnitude, the proportion mediated in the IPW model was higher at all but two time point compared to the original analyses for both PGSs. The main difference observed was that the confidence intervals related to the adult PGS using IPW were slightly larger compared to the initial analyses, leading to non-significant results. After adjusting for attrition, the proportion mediated by fussy eating was still close to null for both PGSs.

4 | DISCUSSION

The goal of the present study was to assess (1) the concurrent associations between two PGSs derived from child and adult GWAS and childhood BMI and (2) if problematic eating behaviours, such as fussy eating and over-eating, mediated these associations. Using LGCMA, we found that the child-derived and the adult-derived PGSs were independently associated with BMI *z*-scores in childhood. This suggests that albeit being correlated, both PGS capture unique genetic contributions to obesity. Furthermore, we confirmed that over-eating mediates the association between both PGSs and BMI *z*-scores. This result is consistent with recent research suggesting that weight-related genes act by changing the expression of proteins in brain regions involved in appetite and satiety regulation.[45]

Our study adds to growing evidence of the emerging role of common genetic variants influencing BMI in childhood in two ways: (1) the presence of independent effects on BMI from PGSs derived from child and adult GWAS, and (2) the increase over time in the direct effect on BMI for the adult-based PGS, but not the child-based PGS. This is consistent with evidence from large-scale GWAS analysis and twin studies. Increasing differences in weight between high- and low-genetic-risk individuals throughout childhood in 300 000 individuals suggest that new genetic sources of variation seem to emerge in childhood.[46] This is also supported by the observation that a PGS developed on an adult GWAS presented an increasing contribution to the variation in BMI throughout childhood, especially between 3 and 8 years.[47] Furthermore, results from a large-scale analysis of over 38 000 twin pairs support that genetic factors affecting variations in BMI change throughout childhood. The study[48] showed that genetic correlations between BMI measures were smaller with increasing age from 1 to 19 years. More specifically, our study added evidence for unique contributions to BMI variation from genetic factors present in early childhood and those emerging later in development. This suggests that PGSs derived from GWAS statistics obtained in child and adult populations contain overlapping, but distinct genetic factors.

The role of appetite as a mediator of the genetic contribution to body weight in childhood is a central tenet of the Behavioural susceptibility theory. Changes in the food environment contribute to the risk of obesity, with genetic susceptibility responsible for shaping the individual's response to such changes. Our study supports this hypothesis. Our over-eating score is composed of two items referring to eating too much and eating too fast. The well-known Children's Eating Behaviour Questionnaire subscales food responsiveness and satiety responsiveness/slowness in eating includes similar items as 'If allowed to, my child would eat too much', and 'My child finishes his/her meal very quickly', respectively.[7] It is likely that our over-eating score relates more to the appetite and satiety compared to emotional behaviours that can also be studied. This also aligns with another study completed using QLSCD data that found an association between the over-eating score and satiety responsiveness measured at 22 years using the adult version of the Children's Eating Behaviour Questionnaire.[49]

Prior studies have used a similar approach to test the mediation of the contribution of genetic factors to BMI by appetite-related traits in children. One study[17] found mediation by satiety responsiveness, while another[18] did not. Differences between prior results could be explained through a contrast in methods. Mainly, cross-sectional data was used in the study that observed mediation at a single point in time, while the other used a longitudinal method designed to identify mediation for BMI growth (slope), rather than BMI at a specific time point. With its longitudinal design focused on identifying mediation at different points in time, our study allowed us to shed light on previous mixed results. We identified mediation through over-eating at sequential time points that was largely driven by the intercept of BMI. The use of structural models allowed us to observe the strong association between over-eating and the intercept of BMI responsible for the high mediation at 6 years, which then decreased due to an inverse effect of over-eating on the slope of BMI.

This result is consistent with another study of children that found decreasing mediation through time. High appetite at 2 years mediated the effect of a 16 variant PGS on BMI *z*-score from 2 to 5 years in over 1000 French Canadian children.[19] The proportions mediated by high appetite decreased from 47% at the 2-year mark to 24% at the 5-year mark. Like our study, the reduction in mediation at older time points likely stems from the passage of time, where appetite at 2 years becomes progressively less influential. Indeed, the inverse effect of over-eating on BMI growth could be explained by the BMI observations becoming further apart from the index measurement of over-eating as time passes. The concordance of our results and those obtained in Lauzon-Guillain et al.[19] is compelling since one study gathered anthropometric data before adiposity rebound (typically around 5 years) and the other after. This suggests that mediation through over-eating remains present at different developmental stages.

The presence of mediation of genetic susceptibility to obesity through over-eating aligns with molecular genetic literature that

emphasizes the importance of the brain in translating the effect of genetic factors on body weight regulation. RNA-based analyses applied to large-scale GWAS have shown that most variants associated with BMI are predominantly expressed in the central nervous system in regions such as the hippocampus and limbic system, involved in cognition and emotion regulation, but also the hypothalamus and pituitary gland, responsible for appetite regulation.[9,10] Additionally, a recent study identified 60 instances where protein concentration in the brain was linked with genetic variants known to influence obesity.[45] More specifically, that study used a combination of gene colocalization and mendelian randomization analyses focused on the left dorsolateral prefrontal cortex. This brain region notably influences appetite, satiety regulation and cognitive functions, including decision-making and executive functioning.[45] These results provide a biological basis for the theory that genetic susceptibility to obesity is expressed through variations in appetite, resulting from changes in protein concentration. We did not observe mediation through fussy eating. This result complements the available but scarce literature on the subject. Avoidant eating and slowness in eating were not found to mediate the effect of a PGS on BMI in a Finnish adult cohort and a Norwegian longitudinal children study, respectively.[15,18]

The additional evidence favouring the Behavioural susceptibility theory further provides incentives for supporting obesity prevention efforts that focus on creating a healthy food environment rather than centered on personal responsibility. This is an important public health issue since a shift from the global food system leading to deteriorating nutritional choices is considered a driving factor behind the rise of obesity.[50] For example, the current food environment in Canada does not favour healthy food consumption patterns on most fronts.[51] A recent evaluation of the policies and actions of multiple levels of government in Canada for creating a healthy food environment highlighted (1) many limitations to Canada's food environment policy landscape and (2) restrictions on marketing aimed at children as a key policy area to improve a healthier food landscape in Canada.[51]

The primary strength of the present study is the use of a longitudinal design. This allows a better characterization of the mediation of genetic susceptibility to obesity by over-eating throughout childhood and adolescence. Many studies only use a singular time point to describe this process, while we were able to detail how the mediation evolves through time. Furthermore, the use of PRS-CS to calculate the PGSs is another strength of the study considering the method captures more genetic variability compared with methods only incorporating a few variants into the scores. The study also has a few limitations. First, even though we adjusted our analysis models for potential confounders, there could still be unmeasured confounders, such as physical activity, blurring the association between the eating behaviours and BMI z-scores specifically. Second, we used inverse probability weighting to attenuate potential selection bias through attrition. However, we could only produce a weighted study sample similar to the initial QLSCD sample in terms of available baseline characteristics; therefore, the study sample could still differ from the initial sample in some capacity. Third, BMI is known to have limits in distinguishing between lean and fat mass and does not distinguish between the location of fat in the body. Fourth, eating behaviour data was not available throughout the childhood period with measured anthropometric collection. Eating behaviours might change as children's age increases, and the strength of the association between eating behaviours exhibited in early childhood and BMI would be expected to diminish as time passes.

5 | CONCLUSION

In our investigation of French-Canadian children, over-eating throughout childhood, but not fussy eating, mediated genetic susceptibility to obesity. This result aligns with the prevailing theoretical construct behind the genetic architecture of obesity that hinges on the role of appetite. The few previous studies identified appetite-related traits as mediators of genetic susceptibility to weight gain but were mostly performed in adult populations and had a cross-sectional design. The present study analysed data from a longitudinal study of Canadian children and used dense genotypic scores as measures of genetic susceptibility to obesity. The finding of mediation by over-eating suggests that global and wide-reaching policies aimed at introducing a healthier food environment will be key in implementing obesity prevention efforts.

AUTHOR CONTRIBUTIONS

Danick Goulet, Lise Dubois and Michel Boivin contributed to the conception and design of the study. Lise Dubois and Michel Boivin were involved in the data collection from QLSCD participants. Danick Goulet performed the statistical analyses with insights and revisions from Christopher A. Gravel. Danick Goulet was responsible for the first draft of the manuscript. All authors participated in discussions concerning the development of the study and were involved in the revisions to the manuscript and approved the submitted version.

ACKNOWLEDGEMENTS

We also acknowledge the contribution of Till Andlauer, Stéphane Paquin, Geneviève Morneau-Vaillancourt, Isabelle Ouellet-Morin and Michel Boivin, who were involved in the quality control of the genetic data of the QLSCD participants that are used in the research. We are grateful to the QLSCD participants and their families who took part in the various data collection rounds over the years.

FUNDING INFORMATION

This work was supported by a CIHR operating grant (#165964). The funders were not involved in the design or conduct of the study; collection, management, analysis, or interpretation of the data or preparation, review or approval of the manuscript. The analyses were performed using data from the Quebec Longitudinal Study of Child Development (QLSCD), conducted by Sante Quebec, a division of the Institut de la Statistique du Quebec (ISQ) and funded by the Ministry of Health and Social Services of Quebec.

CONFLICT OF INTEREST STATEMENT

No conflict of interest was declared.

DATA AVAILABILITY STATEMENT

The data used to support the findings of this study was obtained from the Québec Longitudinal Study of Child Development, conducted by Santé Québec, a division of the Institut de la Statistique du Québec and may be released upon application to the Institut de la Statistique du Québec, through the Zone de recherche at: https://statistique.quebec.ca/fr/institut/services-recherche#/accueil.

ORCID

Danick Goulet https://orcid.org/0009-0005-7534-6553

REFERENCES

1. Haththotuwa RN, Wijeyaratne CN, Senarath U. Chapter 1 – Worldwide epidemic of obesity. In: Mahmood TA, Arulkumaran S, Chervenak FA, eds. *Obesity and Obstetrics (Second Edition)*. Elsevier; 2020:3-8.
2. Statistique Canada. Indice de masse corporelle des enfants et des jeunes, 2012 à 2013. 2015. Accessed November 23, 2020. https://www150.statcan.gc.ca/n1/pub/82-625-x/2014001/article/14105-fra.htm
3. Simmonds M, Llewellyn A, Owen CG, Woolacott N. Predicting adult obesity from childhood obesity: a systematic review and meta-analysis. *Obes Rev*. 2016;17(2):95-107. doi:10.1111/obr.12334
4. Silventoinen K, Jelenkovic A, Sund R, et al. Genetic and environmental effects on body mass index from infancy to the onset of adulthood: an individual-based pooled analysis of 45 twin cohorts participating in the COllaborative project of Development of Anthropometrical measures in Twins (CODATwins) study. *Am J Clin Nutr*. 2016;104(2):371-379. doi:10.3945/ajcn.116.130252
5. Vogelezang S, Bradfield JP, Ahluwalia TS, et al. Novel loci for childhood body mass index and shared heritability with adult cardiometabolic traits. *PLoS Genet*. 2020;16(10):e1008718. doi:10.1371/journal.pgen.1008718
6. Felix JF, Bradfield JP, Monnereau C, et al. Genome-wide association analysis identifies three new susceptibility loci for childhood body mass index. *Hum Mol Genet*. 2016;25(2):389-403. doi:10.1093/hmg/ddv472
7. Wardle J, Guthrie CA, Sanderson S, Rapoport L. Development of the children's eating behaviour questionnaire. *J Child Psychol Psychiatry*. 2001;42(7):963-970. doi:10.1111/1469-7610.00792
8. Silventoinen K, Konttinen H. Obesity and eating behavior from the perspective of twin and genetic research. *Neurosci Biobehav Rev*. 2020;109:150-165. doi:10.1016/j.neubiorev.2019.12.012
9. Locke AE, Kahali B, Berndt SI, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature*. 2015;518(7538):197-206. doi:10.1038/nature14177
10. Yengo L, Sidorenko J, Kemper KE, et al. Meta-analysis of genome-wide association studies for height and body mass index in ~700000 individuals of European ancestry. *Hum Mol Genet*. 2018;27(20):3641-3649. doi:10.1093/hmg/ddy271
11. Konttinen H, Llewellyn C, Wardle J, et al. Appetitive traits as behavioural pathways in genetic susceptibility to obesity: a population-based cross-sectional study. *Sci Rep*. 2015;5(1):14726. doi:10.1038/srep14726
12. de Lauzon-Guillain B, Clifton EA, Day FR, et al. Mediation and modification of genetic susceptibility to obesity by eating behaviors. *Am J Clin Nutr*. 2017;106(4):996-1004. doi:10.3945/ajcn.117.157396
13. Jacob R, Drapeau V, Tremblay A, Provencher V, Bouchard C, Pérusse L. The role of eating behavior traits in mediating genetic susceptibility to obesity. *Am J Clin Nutr*. 2018;108(3):445-452. doi:10.1093/ajcn/nqy130
14. Brunner EJ, Maruyama K, Shipley M, et al. Appetite disinhibition rather than hunger explains genetic effects on adult BMI trajectory. *Int J Obes (Lond)*. 2021;45(4):758-765. doi:10.1038/s41366-020-00735-9
15. Masip G, Silventoinen K, Keski-Rahkonen A, et al. The genetic architecture of the association between eating behaviors and obesity: combining genetic twin modeling and polygenic risk scores. *Am J Clin Nutr*. 2020;112(4):956-966. doi:10.1093/ajcn/nqaa181
16. Herle M, Smith AD, Kininmonth A, Llewellyn C. The role of eating behaviours in genetic susceptibility to obesity. *Curr Obes Rep*. 2020;9(4):512-521. doi:10.1007/s13679-020-00402-0
17. Llewellyn CH, Trzaskowski M, van Jaarsveld CHM, Plomin R, Wardle J. Satiety mechanisms in genetic risk of obesity. *JAMA Pediatr*. 2014;168(4):338-344. doi:10.1001/jamapediatrics.2013.4944
18. Steinsbekk S, Belsky D, Guzey IC, Wardle J, Wichstrøm L. Polygenic risk, appetite traits, and weight gain in middle childhood: a longitudinal study. *JAMA Pediatr*. 2016;170(2):e154472. doi:10.1001/jamapediatrics.2015.4472
19. de Lauzon-Guillain B, Koudou YA, Botton J, et al. Association between genetic obesity susceptibility and mother-reported eating behaviour in children up to 5 years. *Pediatr Obes*. 2019;14(5):e12496. doi:10.1111/ijpo.12496
20. Ge T, Chen C-Y, Ni Y, Feng Y-CA, Smoller JW. Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nat Commun*. 2019;10(1):1776. doi:10.1038/s41467-019-09718-5
21. Orri M, Boivin M, Chen C, et al. Cohort profile: Quebec Longitudinal Study of Child Development (QLSCD). *Soc Psychiatry Psychiatr Epidemiol*. 2020;56:883-894. doi:10.1007/s00127-020-01972-z
22. de Onis M, Onyango AW, Borghi E, Siyam A, Nishida C, Siekmann J. Development of a WHO growth reference for school-aged children and adolescents. *Bull World Health Organ*. 2007;85(9):660-667. doi:10.2471/blt.07.043497
23. childsds: data and methods around reference values in pediatrics. 2022. https://CRAN.R-project.org/package=childsds
24. Golding, Pembrey, Jones, The Alspac Study T. ALSPAC–the Avon Longitudinal Study of Parents and Children. *Paediatr Perinat Epidemiol*. 2001;15(1):74-87. doi:10.1046/j.1365-3016.2001.00325.x
25. Institut de la statistique du Québec. *Enquête sociale et de santé auprès des enfants et des adolescents québécois*. Les Publications du Québec; 1999:520. https://statistique.quebec.ca/fr/document/enquete-sociale-et-de-sante-aupres-des-enfants-et-adolescents-quebecois-1999
26. Dubois L, Farmer AP, Girard M, Peterson K. Preschool children's eating behaviours are related to dietary adequacy and body weight. *Eur J Clin Nutr*. 2007;61(7):846-855. doi:10.1038/sj.ejcn.1602586
27. Perret LC, Boivin M, Monneau-Vaillancourt G, et al. Polygenic risk score and peer victimisation independently predict depressive symptoms in adolescence: results from the Quebec Longitudinal Study of Children Development. *J Child Psychol Psychiatry*. 2022;64(3):388-396. doi:10.1111/jcpp.13706
28. Auton A, Abecasis GR, Altshuler DM, et al. A global reference for human genetic variation. *Nature*. 2015;526(7571):68-74. doi:10.1038/nature15393
29. Delaneau O, Zagury J-F, Marchini J. Improved whole-chromosome phasing for disease and population genetic studies. *Nat Methods*. 2013;10(1):5-6. doi:10.1038/nmeth.2307
30. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet*. 2009;5(6):e1000529. doi:10.1371/journal.pgen.1000529
31. Khera AV, Chaffin M, Zekavat SM, et al. Whole-genome sequencing to characterize monogenic and polygenic contributions in patients hospitalized with early-onset myocardial infarction. *Circulation*. 2019;139(13):1593-1602. doi:10.1161/circulationaha.118.035658

32. Sullivan A, Gunzler D, Morris N, VanderWeele T. Longitudinal mediation analysis with latent growth curves. 2021. doi:10.48550/arXiv.2103.05765
33. Keane E, Layte R, Harrington J, Kearney PM, Perry IJ. Measured parental weight status and familial socio-economic status correlates with childhood overweight and obesity at age 9. *PLoS One*. 2012;7(8):e43503. doi:10.1371/journal.pone.0043503
34. Barclay K, Myrskylä M. Maternal age and offspring health and health behaviours in late adolescence in Sweden. *SSM Popul Health*. 2016;2:68-76. doi:10.1016/j.ssmph.2016.02.012
35. Williams AS, Ge B, Petroski G, Kruse RL, McElroy JA, Koopman RJ. Socioeconomic status and other factors associated with childhood obesity. *J Am Board Fam Med*. 2018;31(4):514-521. doi:10.3122/jabfm.2018.04.170261
36. Gnawali A. Prematurity and the risk of development of childhood obesity: piecing together the pathophysiological puzzle: A literature review. *Cureus*. 2021;13(12):e20518. doi:10.7759/cureus.20518
37. Ochiai H, Shirasawa T, Ohtsu T, et al. Number of siblings, birth order, and childhood overweight: a population-based cross-sectional study in Japan. *BMC Public Health*. 2012;12:766. doi:10.1186/1471-2458-12-766
38. White PA, Awad YA, Gauvin L, et al. Household income and maternal education in early childhood and risk of overweight and obesity in late childhood: findings from seven birth cohort studies in six high-income countries. *Int J Obes (Lond)*. 2022;46(9):1703-1711. doi:10.1038/s41366-022-01171-7
39. French SA, Epstein LH, Jeffery RW, Blundell JE, Wardle J. Eating behavior dimensions. Associations with energy intake and body weight. A review. *Appetite*. 2012;59(2):541-549. doi:10.1016/j.appet.2012.07.001
40. Fogel A, Fries LR, McCrickerd K, et al. Oral processing behaviours that promote children's energy intake are associated with parent-reported appetitive traits: results from the GUSTO cohort. *Appetite*. 2018;126:8-15. doi:10.1016/j.appet.2018.03.011
41. Maneschy I, Jimeno-Martínez A, Miguel-Berges ML, et al. Eating behaviours and dietary intake in children and adolescents: a systematic review. *Curr Nutr Rep*. 2024;13(3):363-376. doi:10.1007/s13668-024-00544-w
42. Kininmonth A, Smith A, Carnell S, Steinsbekk S, Fildes A, Llewellyn C. The association between childhood adiposity and appetite assessed using the Child Eating Behavior Questionnaire and Baby Eating Behavior Questionnaire: a systematic review and meta-analysis. *Obes Rev*. 2021;22(5):e13169. doi:10.1111/obr.13169
43. twang: toolkit for weighting and analysis of nonequivalent groups. 2021 https://CRAN.R-project.org/package=twang
44. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing; 2023. https://www.R-project.org/
45. Gagnon E, Girard A, Gobeil É, et al. Genetic control of body weight by the human brain proteome. *iScience*. 2023;26(4):106376. doi:10.1016/j.isci.2023.106376
46. Khera AV, Chaffin M, Wade KH, et al. Polygenic prediction of weight and obesity trajectories from birth to adulthood. *Cell*. 2019;177(3):587-596.e9. doi:10.1016/j.cell.2019.03.028
47. Helgeland Ø, Vaudel M, Sole-Navais P, et al. Characterization of the genetic architecture of infant and early childhood body mass index. *Nat Metab*. 2022;4:344-358. doi:10.1038/s42255-022-00549-1
48. Silventoinen K, Li W, Jelenkovic A, et al. Changing genetic architecture of body mass index from infancy to early adulthood: an individual based pooled analysis of 25 twin cohorts. *Int J Obes (Lond)*. 2022;46:1901-1909. doi:10.1038/s41366-022-01202-3
49. Dubois L, Bédard B, Goulet D, Prudhomme D, Tremblay RE, Boivin M. Eating behaviors, dietary patterns and weight status in emerging adulthood and longitudinal associations with eating behaviors in early childhood. *Int J Behav Nutr Phys Act*. 2022;19(1):139. doi:10.1186/s12966-022-01376-z
50. Swinburn BA, Sacks G, Hall KD, et al. The global obesity pandemic: shaped by global drivers and local environments. *Lancet*. 2011;378(9793):804-814. doi:10.1016/S0140-6736(11)60813-1
51. Vanderlee L, Goorang S, Karbasy K, Vandevijvere S, L'Abbé MR. Policies to create healthier food environments in Canada: experts' evaluation and prioritized actions using the Healthy Food Environment Policy Index (Food-EPI). *Int J Environ Res Public Health*. 2019;16(22):4473. doi:10.3390/ijerph16224473

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Goulet D, Boivin M, Gravel CA, Little J, Potter BK, Dubois L. Mediation of genetic susceptibility to obesity through eating behaviours in children. *Pediatric Obesity*. 2024;e13180. doi:10.1111/ijpo.13180

Appendix 6: Proof of authorization to reuse manuscript 2 in the thesis.

JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS

Nov 05, 2024

This Agreement between Danick Goulet ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number	5902610203345
License date	Nov 05, 2024
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	Pediatric Obesity
Licensed Content Title	Mediation of genetic susceptibility to obesity through eating behaviours in children
Licensed Content Author	Lise Dubois, Beth K. Potter, Julian Little, et al
Licensed Content Date	Oct 10, 2024
Licensed Content Volume	0
Licensed Content Issue	0
Licensed Content Pages	11
Type of use	Dissertation/Thesis
Requestor type	Author of this Wiley article

Format	Electronic
Portion	Full article
Will you be translating?	No
Title of new work	The interplay between dietary habits and genetic susceptibility to obesity in children and adolescents
Institution name	University of Ottawa
Expected presentation date	Dec 2024
The Requesting Person / Organization to Appear on the License	Danick Goulet
	Danick Goulet
Requestor Location	, NOTSELECTED
Order reference number	11563
Publisher Tax ID	EU826007151
Total	0.00 CAD

Terms and Conditions

TERMS AND CONDITIONS

This copyrighted material is owned by or exclusively licensed to John Wiley & Sons, Inc. or one of its group companies (each a"Wiley Company") or handled on behalf of a society with which a Wiley Company has exclusive publishing rights in relation to a particular work (collectively "WILEY"). By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the billing and payment terms and conditions established by the Copyright Clearance Center Inc., ("CCC's Billing and Payment terms and conditions"), at

the time that you opened your RightsLink account (these are available at any time at http://myaccount.copyright.com).

Terms and Conditions

- The materials you have requested permission to reproduce or reuse (the "Wiley Materials") are protected by copyright.
- You are hereby granted a personal, non-exclusive, non-sub licensable (on a stand-alone basis), non-transferable, worldwide, limited license to reproduce the Wiley Materials for the purpose specified in the licensing process. This license, **and any CONTENT (PDF or image file) purchased as part of your order,** is for a one-time use only and limited to any maximum distribution number specified in the license. The first instance of republication or reuse granted by this license must be completed within two years of the date of the grant of this license (although copies prepared before the end date may be distributed thereafter). The Wiley Materials shall not be used in any other manner or for any other purpose, beyond what is granted in the license. Permission is granted subject to an appropriate acknowledgement given to the author, title of the material/book/journal and the publisher. You shall also duplicate the copyright notice that appears in the Wiley publication in your use of the Wiley Material. Permission is also granted on the understanding that nowhere in the text is a previously published source acknowledged for all or part of this Wiley Material. Any third party content is expressly excluded from this permission.
- With respect to the Wiley Materials, all rights are reserved. Except as expressly granted by the terms of the license, no part of the Wiley Materials may be copied, modified, adapted (except for minor reformatting required by the new Publication), translated, reproduced, transferred or distributed, in any form or by any means, and no derivative works may be made based on the Wiley Materials without the prior permission of the respective copyright owner.**For STM Signatory Publishers clearing permission under the terms of the STM Permissions Guidelines only, the terms of the license are extended to include subsequent editions and for editions in other languages, provided such editions are for the work as a whole in situ and does not involve the separate exploitation of the permitted figures or extracts,** You may not alter, remove or suppress in any manner any copyright, trademark or other notices displayed by the Wiley Materials. You may not license, rent, sell, loan, lease, pledge, offer as security, transfer or assign the Wiley Materials on a stand-alone basis, or any of the rights granted to you hereunder to any other person.
- The Wiley Materials and all of the intellectual property rights therein shall at all times remain the exclusive property of John Wiley & Sons Inc, the Wiley Companies, or their respective licensors, and your interest therein is only that of having possession of and the right to reproduce the Wiley Materials pursuant to Section 2 herein during the continuance of this Agreement. You agree that you own no right, title or interest in or to the Wiley Materials or any of the intellectual property rights therein. You shall have no rights hereunder other than the license as provided for above in Section 2. No right, license or interest to any trademark, trade name, service mark or other branding ("Marks") of WILEY or its licensors is granted hereunder, and you agree that you shall not assert any such right, license or interest with respect thereto

- NEITHER WILEY NOR ITS LICENSORS MAKES ANY WARRANTY OR REPRESENTATION OF ANY KIND TO YOU OR ANY THIRD PARTY, EXPRESS, IMPLIED OR STATUTORY, WITH RESPECT TO THE MATERIALS OR THE ACCURACY OF ANY INFORMATION CONTAINED IN THE MATERIALS, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, ACCURACY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, USABILITY, INTEGRATION OR NON-INFRINGEMENT AND ALL SUCH WARRANTIES ARE HEREBY EXCLUDED BY WILEY AND ITS LICENSORS AND WAIVED BY YOU.

- WILEY shall have the right to terminate this Agreement immediately upon breach of this Agreement by you.

- You shall indemnify, defend and hold harmless WILEY, its Licensors and their respective directors, officers, agents and employees, from and against any actual or threatened claims, demands, causes of action or proceedings arising from any breach of this Agreement by you.

- IN NO EVENT SHALL WILEY OR ITS LICENSORS BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR ENTITY FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, EXEMPLARY OR PUNITIVE DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, PROVISIONING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

- Should any provision of this Agreement be held by a court of competent jurisdiction to be illegal, invalid, or unenforceable, that provision shall be deemed amended to achieve as nearly as possible the same economic effect as the original provision, and the legality, validity and enforceability of the remaining provisions of this Agreement shall not be affected or impaired thereby.

- The failure of either party to enforce any term or condition of this Agreement shall not constitute a waiver of either party's right to enforce each and every term and condition of this Agreement. No breach under this agreement shall be deemed waived or excused by either party unless such waiver or consent is in writing signed by the party granting such waiver or consent. The waiver by or consent of a party to a breach of any provision of this Agreement shall not operate or be construed as a waiver of or consent to any other or subsequent breach by such other party.

- This Agreement may not be assigned (including by operation of law or otherwise) by you without WILEY's prior written consent.

- Any fee required for this permission shall be non-refundable after thirty (30) days from receipt by the CCC.

- These terms and conditions together with CCC's Billing and Payment terms and conditions (which are incorporated herein) form the entire agreement between you and WILEY concerning this licensing transaction and (in the absence of fraud) supersedes all prior agreements and representations of the parties, oral or written. This Agreement may not be amended except in writing signed by both parties. This Agreement shall be binding upon and inure to the benefit of the parties' successors, legal representatives, and authorized assigns.

- In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall prevail.

- WILEY expressly reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

- This Agreement will be void if the Type of Use, Format, Circulation, or Requestor Type was misrepresented during the licensing process.

- This Agreement shall be governed by and construed in accordance with the laws of the State of New York, USA, without regards to such state's conflict of law rules. Any legal action, suit or proceeding arising out of or relating to these Terms and Conditions or the breach thereof shall be instituted in a court of competent jurisdiction in New York County in the State of New York in the United States of America and each party hereby consents and submits to the personal jurisdiction of such court, waives any objection to venue in such court and consents to service of process by registered or certified mail, return receipt requested, at the last known address of such party.

WILEY OPEN ACCESS TERMS AND CONDITIONS

Wiley Publishes Open Access Articles in fully Open Access Journals and in Subscription journals offering Online Open. Although most of the fully Open Access journals publish open access articles under the terms of the Creative Commons Attribution (CC BY) License only, the subscription journals and a few of the Open Access Journals offer a choice of Creative Commons Licenses. The license type is clearly identified on the article.

The Creative Commons Attribution License

The Creative Commons Attribution License (CC-BY) allows users to copy, distribute and transmit an article, adapt the article and make commercial use of the article. The CC-BY license permits commercial and non-

Creative Commons Attribution Non-Commercial License

The Creative Commons Attribution Non-Commercial (CC-BY-NC)License permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.(see below)

Creative Commons Attribution-Non-Commercial-NoDerivs License

Appendix 7: STROBE reporting checklist for Manuscript 2.

Original checklist from: von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ. 2007 Oct 20;335(7624):806-8. doi: 10.1136/bmj.39335.541782.AD. PMID: 17947786; PMCID: PMC2034723.

	Item No.	Recommendation	Page No.	Relevant text from manuscript
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	Abstract p.90	Birth-cohort study
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	Abstract p.90	
Introduction				
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	Introduction p.91-92	
Objectives	3	State specific objectives, including any prespecified hypotheses	Introduction p.93	
Methods				
Study design	4	Present key elements of study design early in the paper	Methods (study design and participants) p.93	
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	Methods (study design and participants) p.93	
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	Methods (study design and participants) p.93	
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	NA	

Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	Methods (Anthropometric measures, Eating behaviours, Genotyping and polygenic score construction) p.94-95	
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	Methods (Anthropometric measures, Eating behaviours, Genotyping and polygenic score construction) p.94-95	
Bias	9	Describe any efforts to address potential sources of bias	Methods (Statistical analysis) p.97	Covariate adjustments and inverse probability weighting for attrition
Study size	10	Explain how the study size was arrived at	Methods (Study design and participants, Genotyping and polygenic score construction) p.94-95	Available sample
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	Methods (Eating behaviours, Genotyping and polygenic score construction) p.94-95	Fussy eating and over-eating score. Standardized polygenic score
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	Methods (Statistical analysis) p.96-97	
		(*b*) Describe any methods used to examine subgroups and interactions	Methods (statistical analysis) p.97	Sex subgroup
		(*c*) Explain how missing data were addressed	NA	Complete case
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	Methods (statistical analysis) p.97	Inverse probability weighting for attrition
		(*e*) Describe any sensitivity analyses	Methods (statistical analysis) p.97	Inverse probability weighting for attrition
Results				
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	Methods (Study design and participants, Genotyping and polygenic score construction) p.94-95	
		(b) Give reasons for non-participation at each stage	Methods (Study design and participants) p.94	Described in cohort profile cited.

		(c) Consider use of a flow diagram	Supplementary Figure 5-1	
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	Results p.98, Table 5-1	
		(b) Indicate number of participants with missing data for each variable of interest	Supplementary Figure 5-1	
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	-	
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	Table 5-1	Weight category % at 6 and 13 years old
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	NA	
		Cross-sectional study—Report numbers of outcome events or summary measures	NA	
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	Results p.98 Tables 5-2 and 5-3	Adjusted estimates for direct associations and mediation analyses
		(*b*) Report category boundaries when continuous variables were categorized	NA	
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	NA	
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	Results p.99, Supplementary Table 5-1	Sex subgroup and inverse probability weighting
Discussion				
Key results	18	Summarise key results with reference to study objectives	Discussion p.100	
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	Discussion p.102-103	
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	Discussion p.100-102	
Generalisability	21	Discuss the generalisability (external validity) of the study results	-	
Other information				
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	p.111	Funding statement

Appendix 8: Proof of submission email for manuscript 3, Journal of Nutritional Science.

Outlook

Fwd: Journal of Nutritional Science - JNS-2024-0307: Manuscript Submission

From Lise Dubois <ldubois@uottawa.ca>
Date Tue 2024-12-03 2:12 PM
To Danick Goulet <dgoul082@uottawa.ca>

> 19-Nov-2024
>
> Dear Dr. Dubois,
>
> Your manuscript entitled "Interaction between polygenic scores and dietary intake in relation to body mass index in Canadian children" has been successfully submitted online for consideration for publication in Journal of Nutritional Science. Your manuscript ID is JNS-2024-0307. If we have any queries regarding your submission we will contact you within the next few days.
>
> Please mention the above manuscript ID in all future correspondence. If there are any changes in your contact details, please log in to ScholarOne Manuscripts at https://mc.manuscriptcentral.com/jns and edit your user information as appropriate. You can view the status of your manuscript at any time by checking your Author Centre after logging in to https://mc.manuscriptcentral.com/jns
>
> As Journal of Nutritional Science is an Open Access (OA) publication, you might also like to find out more about publishing OA such as the benefits of doing so and things you need to know by visiting our Publishing Open Access information page:
>
> https://www.cambridge.org/core/services/authors/journals/publishing-open-access
>
> Thank you for submitting your manuscript to Journal of Nutritional Science.
>
> Kind regards,
> Emilie McDermott
>
> Editorial Assistant
> Journal of Nutritional Science
> jns.edoffice@cambridge.org

Appendix 9: STROBE reporting checklist for Manuscript 3.

Original checklist from: von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ. 2007 Oct 20;335(7624):806-8. doi: 10.1136/bmj.39335.541782.AD. PMID: 17947786; PMCID: PMC2034723.

	Item No.	Recommendation	Section and page No.	Relevant text from manuscript
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	Abstract p.122	Secondary analysis of birth-cohort study.
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	Abstract p.122	
Introduction				
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	Introduction p.123-124	
Objectives	3	State specific objectives, including any prespecified hypotheses	Introduction p.124	
Methods				
Study design	4	Present key elements of study design early in the paper	Methods (study design and participants) p.125	
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	Methods (study design and participants) p.125	
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	Methods (study design and participants) p.125	
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	NA	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	Methods (Anthropometric measurements, Food, macronutrient and energy intake,	

			Genotyping and polygenic score construction) p.125-127	
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	Methods (Anthropometric measurements, Food, macronutrient and energy intake, Genotyping and polygenic score construction) p.125-127	
Bias	9	Describe any efforts to address potential sources of bias	Methods (Statistical analysis) p.127-128	Covariate adjustments and inverse probability weighting for attrition
Study size	10	Explain how the study size was arrived at	Methods (Study design and participants) p.125, Supplementary Figure 6-1	Participants with available BMI, covariates, and food, macronutrient and energy data.
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	Methods (Food, macronutrient and energy intake, Genotyping and polygenic score construction) p.126-127	Single average dietary intake per participant. Standardized polygenic score
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	Methods (Statistical analysis) p.127-128	
		(*b*) Describe any methods used to examine subgroups and interactions	Methods (statistical analysis) p.127-128	Interaction with sex tested for all statistical models
		(*c*) Explain how missing data were addressed	NA	Complete case
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	Methods (statistical analysis) p.128	Inverse probability weighting for attrition
		(*e*) Describe any sensitivity analyses	Methods (statistical analysis) p.128	Inverse probability weighting for attrition. Adjustment for total energy intake in macronutrient models.
Results				

Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	Methods (Study design and participants) p.125. Supplementary Figure 6-1.	Flow chart for study sample available.
		(b) Give reasons for non-participation at each stage	Methods (Study design and participants) p.125	Described in cohort profile.
		(c) Consider use of a flow diagram	Supplementary Figure 6-1	Flow chart for study sample.
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	Results p.129, Table 6-1 p.136-137	Supplementary Table 6-1 and 6-2 also have comparative information to non-participants.
		(b) Indicate number of participants with missing data for each variable of interest	NA	Complete-case analyses.
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	Methods (Study design and participants) p.125	
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	Table 6-1 p.136-137	Mean BMI at 4, 8 and 13 years old
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	NA	
		Cross-sectional study—Report numbers of outcome events or summary measures	NA	
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	Results p.129-131 Tables 6-2 to 6-4 p.138-140	Adjusted estimates for direct associations and interaction analyses.
		(*b*) Report category boundaries when continuous variables were categorized	NA	
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	NA	
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	Results p.129-131, Supplementary Tables 6-3 to 6-9	Sex interaction analyses, inverse probability weighting, and energy adjustment sensitivity analysis.
Discussion				
Key results	18	Summarise key results with reference to study objectives	Discussion p.131-132	
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	Discussion p.134-135	
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	Discussion p.131-134	
Generalisability	21	Discuss the generalisability (external validity) of the study results	Discussion p.135	

Other information				
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	p.142	Funding statement