uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Céline Pullen
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.Sc. (Epidemiology)
GRADE / DEGREE

Department of Epidemiology and Community Medicine
FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Dietary Intake and Frequency of Physician Consultations, Infections, and Antibiotic Treatments in the Québec Longitudinal Study of Child Development

TITRE DE LA THÈSE / TITLE OF THESIS

Lise Dubois
DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

Monica Taljaard
CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

Yue Chen

Elizabeth Potter

Gary W. Slater
Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

Dietary Intake and Frequency of Physician Consultations, Infections, and Antibiotic Treatments in the Québec Longitudinal Study of Child Development

Céline C. Pullen

Thesis submitted to the Faculty of Graduate and Postdoctoral Studies in partial fulfillment of the requirements for the MSc degree in Epidemiology

Epidemiology and Community Medicine,
Faculty of Medicine,
University of Ottawa

Your file Votre référence
ISBN: 978-0-494-73821-4
Our file Notre référence
ISBN: 978-0-494-73821-4

Abstract

Many Canadian children are not eating in accordance with the recommendations of Canada's Food Guide. The health effects of such behaviours are widely unknown.

Using the Generalized Estimating Equations technique and data from the Québec Longitudinal Study of Child Development, this study longitudinally explored the effects of low, medium, and high dietary intake of grain products, vegetables, fruits, milk and alternatives, and meat and alternatives on three indicators of child health, namely: the frequency of physician consultations, infections, and antibiotic treatments, in children from age1 to 7.

Using multivariable models, high intake of vegetables (two or more times per day) was found to significantly reduce the risk of physician consultations and antibiotic treatments. Low intake of meat and alternatives (less than once per day) reduced the risk of physician consultations, and high intake of fruits (two or more times per day) reduced the risk of general infections.

(Word count = 147; 150 allowed)

Acknowledgements

I would like to thank my thesis supervisors Dr. Lise Dubois and Dr. Monica Taljaard for their guidance, patience, and expertise. This project has been a tremendous learning experience and I am grateful for the time they dedicated to me and for their continuous gentle encouragements. Their enthusiasm for their areas of expertise is inspirational and motivating.

I would also like to thank the professors in the Department of Epidemiology who gave me the formal training and epidemiological knowledge necessary to complete this project.

Furthermore, this study would not have been possible without the children, families, and researchers involved in the QLSCD as well as the *Institut de la statistique du Québec*. I appreciate the time and effort they dedicated to the study.

Finally, I would also like to thank my family, friends and Jocque, as they never doubted my abilities and supported me both emotionally and with their time throughout this project.

Table of Contents

List of Tables

Table of Figures

List of Appendices

Chapter 1: Introduction

Nutrition is an important part of every child's growth and development. In children, following the nutritional recommendations of "Eating Well with Canada's Food Guide" (2007) ensures that adequate energy, macronutrients (carbohydrates, lipids and proteins) vitamins, minerals and other nutrients are consumed.[1] Many studies have shown that Canadian children are not eating in accordance with this recommendation but the health impacts of such behaviour are largely unknown.

This chapter defines diet quality, discusses the negative health effects of not eating in accordance with Canada's Food Guide, and provides a rationale for studying the relationship between diet quality and health in children.

1.1 Background

1.1.1 Definition and Aetiology of Diet Quality

Diet quality can be defined in many different manners for different purposes. According to a review conducted in 2008 by Arvaniti and Panagiotakos[2], a good measure of diet quality should account for inter-relationships between diet components and also have the ability to estimate specific health outcomes. For instance, analyzing a single nutrient, food, or food group does not provide a realistic account of diet quality because single nutrients and food groups are not consumed by themselves in a typical diet.[2] Essentially, dietary indices can be based on the intake of nutrients, food or food groups, or a combination of both.[2]

This study assessed diet quality on the basis of frequency of food groups consumed, and these food groups were grouped similarly to the recommendations of Canada's Food Guide.

Canada's Food Guide has been in existence since 1942 and has undergone modifications to reflect new healthy eating knowledge and changes of the population.[1] The purpose of the food guide was to act as an educational tool to "help people follow a healthy diet".[1] The recommended servings per day for children in the most recent version of Canada's food guide, Eating Well with Canada's Food Guide (2007) are: four servings of vegetables and fruit for children aged 2 to 3 and five servings for children aged 4 to 8, three servings of grain products for children aged 2 to 3 and four servings for children aged 4 to 8, two servings of milk and alternatives for children 2 to 8, and one serving of meat and alternatives for children aged 2 to 8.[1]

1.1.2 Canadian Prevalence Estimates

The Canadian Community Health Survey gathered specific information on the nutritional habits of Québec children from ages 4 to 8 in 2004. This study revealed that 58% did not consume the recommended servings of vegetables and fruit, 22% did not consume enough grain products, 40% did not consume enough milk and alternatives, and 13% did not consume enough meat and alternatives.[3] Other studies also revealed that a high proportion of Canadian children are consuming fewer servings than recommended in Canada's Food Guide.[3-9]

1.1.3 Diet and Health

Eating well in adults has been associated with several well known health benefits such as less overweight and obesity, cardiovascular diseases, cancer, and all-cause mortality.[10-12] The World Health Organization states that "poor nutrition can lead to reduced immunity, increased susceptibility to disease, impaired physical and mental development,

and reduced productivity".[13] Furthermore, the World Health Organization specifies the importance of essential vitamins and minerals such as Vitamin A, zinc, iron and iodine in healthy development and boosting immunity.[13]

The benefits of eating selected food groups have been noted in the American 2005 Dietary Guidelines Advisory Committee Report[14], a report used to build the Dietary Guidelines for Americans. Concerning grain products, adults consuming at least three servings of grain products have a reduced risk of coronary heart disease and diabetes.[14] Furthermore, consuming adequate grain products is known to help maintain a healthy weight.[14] Adults consuming adequate vegetables and fruits may reduce the risk of cardiovascular diseases, and may improve and control established diabetes, hyperlipidemia, and blood pressure.[14] This report also found that adequate intake of milk products in adults is protective against osteoporosis.[14] Eating according to the American recommendations (which are similar to Canada's Food Guide recommendations) shows some distinct health benefits in adults.

In children, following the nutritional recommendations of Canada's Food Guide ensures adequate vitamins, minerals and other nutrients are obtained which contribute to overall health and development.[1] It is known that the majority of nutrients such as vitamins and trace minerals are generally abundant in vegetables and fruits.[15] It has been found that Canadian children are not consuming enough of the recommended food groups, such as vegetables and fruit, and too much 'other' foods (foods not in the four recommended food groups of Canada's Food Guide), which tend to be higher in fat and/or calories, sugar, and are generally nutrient poor. Two studies found that on average, one quarter of the daily energy intake of Canadian adolescents comes from the 'other' food category.[4,16] Diets consisting of a high consumption of the 'other' food category are associated with high

energy intake and low intake of vitamin A, B6, B12, folate, calcium and magnesium, iron, and zinc.[17]

Undernutrition has distinct effects on the immune system, and has been linked to increased susceptibility to infection, particularly under extreme undernutrition circumstances. For instance, it has been shown that deficiencies in vitamin A in children may cause greater susceptibility to respiratory diseases, diarrhea, severe measles and high mortality in general.[13,18-23] Also with vitamin A deficiencies, mucosal linings tend to be weaker, which increases susceptibility to pathogens that infect via the eye, respiratory and gastrointestinal tracts.[23,24] A review conducted by Douglas et al.[25] demonstrated that Vitamin C supplementation was associated with reduced duration and severity of common cold symptoms. It has been seen that having insufficient folate is associated with decreased resistance to infections, especially in elderly people.[26,27] A review conducted by Ekweagwu et al.[23] indicated that zinc is important in the proper functioning of the immune system, and also that iron deficiency in malrourished children increases susceptibility to infections. And finally, selenium is a trace mineral only required in small amounts but is known to play a role in the immune system.[23] Overall, these studies have shown that certain vitamins and nutrients are associated with the immune system and that deficiencies can be associated with increases in susceptibility to certain infectious diseases.

Although these extreme examples of nutrient deficiencies leading to increased infection exist under extreme undernutrition circumstances, it is unknown if poor diet quality, like many Canadian children experience, also has a negative effect on the immune system.

1.2 Rationale

Children who eat a well balanced diet, consisting of adequate servings of food from each of the four food groups recommended in Canada's Food Guide, should receive adequate nutrients, have adequately functioning immune systems, and generally should not be at risk of having poor health outcomes due to poor diet quality. Since many Canadian children are not consuming enough servings from each food group, especially vitamin and trace mineral rich vegetables and fruits, it can be hypothesized that they will have weakened immune systems. This weakness thereby may cause the children to be less healthy and more susceptible to general infection, have a higher intake of antibiotic treatments, and therefore consult physicians more frequently.

There are very few studies involving dietary intake of the recommended food groups in western countries and health outcomes in children. Studies of diet quality in adults generally focus on chronic diseases and little is known about its effect on infectious diseases. Indicators of health outcomes such as frequency of physician consultations, general infections, and antibiotic treatments, although all surrogate measures of health, may provide some indication of the health status of a child.

If poor dietary intake of the recommended food groups does indeed have adverse effects on surrogate health outcomes such as frequency of physician consultations, infections, and antibiotic treatments, it could provide a foundation for nutritional policies and also strengthen the evidence for children to eat according to Canada's Food Guide as recommended. For example, educational public health programs promoting proper nutrition in children could be strengthened, which could in turn decrease health care burdens.

Chapter 2: Literature Review

Diet quality has the potential to impact the health of children. In this study, the health outcomes considered are physician consultations, general infections and antibiotic treatments (see section 4.3.1 for justification). Therefore, this chapter will consider both a literature review and common correlates of these specific health outcomes.

2.1 The Relationship Between Diet Quality and Physician Consultations, Infections, and Antibiotic Treatments: Critical Appraisal of the Literature

A literature search followed by a critical appraisal of the resulting studies was conducted to describe what is currently known about the relationship between diet quality and physician consultations, infections, and antibiotic treatments. The following section describes the search strategy and the results of the appraisal process.

2.1.1 Literature Search Strategy

Inclusion and exclusion criteria were formulated to ensure that only relevant articles would be included in the results. Studies were included if they examined the following relationships: the relationship between diet quality and physician consultations, between diet quality and general infections, or between diet quality and antibiotic treatments. Studies were required to have diet quality as a measure of food group or similar grouping. Studies on children aged 0 to 18 were included, and studies were excluded if they were published earlier than 1999 to ensure that the literature found was current. Only human studies, and studies in French or English were included. The search was not restricted to study type. Finally, studies performed in developing countries that involved malnutrition were excluded. This restriction

is justified as diet and nutrition in developing countries consists mostly of malnutrition, which may not be comparable to diet and nutrition in western countries.

A search strategy was created using the most relevant search terms and by searching the most appropriate databases, all while taking into account the inclusion and exclusion criteria noted above. MeSH searches for all terms was performed in order to obtain the best suited search terms, headings, and subheadings. The search strategy for MEDLINE can be found in Appendix A. Slight modifications of the MEDLINE search strategy was used for the following databases: EMBASE, CAB Abstracts and CINAHL. MEDLINE's search resulted in 1,666 studies, EMBASE was 1,237, CAB Abstracts was 288, and CINAHL uncovered 329 studies for a total of 3,520 studies. All 3,520 studies were moved into RefWorks and exact duplicates were removed leaving 3,354 studies. Note that not all duplicates were removed at this stage as in some cases the same article was published in a different journal and was not identified by RefWorks as an exact match.

All 3,354 studies were screened by looking over the title and abstract and 73 articles were retained for closer investigation. The majority of the excluded articles involved foodborne illnesses, malnutrition in developing countries, considered nutrition in terms of individual nutrients, and used dietary intake measures that did not involve a food group type measure. This search resulted in no articles that studied the relationship between diet quality and physician consultations, two articles which considered the relationship between diet quality and general infections, and one article that examined the relationship between diet quality and antibiotic use. The references and citations of any articles that met or were close to meeting the inclusion and exclusion criteria were assessed to identify other potential relevant articles, but this search resulted in no additional articles.

Since this search resulted in so few articles, the search was expanded to include studies published since 1989 and also adult studies using the method described above with a date and age modification. This expansion resulted in one additional relevant study involving diet quality and infection.

The literature search was updated as of April 2010 by screening email notifications of all new articles that fit the search criteria since the initial search. All articles were scanned for relevance but no articles fit the search criteria.

2.1.2 Results and Discussion

Table 1 to Table 3 summarizes the results of the four relevant articles. One study was of Australian children, one of Chinese children, one of Japanese children, and the final study concerns pregnant women. Three studies had sample sizes which ranged from 814 to 2,288 participants, and one study was a small exploratory intervention study that had sample size of only 15. Two studies used logistic regression analysis, one used Cox proportional hazard models and these three studies controlled for different confounders. None of the studies were longitudinal in nature.

No studies were found which analyzed the relationship between diet quality and physician consultations.

Two studies assessed the relationship between diet quality and infection in children (Table 1). He et al.[57] conducted a study which considered fruit, vegetable and milk consumption of 2,228 children in China and their respiratory health. A food frequency questionnaire completed by the mother was used to assess the child's consumption of fresh fruit, leafy vegetables, root vegetables, and milk.[57] The study found a decrease in odds of colds with phlegm if milk is consumed one or more times per day compared to less than once

per week.[57] Additionally, a decrease in odds of cold with cough is seen if root vegetables are consumed one or more times per day (compared to less than once per week), and also if milk is consumed more than once per day or one to six times per week compared to less than once per week.[57] A limitation of this study may be the uncertainty in extrapolating the results to a western country's population. Another limitation may be that other food groups such as grain products and meats and alternatives were not considered.

A study concerning dietary intake and mild and severe influenza infections (Table 1) was conducted by Hirota et al.[58] in 1992. This study consisted of 814 school children aged 6 to 12 in the city of Kasuga, which has a high population density.[58] The study was conducted immediately after the peak influenza season by means of parental questionnaires.[58] This study considered the dietary frequency of the following foods: meat products, fish products, eggs, dairy products, green/yellow vegetables, and other vegetables and fruits.[58] This study found decreased odds of mild influenza with a frequent intake of vegetables or fruits other than green/yellow vegetables (greater than or equal to once per day compared to less than three or four times per week) with an odds ratio of 0.5 (95% CI: 0.3, 1.0).[58] The classification of 'green/yellow vegetables' was intended to represent vegetables and fruits with high vitamin A, and the 'other vegetables and fruits' was intended to represent vegetables and fruits high in vitamin C.[58] The study also found a decreased odds of severe influenza with frequent intake of dairy products (greater than or equal to once per day compared to less than three or four times per week) with an odds ratio of 0.3 (96% CI: 0.1, 0.6).[58] A strength of this study is its high response rate of 98.6%, making non-response bias negligible.[58] A potential weakness of the study is it did not control for any indicators of socioeconomic status with the exception of room space per capita being controlled for only

in the mild influenza model. Furthermore, due to the study's age and also its location being in Japan, it is unknown if the results can be extrapolated to children in Canada.

The one adult study, conducted by Li et al.[59], was a retrospective cohort of 1,034 pregnant women from 1996 to 2002 in Canada and the USA (Table 2). This study used a semi-quantitative food frequency questionnaire to collect information on vegetable and fruit consumption during the six months before pregnancy.[59] The number of upper respiratory infections the women acquired by the three month follow-up and the six month follow-up time point during pregnancy was collected.[59] Cox proportional hazard models were created for vegetables and fruits, fruits alone, and vegetables alone for the two data collection times.[59] Li et al.[59] found that women at the five month follow-up period who were in the highest quartile of vegetables and fruit consumption had a moderate reduction in risk of upper respiratory tract infections with a hazard ratio of 0.61 (95% CI: 0.39, 0.97).[59] No significant associations were found for the three month follow-up and dietary intake nor in models with vegetables and fruit alone.[59] This study controlled for a wide variety of factors, and was also conducted relatively recently. Limitation of this study may be the uncertainty in extrapolating the results to children and also that the study did not consider different food groups as food groups are generally not eaten in isolation. Furthermore, misclassification of dietary intake could have resulted as women may have changed their dietary habits once they realized they were pregnant and dietary intake was measured pre-pregnancy.[59]

To summarize the above outlined relationships between dietary intake and infection in children, milk consumption was protective against colds with phlegm and colds with a cough, and vegetable consumption was protective against colds with a cough.[57] Dietary intake was also associated with infections as frequent intake of non green or yellow vegetables and fruits decreased the odds of mild influenza, and frequent intake of milk

products decreased the odds of severe influenza.[58] Finally, in pregnant women, high vegetable and fruit consumption was associated with a lower risk of upper respiratory tract infections.[59]

One study considered the relationship between fruit consumption and antibiotic intake (Table 3). An exploratory study was conducted in a small and remote Australian Aboriginal community that implemented an intervention which supplied two fresh fruit snacks per day to 15 school children.[60] There was a pre-intervention average of seven antibiotic prescriptions written per month, which decreased to one per month after six months of the intervention.[60] This study has many limitations such as being small, does not control for any confounders, is not a blinded intervention, did not have a control group, did not report a response rate, and did not mention if the antibiotic prescription rate was self reported or reported by a physician.[60] Due to the large limitations of the study, it is not possible to assess if diet quality will have an effect on antibiotic intake.

2.1.3 Limitations

There were several limitations to the critical appraisal of the literature. The appropriate studies that were published in a language other than French or English could have been missed. Also, grey literature was not searched. Furthermore, it is recommended that two reviewers assess the articles for inclusion and exclusion and in this study and only one person assessed the articles. This could have resulted in studies that fit the inclusion criteria being overlooked.[61] However, two reviewers assessed the construction of the inclusion/exclusion criteria and also the search strategy, which could have helped to reduce bias. It is possible that publication bias exists since the literature review was performed solely on published literature. It is known that studies which have significant effects are

Table 1: Child studies concerning the relationship between dietary intake and infection.

Study	Study design, sample size, and children's age	Dietary intake definition	Type and definition of infection	Statistical method	Statistical adjustments	Main effect of dietary intake on infection
He et al.[57] 2008. Nutrition and children's respiratory health in Guangzhou, China.	Cross-sectional N=2,288 Age=10.07 ± 0.86 years	Food frequency questionnaire for fresh fruit, leafy vegetables, root vegetables, and milk consumption	Parental reported (in last 12 months) 1) phlegm with cold 2) cough with cold	Logistic regression	Age, sex, BMI, allergy status, physical activity, toy composition, district, passive smoke exposure, parental asthma/allergy status, parental education, home contains: pets, mould, new furniture or decorations	Ø for all food groups except: 1) cold with phlegm *Milk*: OR=0.62 (95% CI: 0.39,0.97) if >1/day and OR=0.60 (95% CI: 0.39,0.93) if 1/day (ref=<1/week) 2) cold with cough *Root vegetables:* OR=0.61 (95% CI: 0.42,0.91) if >1/day and OR=0.61 (95% CI: 0.40,0.90) if 1/day (ref=<1/week) *Milk:* OR=0.53 (95% CI: 0.36,0.79) if >1/day and OR=0.58 (95% CI: 0.40,0.85) if 1-6/week (ref=<1/week)
Hirota et al.[58] 1992. Various factors associated with the manifestation of influenza-like illness. Japan	Case-control N=814 Age=6-12 years	Food frequency questionnaire, parental reported intake of: meat products, fish products, eggs, dairy products, green/yellow vegetables, and other vegetables and fruits	Parental report of symptoms classified into: 1) Mild influenza-like illness 2) Severe influenza-like illness Control group was those with no symptoms	Logistic regression	School grade, easily inflamed tonsils, asthma, family size, room space per capita, previous influenza vaccination	1) Mild influenza: increased odds for children with easily-inflamed tonsils OR=3.0 (95% CI: 1.7,5.4), and larger family size OR=1.9 (95% CI: 1.1,3.4). Decreased odds for higher school grade children OR=0.4 (95% CI: 0.2,0.9), frequent intake of vegetables or fruit other than green/yellow vegetables OR=0.5 (95% CI: 0.3,1.0), and large room space per capita OR=0.4 (95% CI: 0.2,0.9). 2) Severe influenza: increased odds for children with easily-inflamed tonsils OR=3.8 (95% CI: 1.8,8.1), and asthma OR=2.9 (95% CI: 1.2,6.7). Decreased odds for higher grades OR=0.2 (95% CI: 0.1,0.6), frequent intake of milk products OR=0.3 (95% CI: 0.1,0.6), and vaccination OR=0.3 (95% CI: 0.1,0.8).

Table 2: Adult studies concerning the relationship between dietary intake and infection.

Study	Study design, sample size, and age	Dietary intake definition	Type and definition of infection	Statistical method	Statistical adjustments	Main effect of dietary intake on infection
Li et al.[59] 2009. Fruit and vegetable intake and risk of upper respiratory tract infection in pregnant women. USA and Canada	Retrospec-tive cohort N=1034 pregnant women Age=13 to 45 (mean age=28 years)	Semi quantitative food frequency questionnaire. This study considered specifically: a) vegetables and fruits b) fruits alone c) vegetables alone	Upper respiratory tract infection: nasal stuffiness or congestion, headache, sore throat, cough, achiness and fever, with duration of no more than six weeks. 1) # for 3 month follow up period 2) # for 5 month follow up period	Cox proportio-nal hazard models used to generate hazard ratios (HR)	Age at pregnancy, race, BMI, marital status, employment status, education, income, vitamin intake, daily energy intake, smoking status, drinking status, number of previous live births, if pregnancy was planned, and season	The hazard ratio (HR) for women in the highest quartile of dietary intake (compared to the lowest quartile): a) vegetables and fruits: 1) 3 month follow up: HR=0.61 (95% CI: 0.39, 0.97) 2) 5 month follow up: HR=0.74 (95% CI: 0.53,1.05) b) fruits alone: 1) 3 month follow up: HR=0.84 (95% CI: 0.53, 1.33) 2) 5 month follow up: HR=0.85 (95% CI: 0.60, 1.20) c) vegetables alone: 1) 3 month follow up: HR=0.98 (95% CI: 0.63, 1.52) 2) 5 month follow up: HR=1.17 (95% CI: 0.84, 1.64)

Table 3: Child studies concerning the relationship between dietary intake and antibiotics.

Study	Study design, sample size, and children's age	Dietary intake definition	Type and definition of infection	Statistical method	Statistical adjustments	Main effect of dietary intake on infection
Jones et al.[60] 2006. Are there health benefits from improving basic nutrition in a remote Aboriginal community. Australia	Exploratory analysis, N=15 Aboriginal students Age = kindergarten to grade 6	Intervention included providing fresh fruit twice daily	# of antibiotics taken per month for skin infections or otitis media	None	None	Went from a pre-intervention average of 7 antibiotic prescriptions written per month down to 1 per month after 6 months of the intervention

published more often than studies which result in null effects, which may lead to an overestimation of the effect.[62]

2.2 Common Correlates of Diet and Physician Consultations, General Infections, and Antibiotic Treatments

This section identifies some common correlates of diet quality, physician consultations, general infections, and antibiotic treatments identified in the literature, and summarizes their relationships in Table 4. Accounting for correlates is imperative as they may influence the relationship between diet quality and health outcomes.

2.2.1 Child Related Variables

2.2.1.1 Sex

Sex is associated with diet quality and health in many studies. A review conducted by Rasmussen et al.[28] in 2006 examined the determinants of vegetable and fruit consumption in children from the age of 6 to 18. This review found that 27 out of 49 papers documented that girls consume vegetables and/or fruit more often than boys.[28] Only 4 papers observed more frequent intake among boys, which was explained as being due to different study geographical locations and methodological bias.[28]

Sex is also related to different health outcomes. A review concerning determinants of children's primary health care found mixed results or no relationship between sex and the frequency of physician consultations.[29] The relationship between sex and general infection is dependent on the type of general infection being considered. For instance, boys had a higher risk than girls of developing a respiratory and otitis infection, but girls had a higher risk than boys of developing urinary tract infections.[30-34] Finally, a review on antibiotic prescription

prevalence rates by Rossignoli et al.[35] indicated that a greater proportion of preschool boys than girls required antibiotics; however, it was noted that the relationship tended to be opposite in older children in some countries.

2.2.1.2 Age

Age influences diet quality, the frequency of physician consultations, the risk of general infection, and antibiotic intake. Concerning diet, it is expected that diet quality is less influenced by parents as the child gets older and eats more independently. A literature review on fruit and vegetable consumption showed a general trend of increasing age and decreasing fruit and vegetable consumption.[28] However, the authors of the review suspected this difference was due to response bias because the majority of the papers that measured food intake via a food frequency questionnaire found a negative association with age, and the papers that measured food intake via a 24 hour recall usually found no age association.[28] The author hypothesized that all age groups eat the same amount of vegetables and fruits, but that younger children may in general, eat them more frequently.[28]

Age also has distinct effects on different health outcomes. In a literature review and several studies, an association is reported between age and physician consultations whereby younger ages are associated with higher numbers of physician consultations.[29,36-38] The same trend was observed for general infections. Younger children tend to have a higher incidence of a wide range of general infections in comparison with older children.[31,32,39] Finally, age is associated with antibiotic intake as children of preschool age received antibiotics more frequently than school aged children.[35,36,40,41] Some potential reasons why younger children experience worse health outcomes include that their immune system is less developed, younger children could be perceived as needing more care by their parents and physicians, or

that younger children generally do not communicate their problems as clearly as older children thereby causing the parent to seek professional medical advice more often.

2.2.1.3 Body Mass Index (BMI)

According to the literature, BMI is associated both with diet quality and infection. Although many factors play a role in weight gain, a sedentary lifestyle in combination with poor diet quality (such as not eating according to recommendations) can lead to an imbalance between energy intake and energy expenditure, thereby, contributing to weight gain.[42-44] BMI's relationship with infection was identified in one study involving 1,129 Polish preadolescents where obese children have twice the odds (OR=2.02, 95% CI=1.13, 3.59) of respiratory infection in comparison to children with lower BMIs.[45]

2.2.2 Parental Related Variables

2.2.2.1 Socioeconomic Status

Socioeconomic status (SES) is an important determinant of health, and is associated with both diet quality and different health outcomes. A review and many studies show that generally, SES has a positive association with diet quality.[28,46-49] Family income, a constituent of SES, also has a positive association with diet quality.[48,49] A review conducted by Taylor et al.[49] which specifically concerned the determinants of healthy eating in children and youth, noted that "food price becomes the most important consideration in food choice when income is restricted, often leading to the selection of foods that are higher in sugar and fat because they are among the least expensive sources of dietary energy".[49] Parental education has a positive association with diet quality in many studies.[28,47-49] Finally, there is a positive association between parental occupation and diet quality.[28]

The relationship between SES and the frequency of physician consultations varied by study. Some studies found no association between SES and physician consultations, whereas others found an association between low SES and lower primary health care utilization.[29,50] Janickle et al.[29] noted in a review that low SES is associated with low primary health care utilization and high use of emergency services. The type of physician appointment is also influenced by SES. Gorman et al.[51] found that the more educated the parents, the more likely the child would attend an annual check up appointment. Associations between parental education and physician consultations varied, as in some instances there were no associations, and others reported increasing visits with increasing maternal education.[29,50]

Regarding infections, low SES, low income, and low parental education are associated with increased common childhood infections such as otitis media and respiratory infection.[31,32,39,52] When considering the effects of SES on antibiotics prescription, higher SES is associated with a decrease in antibiotic intake, and low income is associated with an increase in antibiotic intake.[52,53] The effects of parental education on the frequency of antibiotic prescription is both positive and negative, a variation which may be attributable to the country of the study.[50,52,53]

2.2.2.2 Ethnicity

Race or ethnicity is important when considering diet and health outcomes. Two reviews by Rasmussen et al.[28] and Patrick et al.[48] on diet quality showed that the association between ethnicity and diet quality is generally dependent on the ethnic group being studied.

For primary health care utilization such as physician consultations, race groups other than white use health care less often.[51,54] Concerning general infection, children from ethnic minority groups are at greater risk of developing respiratory illnesses.[31] Some ethnic groups

have an increased risk of ear infections.[32] Finally, two studies showed no significant association between ethnicity and antibiotic intake.[40,55]

2.2.2.3 Family Structure

Family structure (i.e. single parent versus two parent households) is associated with diet quality, as well as health. One review indicated that children from single parent families have lower fruit and vegetable consumption in comparison to children from two parent families.[28]

Family structure is also associated with physician consultations and antibiotic intake. When the child's caregiver is a single father, the child has less physician consultations.[51] A study of 63,054 U.S. children showed that single mothers with low maternal education have better access to health care than two parent families with similar maternal education.[56] A study of 5,024 Danish children showed that the children of single mothers have an increased risk of receiving antibiotics.[52]

2.2.2.4 Living Location

Where the child lives influences both diet quality and health. A review found that in three out of four studies, rural living environment is associated with a higher consumption of vegetables and fruits in children in comparison to living in an urban location.[28]

In a large study of Swedish children, and also in a review done by Janicke et al.[29], rural children consult physicians less often than urban children.[51,56] As for general infections, increased exposure to pollution, such as in a city, increases the risk of otitis media and respiratory illnesses in children.[31,32] And finally, children living in a rural location have a lower risk of taking antibiotics compared to children living in a big city.[50]

All relationships in section 2.2 are summarized below in Table 4.

Table 4: Relationship between various covariates and diet quality, physician consultations, general infections and antibiotic treatments.

	Diet quality (poor)	Physician consultations (high)	Infections (high)	Antibiotic treatments (high)
Child related variables				
Sex (male vs female)	positive	positive, negative, & null	positive & negative	positive
Age (younger vs older children)	negative	positive	positive	positive
BMI (obese vs normal weight)	positive	-	positive	-
Parental related variables				
Socioeconomic status (high vs low)	negative	positive & null	negative	negative
Parental income (high vs low)	negative	positive	negative	negative
Parental education (high vs low)	negative	positive & null	negative	positive & negative
Maternal ethnic origin (immigrant vs non immigrant)	positive & negative	negative	positive	null
Family structure (single parent vs two parents)	positive	dependent on mother's education	-	positive
Living location (urban vs rural)	positive	positive	positive	positive

Chapter 3: Study Goal, Objectives

3.1 Goal

The goal of this study is to analyze the relationship between dietary intake and health in childhood.

3.2 Objectives

The general objective of this study is to explore longitudinally the effects of low, medium, and high dietary intake of grain products, vegetables, fruits, milk and alternatives, and meat and alternatives on three indicators of childhood health, namely: (a) the frequency of physician consultations, (b) the frequency of infections, and (c) the frequency of antibiotic treatments, in children from age 1 to 7. For the definitions of low, medium and high dietary intake of each food group see section 4.3.2.

3.3 Hypotheses

Hypothesis 1: There is an association between frequency of consumption (high, medium, or low) of each of the food groups (grain products, vegetables, fruits, milk and alternatives, and meat and alternatives) over time and frequency of physician consultations.

Hypothesis 2: There is an association between frequency of consumption (high, medium, or low) of each of the food groups (grain products, vegetables, fruits, milk and alternatives, and meat and alternatives) over time and frequency of infections.

Hypothesis 3: There is an association between frequency of consumption (high, medium, or low) of each of the food groups (grain products, vegetables, fruits, milk and

alternatives, and meat and alternatives) over time and frequency of antibiotics treatments.

3.4 Conceptual Framework

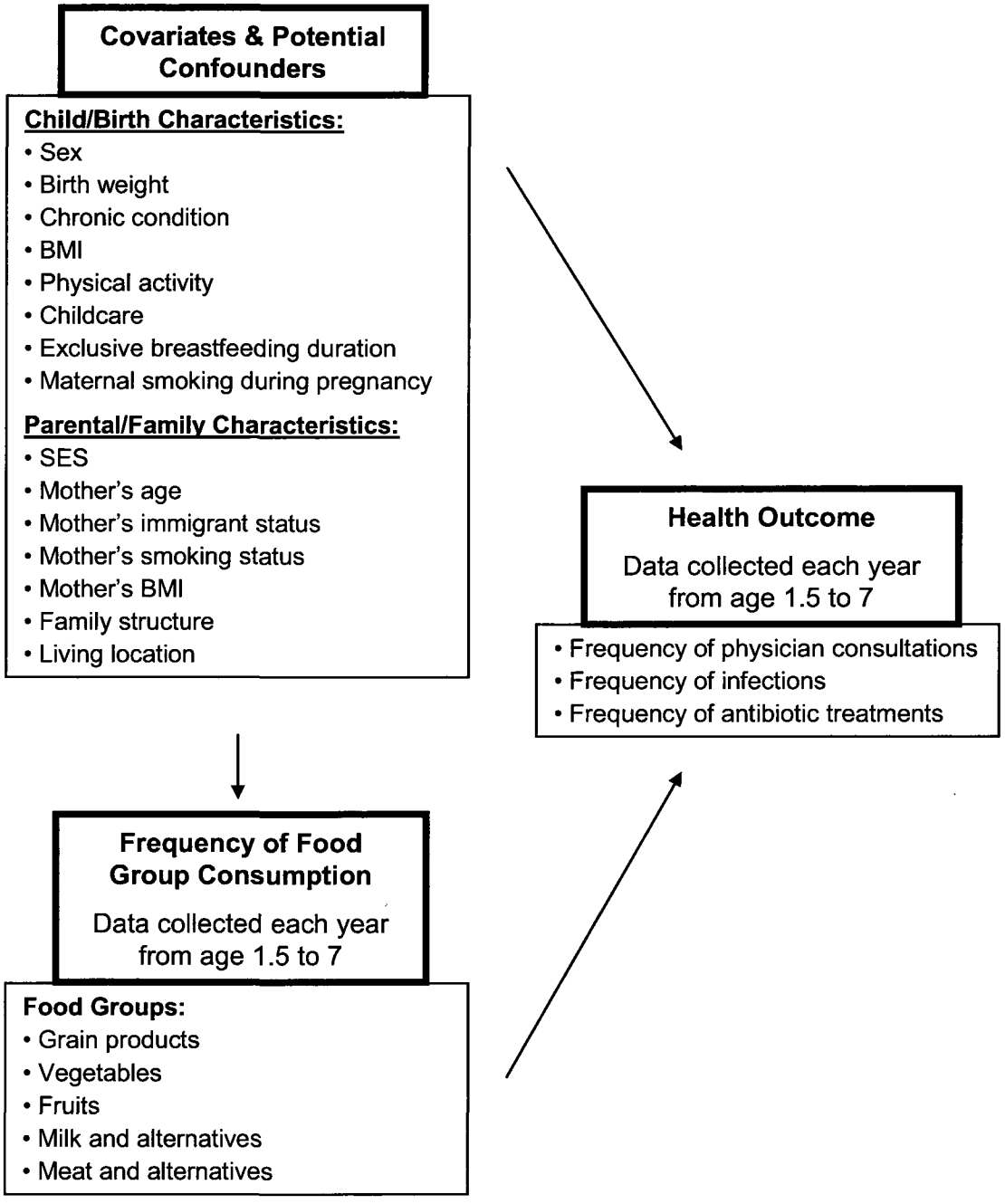
The common correlates and confounders of similar studies found in the literature review in Chapter 2 were searched for in the Québec Longitudinal Study of Child Development (QLSCD) database in order to find similar variables (see section 4.1 for more details on the QLSCD). These variables were considered as covariates or potential confounders, and are presented in Table 5.

Table 5: Potential confounders and covariates.

Child/Birth Characteristics	Prenatal Characteristics	Parental/Family Characteristics
Sex	Exclusive breastfeeding duration	SES
Birth weight	Maternal smoking during Pregnancy	Mother's age
Chronic condition		Mother's immigrant status
BMI		Mother's smoking status
Childcare		Mother's BMI
Physical activity		Family structure
		Living location

Figure 1 represents the simplistic schematic of the relationship of these variables with the health outcomes and exposures to aid in conceptualizing their relationships.

Figure 1: Conceptual framework



Chapter 4: Methods

4.1 Data Source and Design

4.1.1 General Information

The Québec Longitudinal Study of Child Development (QLSCD) was conducted by *Santé Québec* (a division of *l'Institut de la statistique du Québec*) and funded mainly by the *ministère de la Santé et des Services Sociaux* during phase 1 (1998-2002). In phase 2 (2003-2010) the QLSCD had additional funding agents: the Lucie and André Chagnon Foundation, the *ministère de la Famille et des ainés* and the *Institut de la statistique du Québec.*[63] The main objective of phase 1 of the QLSCD was to "identify factors that, coming into play during early childhood, affect the social adjustment and academic performance of young Quebeckers".[64] Phase 2 of the study furthers this objective by helping in "understanding the factors that contribute to academic success in primary school, while taking into account children's life experiences".[64]

4.1.2 Sampling Frame

Singleton babies of gestational age (sum of the length of pregnancy and the age of baby) of 59 or 60 weeks were sampled using the master birth registry of Québec (*Fichier maître des naissances*).[63] Birth registries without information pertaining to the duration of pregnancy were excluded (1.3%), as well as babies with missing sex information (0.1%), and babies with a gestational age of less than 24 weeks (premature) or greater than 42 weeks (0.1%).[63] Since babies were sampled at the beginning of the data collection periods throughout the year, inconsistencies with the birth registry (such as not being completely up to date at the time) caused some under-coverage of the target population.[63] Also, an under-

coverage of 0.6% was caused by the logistics surrounding the duration of pregnancy in conjunction with the requirement of babies born after October 1, 1997 as to ensure that all the children in the sample would enter school the same year.[63] To be eligible for the study, the baby's mother at the time of the study must have lived in Québec, and not have been living in northern Québec, Cree territory, Inuit territory, or in Indian reserves.[63] These exclusions represented 2.1% of the target population. Overall, it was anticipated that 94.5% of the target population was included in the sampling frame.[63]

4.1.3 Sampling Method

From 1999 to 2001 the measurement method ensured that each child was approximately the same age when the initial measurement was taken and then measured exactly 12 months later until 2001.[63] This was performed by measuring the children in five or six week 'waves' that occurred in groups of three. The child was assigned a particular wave depending on their age.[63] In each year of the study there were three waves between March and June, and another three waves between September to December.[63] From 2002 to 2005 the survey was conducted every spring.[65]

4.1.4 Sampling Design

A stratified three-stage sampling design was used in the implementation of the QLSCD.[63] The area covered by the QLSCD was divided into primary sampling units by region, which were subdivided into four remote and eleven non-remote groups.[63] In stage 1 (1998-2002), all eleven non-remote regions were chosen and two out of the four remote regions were chosen.[63] In stage 2 (2003-2010), second stage units were made by dividing the regions above by one or two county regional municipalities.[63] These second stage units were

then divided into either high 1996 birth numbers or low 1996 birth numbers.[63] A fixed amount was sampled from the low birth number group and all were sampled in the high birth number group.[63] Finally, for stage 3, infants were selected from the secondary stage units.[63] A mean design effect of 1.3 was anticipated for the study.[63]

4.1.5 Survey Tools

In phase 1 of the study (1998-2002) data collection and longitudinal monitoring was completed using many tools. In the first year of the study (1998) the tools included (in order): a letter of introduction, general brochure, souvenir folder, consent form, computerized questionnaire completed by the interviewer (CQCI), paper questionnaire completed by the interviewer (PQCI), authorization form to access the mother's and infant's medical records, "1,2,3 hand games" (imitation sorting task), self-administered questionnaire for the mother (SAQM), self-administered questionnaire for the father (SAQF), baby diary, observations of family life (OFL), and a birthday card for the child.[63] All of the questionnaires were completed by the individual who is the most knowledgeable about the child, usually the mother.[63] For simplification, the person most knowledgeable about the child will be identified as the mother throughout the remainder of this document.

In phase 1 there were four questionnaires used consistently each year; CQCI, ICPQ, SAQM, and SAQF. Various activity tasks which fluctuated with the development and age of the children were also implemented each year.

The CQCI is of particular interest because it was the primary tool of the QLSCD. If the completion of this questionnaire did not occur the participant was considered to be a non-respondent for that year.[63] The majority of the questions on the CQCI were extracted from Statistic Canada's National Longitudinal Study of Children and Youth.[63] The CQCI

questionnaire was administered by an interviewer in person, and took approximately 50 to 60 minutes to complete.[63] The CQCI results consist of approximately 600 variables and covers topics such as lifestyle, sociodemographic and socioeconomic characteristics, living conditions, health status of parents, parenting behaviours, family history, mother's health during pregnancy, the temperament of the child, and the motor development of the child.[63]

In phase 2 (2003-2010) the questionnaires administered varied each year with the exception of the CQCI and the SAQF. Other questionnaires included ones such as a questionnaire to measure the child's physical activity level, various activity questionnaires such as "the test of gross motor development" and also a questionnaire completed by the child.[64]

In 2002, a 24 hour recall interview was conducted on all the children by a trained nutritionist where mothers specified exactly what foods, and the volume of food that the child consumed in the last 24 hours.

4.2 Sample

4.2.1 General Response Rate

The overall response rate for cycle 1 in 1998 was 75.8%.[63] The response rate was defined as the "ratio of the number of responding units (participating households) to the total number of units eligible for the survey (responding and non-responding units and unresolved cases)".[63] Overall, 2,223 units responded to the CQCI, and there were 444 non-responding units, 265 total unresolved cases (such as respondents who were out of province, spoke a foreign language), and eight who were ineligible for the survey (twin, physical handicap, or death).[63]

Initial contact was done through mail, then telephone, and finally by a personal visit.[63] If there was no response, the neighbours were asked to aid in finding the families.[63] If there was still no contact, the child was considered to be a non-respondent.[63]

Several steps were taken to aid in the participation and retention of families in the study. A $20 incentive was offered each year to compensate for the participant's time, as well as a personalized souvenir album at the end of phase one.[66] Furthermore, an annual follow-up was planned with the families and three contacts occurred throughout the year: a letter announcing an upcoming call, newsletters, and a birthday card for the child.[66] Change of address cards were also sent with all the mailings in attempt to maintain contact with the families.[66]

4.2.2 Longitudinal Sample Details and Longitudinal Response Rates

In 1998, 2,223 children were surveyed, but only 2,120 of these children were retained for the longitudinal study as 103 children were removed due to oversampling measures to assess the impacts of the ice-storm that occurred in January of 1998.[63,67] A total of 1,537 children were still participating in 2005 (cycle 8).[68] The first round of the study (done in 1998) will not be analyzed in this project as the children in the study were only 5 months old and may not have been exclusively consuming solid food, therefore, their diets cannot be assessed.

Table 6 depicts the study years considered in this analysis, the ages of the children at each measurement time, and the percent response rates.[65,67-74]

Table 6: QLSCD years, child ages and participation rates.			
Year survey was conducted	Age of child	Number of children who participated in the study	Response rate (%) relative to the initial survey in 1998*
1998	5 months	2,103	75.8
1999	17 months	2045	72.6
2000	29 months	1997	70.9
2001	41 months	1950	69.2
2002	4 years	1944	69.0
2003	5 years	1759	62.3
2004	6 years	1529	54.0
2005	7 years	1537	53.9

*Response rate is determined by participant answering the CQCI

4.3 Measures

4.3.1 Dependent Variables

Available health variables in the QLSCD were frequency of physician consultations, infections, antibiotic treatments, hospitalizations, chronic conditions present, and the subjective assessment of the child's health by the mother. The health of the child according to the mother was not used in this study due to its highly subjective nature. The hospitalization variable was modified to exclude hospitalization due to injuries as injuries are not expected to be associated with diet. This exclusion caused the variable's frequency to be extremely low, particularly in the later cycles, resulting in its inability to be analyzed. The presence of a chronic condition was used as a covariate to control for the effect of having a chronic condition on the remaining health outcomes. Thus, three variables were identified as the dependent variables for this study: frequency of physician consultations, general infections and antibiotic treatments.

4.3.1.1 Physician Consultations

The physician consultation variable was created based on similar questions from the National Population Health Survey, and developed by a project team from the National Longitudinal Survey of Children and Youth.[75] The number of physician consultations was asked to the mother from cycle 2 to 8 in the CQCI.

The number of physician consultations was derived from adding up the results of a few open ended questions regarding the child's consultations to health professionals in the last year. Specifically, the mother was asked "In the past 12 months, how many times have you seen or talked on the telephone about [child's name]'s physical, emotional or mental health with" and then specified the number of consultations with each of the following: a) a general practitioner, family physician b) a pediatrician c) another medical doctor (such as an orthopedist, or eye specialist). The wording of this question remained the same on all the questionnaires from cycle 2 to 8.

Physician consultation was treated as a count variable in this study, and represents the number of times the child had a physician consultation in the previous year.

4.3.1.2 Infections

The infection variable was derived from summing the results of questions concerning different types of general infections the child incurred. The mother was asked specifically "In the past three months (namely since [insert date], how many times has [child's name] suffered from" and then specified the number of illnesses in each of the following categories: a) "gastro-intestinal infection (gastroenteritis 'stomach flu' lasting one day or more, vomiting and/or diarrhea)" b) "ear infections (otitis)" c) "respiratory infections with fever (cold, flu, pneumonia)" d) "another infection (example: urinary tract infection)". For each type of

infection, the mother could indicate none, one, two, three, or four or more times. Since the frequencies of the 'four or more' category were very low, these responses were recoded as four times, which allowed us to sum the responses across multiple items in order to create a count of the total number of infections.

Over the years, there were slight changes in the wording of the questions regarding infection. From cycle 6 to 8, the parents were asked to distinguish between "bronchitis or pneumonia" and "cold, flu, pharyngitis or laryngitis" as opposed to only one question from cycle 2 to 5 that asked simply for "respiratory infection with fever (cold, flu, pneumonia)'. In this study these two questions were collapsed into the 'respiratory infection' category. Another change in the questionnaire was in cycle 6 and 7 where the parents were asked to specify if their child had a cutaneous infection. In cycle 8 the 'other' infection category was changed to include cutaneous infections as an example of an 'other' infection. In this study, the cutaneous responses were collapsed into the 'other infection' category to be consistent.

This time-varying outcome was treated as a count of the number of general infections the child incurred in the past three months.

4.3.1.3 Antibiotic Treatments

For cycles 2 and 3, the number of antibiotic treatments was derived from the question: "In the past six months (namely since last [date inserted]), how many times has [Child's name] taken antibiotics?" The answers were categorical as none, one, two, three, or four or more times. Since the frequencies of the 'four or more' category were very low, responses in this category were classified as 'four times'.

From cycles 4 to 8 the format of the question was expanded to include two more possible choices: "one or more long-term (more than a month) antibiotic" and "continuous treatment".

Along with these two additional responses, an explanation was included. This explanation distinguished between an antibiotic treatment as "an antibiotic treatment generally lasts less than 15 days", and long-term treatments which "have a duration of more than a month without interruption". Since the two additional categories from 2001 to 2005 concerning long term antibiotic usage and continuous treatment reflect heavy antibiotic usage, they were categorized with the 'four times' category for the analysis. This was necessary as the question was not asked in 1999 or 2000 and also because there were too few children who were on long-term or continuous antibiotic treatments (ranged from 0.14% to 0.47% of children over cycle 4 to 8). It is important to note that as of 1997 it was mandatory that everyone in Québec possesses drug insurance, therefore, the cost of antibiotics should not influence rates.[76]

The number of antibiotic treatments was treated as a count variable representing the number of antibiotic treatments taken in the last six months.

4.3.2 Independent Variable: Dietary Intake

The child's dietary intake was obtained through Food Frequency Questionnaires administered to the mother. The results were compiled into the following food categories: grain products, vegetables, fruits, milk and alternatives, and meat and alternatives. The answers indicated how many times per day the child ate from each food group. This information was used to create a variable with three categories that were easily interpretable and applicable to real life. An example of easily interpretable categories for a variable would be: "under one time per day", "one to two times per day" and "over two times per day". Each dietary pattern variable may have a different range (such as the lowest category being under two times per day rather than under one time per day) as the creation of the categories were

also data driven to ensure that each category had adequate frequencies in order to perform analyses.

The data concerning the child's dietary intake represents the frequency with which the child consumes a particular food group, and not the amount or number of servings of food consumed. The Food Frequency Questionnaire method is frequently used to collect information regarding how often foods are consumed.[77,78]

Note that Canada's Food Guide specifies four food groups in which vegetables and fruits are grouped into one category whereas in this study, fruits and vegetables are assessed on their own to determine their unique relationships with the health outcomes.

Below is a description of each food group's categories for each year of data collection:

- Grain products: low is < 2 times per day, medium is ≥ 2 to < 3, and high is ≥ 3 times per day.
- Vegetables: low is < 1 time per day, medium is ≥ 1 to < 2, and high is ≥ 2 times per day.
- Fruits: low is < 1 time per day, medium is ≥ 1 to < 2, and high is ≥ 2 times per day.
- Milk and alternatives: low is < 2 times per day, medium is ≥ 2 to < 4, and high is ≥ 4 times per day.
- Meat and alternatives: low is < 1 time per day, medium is ≥ 1 to < 2, and high is ≥ 2 times per day.

Although the food groups used in this study do not measure serving sizes, the frequency with which a child consumes a particular food is associated with serving sizes. The number of servings of each food group gathered from a 24 hour recall survey (children

aged 4 years) was compared to the results of the Food Frequency Questionnaire and resulted in an 80% or greater correspondence for grain products, vegetables and fruits, and milk and alternatives.[79] For meat and alternatives there was a 62% correspondence between servings and food frequency.[79] This association between servings and food frequencies depicts that responses obtained from the food frequency questionnaire can provide valuable information.

4.3.3 Covariates

As this study is longitudinal in nature, the year variable is time-varying. Year 0, the baseline year, reflects the children in cycle 2, which is when the children were 1.5 years of age in 1999. Depending on the model and the graphical exploration of observed responses over time, year is treated either as categorical (i.e., fitting an arbitrary trend) or continuous (i.e., fitting a parametric trend over time).

Child's **sex** and **birth weight** are accounted for throughout the study. Birth weight, obtained from hospital records, was dichotomous: low birth weight (under 2.5 kg), and normal and high birth weight (2.5 kg or more).

Chronic condition (a time-varying dichotomous variable) includes one of the following which lasted or is expected to last more than six months: bronchitis, heart condition or disease, epilepsy, asthma, cerebral palsy, kidney condition or disease, mental handicap, or any other long term condition.

The child's **BMI** was calculated according to Cole's BMI criteria using the child's height and weight reported by the mother, and is a time-varying variable.[80] This variable has the following categories: normal and under weight (referred to in this study simply as 'normal weight' for simplification), overweight, and obese.

The QLSCD questions concerning **childcare** were derived from the Canadian National Child Care Study and formulated by the National Longitudinal Survey of Children and Youth project team.[75] Due to the nature of the child's daycare changing year to year, the variable is treated as time-varying. This variable has the following categories: the child is cared for at home (could be by parent(s), relative or anyone else), the child is in a daycare center, or the child is cared for outside of home (including a family member's home).

The **physical activity** level was measured from cycle 3 to 7. It was determined by asking the mother "In your opinion, how physically active is [child's name] compared to other children of the same age and sex?". This variable is time-varying and dichotomous, and indicates if the child has either a 'low or medium', or 'high or very high' activity level.

The **breastfeeding** variable indicates whether or not the child was breastfed exclusively for at least three months. Although Health Canada recommends exclusive breastfeeding for six months, there were too few children that were breastfed exclusively for that duration, hence exclusive breastfeeding in this study was cut at three months.[81]

Socioeconomic status (SES) was defined based on the parental education, work prestige and income level.[75] This is a categorical variable based on tertiles and is time-varying.

Mother's age, was collected in cycle 2 and is categorized as: 'less than 25 years', '25 to less than 30 years', '30 to less than 35 years', and '35 years old and over.'

The **mother's immigrant status** is a time fixed dichotomous variable indicating whether or not the mother is an immigrant. The QLSCD's questions regarding immigration were derived from the 1991 Census of Canada.[75]

The **mother's smoking status** is a time-varying dichotomous variable indicating whether or not she was a smoker at the time of the interview.

Mother's BMI categories are in accordance with Health Canada and consist of underweight or normal weight (BMI of under 25 kg/m^2, referred to in this study simply as 'normal weight' for simplification), overweight (BMI from 25 kg/m^2 to under 30 kg/m^2) and obese (BMI of 30 kg/m^2 or over).[82] This variable is based on reported height and weight data that was collected in cycle 2.

The child's **family structure** is a time-varying categorical variable consisting of: intact family (living with two biological parents), single parent family, and reconstituted family (consists of a couple with at least one child not born to them).

Living location describes if the family lived in a rural or urban area during cycle 1.

4.4 Statistical Analysis

All statistical analyses were conducted using Statistical Analysis Software (SAS, version 9.2). All tests were set at a significance level of $p \leq 0.05$.

4.4.1 Missing Data Analysis

In longitudinal studies, it is of the utmost importance to assess to the extent possible whether or not participants who have missing values are different than those who do not have missing values.[83] In studies where health outcomes are assessed, it has been shown that bias can result from ignoring the missing values as participants who are ill may be more likely to be missing or drop out of the study.[84] Consequently, an in depth assessment of intermittent missing data and dropouts was conducted.

4.4.1.1 Multivariable Missing Outcome Analysis using Logistic Regression

Each missing outcome in cycle 8 was assessed in a multiple logistic regression model with the missing indicator variable being treated as the dependent variable. This allowed the

assessment of which variables and exposures were significantly associated with the prevalence of missingness in cycle 8. Modeling all variables together is a more realistic approach than a simple bivariable approach, as the variables will always be used in combination with other variables. To account for any potential bias due to missingness, any variables that were significantly associated with missingness were included in the outcome analyses regardless of whether or not they were significant predictors. This implies the assumption of 'missing at random'.[85]

4.4.1.2 Longitudinal Dropout Analysis

As this study is longitudinal, a high amount of attrition was expected. According to Weiss[86], intermittent missingness in longitudinal data are often thought to be ignorable, as it could result from the participant being on vacation or being too busy one year. Conversely, dropouts may need to be accounted for because participants who drop out are thought to be different than those who stay in the study.[86]

Dropping out of the study was assessed by the creation of a categorical 'pattern' variable which represented each dropout pattern as a separate category. A graphical method suggested by Song[87] depicts the relationship between the mean profile of those who remained in the study and those who drop out at different time points. Using this method, the mean of the outcomes was plotted by year and by dropout pattern in order to visualize the relationship between dropping out in particular years and the health outcomes.[87]

Song's[87] graphical method was complimented by a slightly modified method described by Stokes et al.[88] where dropout patterns were modeled longitudinally against the heath outcomes to assess whether dropping out of the study was significantly associated with the health outcomes. In this method, the 'pattern' variable was used as a predictor in a

longitudinal model with each health outcome as the dependent variable. If the pattern variable was significant it would suggest that dropouts may be associated with the main outcome. This is referred to as 'missing not at random'.[85]

4.4.1.3 Missing Years of Time-varying Covariates

There is much value in using time-varying covariates in longitudinal models as it accounts for changes in the covariates over time.[86] However, due to the manner in which data are collected in the QLSCD, certain covariates were not collected in particular years. In cases where this occurred, the Last Value Carried Forward (LVCF) method of imputation was used. In the LVCF method, it is assumed that the variable is most likely constant over time; therefore, the response of a subject at a particular time is carried forward to the next time.[89] Similarly, the Next Observation Carried Backwards (NOCB) is the same concept but the variable at a particular time is carried backwards. The LVCF imputation method is known to cause bias in situations where LVCF is used to impute the dropouts of clinical trials, however, bias is not expected in this situation as missing covariates in certain years is dependent on the study design, not on the participants.[90,91] For each of the covariates being considered, only a small number of years were missing (one to a maximum of three years) and none of the missing years were consecutive. Consequently, LVCF or NOCB was used on the following covariates for the following missing cycles: chronic condition present in cycles 3, 5 and 8, physical activity in cycles 2 and 8, daycare in cycle 8, and finally, SES in cycle 5.

4.4.2 Descriptive Analysis

Exploratory analyses were performed both cross-sectionally, and longitudinally where appropriate. Descriptive tables, graphs of the health outcome constituents (if

applicable), histograms, and time plots were constructed for each health outcome. Time plots were also created to visualize the relationship between each health outcome and the food group exposures, as well as with all the covariates.

4.4.3 Generalized Estimating Equations

The Generalized Estimating Equation (GEE) modeling approach results in a population-averaged model in which inferences can be made to the total population.[85,92] All health outcomes in this study were modeled using GEE via Poisson regression with a log link function. As recommended by Fitzmaurice et al.[85], robust standard errors were used (also known as empirical or sandwich estimators) as the number of independent subjects was relatively large in comparison to the number of repeated measures. The manner in which time was modeled (i.e., parametric trend (linear or quadratic with time coded as continuous) or arbitrary trend (with time coded as categorical)) was determined by visual inspection of the time trends in the graphs generated during the exploratory analysis.

The sample size was ample, which allowed for the use of an unstructured correlation matrix to maximize the flexibility of the correlation structure throughout the bivariable analysis. For the multivariable analyses three different covariance structures were evaluated: unstructured covariance, exchangeable correlation structure, and the first order auto-regressive structure. The most appropriate structure was determined by considering the following factors: the sample size and whether or not it was large enough to handle a more complicated structure, a visual inspection of the correlation matrix pattern for trends over time, and finally, the QIC (quasilikelihood under the independence model information criterion) number as the model with the smallest QIC number is usually the best fit.[93]

Models were reduced using backwards elimination where covariates with $p \leq 0.05$ were retained. Higher order interaction terms were removed from the model first.

Interpretation of the model parameter estimates allowed for determination of the influence of each food group's intake on the outcome longitudinally. Statistical significance was set at $p \leq 0.05$ and risk ratios were calculated with 95% confidence intervals. The model included main effects for time, food groups and interactions between time and food groups to allow for the determination of whether changes over time in food intake were associated with the health outcome.

4.4.3.1 Bivariable Analysis

Bivariable analysis was conducted to examine the crude association between each health outcome and the individual food group exposures. The term 'bivariable' is used loosely here to refer to an analysis in which each of the food group variables is included on its own (without the other food group variables) but with the other (non-food group) covariates. These other covariates were selected based on the literature, the results of missing data analysis, and the results of the descriptive analyses.

Bivariable analysis included several steps for each food group.

- **Model 1 - Full unadjusted bivariable model:** included only the food group, year, and a food group interaction term with year. This model allowed the determination of the unadjusted longitudinal association between each food group and the outcome.
- **Model 2 – Full adjusted bivariable model:** same model as above with the addition of all the 'other' covariates. This model allowed the determination of the effect of the covariates on the crude longitudinal association between each food group and the outcome estimated in model 1.

- **Model 3 – Reduced adjusted bivariable model:** This step involved making Model 2 more parsimonious by backward elimination of (a) non-significant covariates as well as (b) non-significant interactions between the food group variable and time; thus, it allowed us to determine whether there is a longitudinal trend over time and whether the trend varies among the high, medium and low consumption groups.

4.4.3.2 Multivariable Analysis

The term 'multivariable' is used to refer to the inclusion of all food group variables in the same model (with or without the 'other' covariates). The multivariable analysis included the same steps as Models 1-3 above, but with inclusion of all food groups simultaneously:

- **Model 4: Full unadjusted multivariable model:** included all the food groups and their interaction terms with time. This model allowed the determination of the longitudinal association between each food group and the outcome, while accounting for the other food groups.
- **Model 5: Full adjusted multivariable model:** all the covariates added to model 4. This model allowed evaluating the effect of the other covariates on the multivariable food group relationships identified in model 4.
- **Model 6: Reduced adjusted multivariable model (final model):** this model is the final model for making inferences and involved making model 5 more parsimonious by backward elimination of non-significant terms (including covariates, food group variables and interactions with time).

4.4.5 Model Assessment

4.4.5.1 Multiple Imputation for Missing Data

Multiple imputation was used as a sensitivity analysis for the final models fitted. Imputation replaces the missing values with likely values based on similar variables.[83,94] In multiple imputation, more than one imputation is calculated for each missing value, which leads to several complete datasets.[83,89] Analysis is then performed on each completed dataset separately and the results are combined using the formulas provided by Rubin.[83,89] Multiple imputation is a highly recommended method as it minimizes bias and produces non inflated standard errors in comparison to other single imputation methods.[83,89,95] The SAS procedure of PROC MI was used to create ten complete datasets using the Markov Chain Monte Carlo method for arbitrary missing data.[96]

It is documented that biased datasets could result from rounding imputed categorical variables.[97] To avoid this bias, original variables (pre-categorized) were used for imputation (where possible) then re-categorized after imputation. Furthermore, the imputation process assumes all variables are normally distributed; therefore, variables were visually assessed for normality using histograms and transformed when necessary before imputation.[96]

Combining the imputed results was performed using SAS's PROC MIANALYZE procedure.

4.4.5.2 Model Assessment using Survey Weights

Using sample weights in longitudinal analysis can be a very complex process. There are certain characteristics about a study to be considered when deciding if weights are necessary. Weights are generally utilized if the study is design based, the results will be used

for official statistics or to influence public policy, or if there is minimal evidence that the model is correct.[98] Since this study is model based, looking for association rather than being used to produce official statistics, and there is sufficient evidence on factors related to the exposure and outcome (thereby all important variables should be accounted for), it is not pertinent to use longitudinal weights.

Statistical weighting adjusts for differences between the sample and the actual population as well as for non-response and post-stratification.[99] If the sample size is large enough, such as in this study, the results will generally be the same as if weighted.[99] Even though the advantages of using weighting as a means of adjustment can be important, it is also known in certain cases to severely increase the variability of the estimated parameters.[99]

It is well documented that if the model based assumptions are correct that the results, whether weighted or not, will be similar.[98,99] If the results are not similar, then the model does not fit the population.[98] Longitudinal weights from the QLSCD were used in a weight statement in PROC GENMOD on each model to ensure the weighted results were similar to the unweighted results.[70]

4.4.5.3 Residual and Cook's D

The GEE method generally relies on graphical approaches to assess model fit and to identify potential outliers.[93] Histograms of the residuals were plotted for each model. If the histogram of residuals yields an approximately normal distribution, it is an indication of adequate model fit.[85,93] This plot was also used to detect potential outlying observations, which were in turn investigated to ensure they were not data input errors nor real outliers.[93] Residuals were plotted against adjusted predicted mean values (adjusted for Poisson distribution models) to assess model fit. A loess smoothed line with a constant mean of zero

and no systematic pattern in the scatter plot indicates good model fit.[85] The residuals were also plotted against all covariates with a smoothed line to ensure the covariates adequately explained the model. Finally, Cook's cluster D was used to assess influential clusters (individual children). Influential children were assessed to ensure that they were not data input errors or outliers.

Pearson's residuals were used in model assessment as adjusted residuals are not produced with these particular longitudinal GEE models. Typically in longitudinal residual analysis, Cholesky's decomposition is utilized to "de-correlate" residuals, however, the GEE method does not allow for this de-correlation option.[94]

Chapter 5: Results

5.1 Characteristics of the Study Sample

There were initially 2,103 children who participated in the longitudinal study during cycle 1. The GEE method permits missing values as long as the value is not missing every year, therefore no subjects were excluded unless they were missing data for each year of the study.[85] Consequently, there were 2,057 children who participated at least once from cycle 2 to cycle 8.

Frequencies and proportions of all independent variables and covariates are presented by year in Table 7. Note that the proportion of children at age 1.5 who are overweight or obese is high, which may be attributed to BMI measurements being not accurate in very young children and also that Cole's BMI calculation methods begin at the age of 2.

5.1.1 Dietary Intake

The proportion of children in each dietary intake category by each food group over time is depicted in Figure 2. The proportion of each grain product consumption category appears to remain relatively constant over the seven years of the study. The proportion of children who ate vegetables less than one time per day appears to increase slowly over the seven years of the study. The proportion of children who ate vegetables two or more times per day seem to be higher in the first year and last two years of the study; and this trend was similar for fruit consumption for the same category (Figure 2). Concerning milk and alternatives consumption, there are few children aged 1.5 who consumed milk and alternatives less than two times per day, but the proportion of low milk and alternative

Table 7: Number and proportion of all variables by child's age.

	Age 1.5		Age 2.5		Age 3.5		Age 4.5		Age 5		Age 6		Age 7	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Independent Variables														
Grain products:														
Low (< 2 times per day)	562	27.5	698	35.0	637	32.7	629	32.4	618	35.2	506	33.9	519	39.3
Med (≥ 2 to < 3 times per day)	918	44.9	844	42.3	848	43.5	828	42.6	784	44.6	588	39.4	467	35.3
High (≥ 3 times per day)	564	27.6	454	22.8	464	23.8	487	25.1	356	20.3	398	26.7	336	25.4
Vegetables:														
Low (< 1 time per day)	232	11.4	395	19.8	398	20.4	441	22.7	389	22.2	355	23.8	334	25.4
Med (≥ 1 to < 2 times per day)	545	26.7	618	31.0	645	33.1	644	33.1	617	35.1	381	25.5	250	19.0
High (≥ 2 times per day)	1266	62.0	982	49.2	904	46.4	858	44.2	750	42.7	756	50.7	730	55.6
Fruits:														
Low (< 1 time per day)	347	17.0	516	25.9	533	27.4	550	28.3	490	27.9	332	22.3	311	23.7
Med (≥ 1 to < 2 times per day)	553	27.1	662	33.2	690	35.4	686	35.3	626	35.7	387	26.9	269	20.5
High (≥ 2 times per day)	1141	55.9	818	41.0	726	37.3	706	36.4	640	36.5	773	51.8	732	55.8
Milk and alternatives:														
Low (< 2 times per day)	36	1.8	123	6.2	225	11.5	207	10.7	205	11.7	201	13.5	266	20.1
Med (≥ 2 to < 4 times per day)	1177	57.6	1157	58.0	1129	57.9	1177	60.6	1085	61.7	865	58.0	743	56.2
High (≥ 4 times per day)	831	40.7	716	35.9	595	30.5	560	28.8	468	26.6	426	28.6	313	23.7
Meat and alternatives:														
Low (< 1 time per day)	780	38.2	917	45.9	953	48.9	944	48.6	478	27.2	580	38.9	436	33.0
Med (≥ 1 to < 2 times per day)	1116	54.6	983	49.3	923	47.4	905	46.6	1097	62.4	781	52.4	674	51.0
High (≥ 2 times per day)	148	7.2	96	4.8	73	3.8	95	4.9	183	10.4	131	8.8	212	16.0
Child/Birth Characteristics														
Sex:														
Male	1043	50.7												
Female	1014	49.3												
Birth weight:														
< 2.5 kg	67	3.3												
≥ 2.5 kg	1965	96.7												
Chronic condition:														
Yes	244	11.9	-	-	116	5.95	-	-	202	11.5	200	13.4	-	-
No	1801	88.1	-	-	1834	94.1	-	-	1557	88.5	1292	86.6	-	-
BMI (by Cole definition):														
Normal weight	1041	58.8	1479	79.2	1532	83.6	1530	82.0	959	77.5	975	86.0	1249	84.7
Overweight	396	22.4	281	15.1	230	12.6	255	13.7	172	13.9	104	9.2	155	10.5
Obese	334	18.9	107	5.7	70	3.8	80	4.3	106	8.6	55	4.9	70	4.8

(continued on the next page)

Table 7 *(continued)*

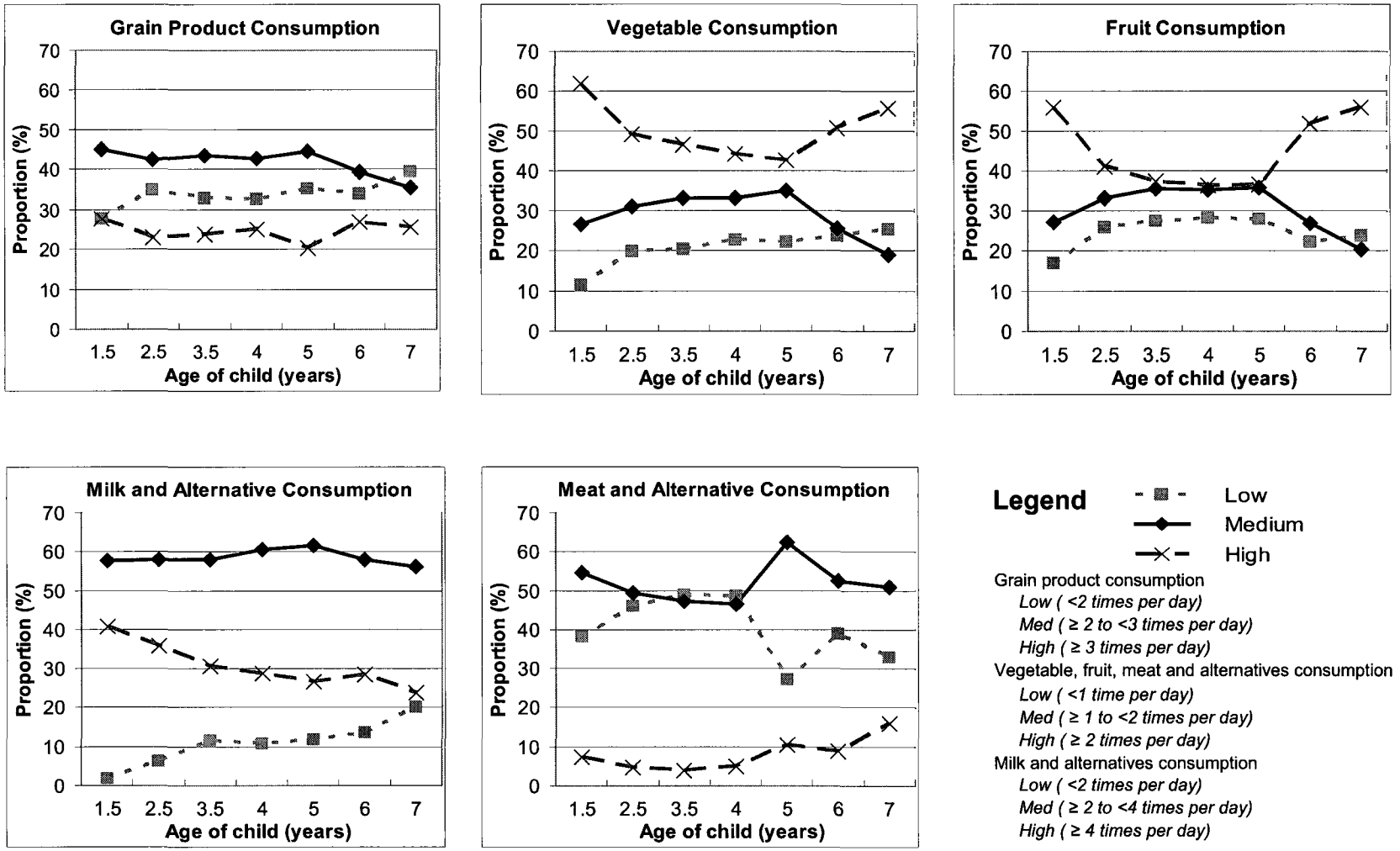
	Age 1.5		Age 2.5		Age 3.5		Age 4.5		Age 5		Age 6		Age 7	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Physical activity:														
Low or medium	-	-	1288	64.6	1346	69.0	1394	71.7	1248	71.0	995	66.8	-	-
High	-	-	707	35.4	605	31.0	550	28.3	511	29.1	495	33.2	-	-
Child care:														
Cared for at home	1051	51.4	854	42.8	556	28.5	569	29.3	413	24.3	559	37.5	-	-
Daycare center	195	9.5	339	17.0	573	29.4	658	33.9	662	39.0	711	47.7	-	-
Cared for outside of home	799	39.1	802	40.2	821	42.1	714	36.8	623	36.7	222	14.9	-	-
Breastfed exclusively for ≥ 3 months:														
Yes	520	25.3												
No	1537	74.7												
Parental/Family Characteristics														
SES (tertiles):														
Lowest	611	30.1	597	30.2	576	29.8	-	-	504	28.9	426	28.7	455	29.9
Middle	681	33.6	657	33.3	653	33.8	-	-	597	34.2	512	34.6	508	33.4
Highest	737	36.3	720	36.5	702	36.4			646	37.0	544	36.7	558	36.7
Mother's age:														
< 25 years	345	16.9												
≥ 25 and < 30 years	605	29.6												
≥ 30 and < 35 years	713	34.9												
≥ 35 years	379	18.6												
Mother's immigrant status:														
Immigrant	218	10.7												
Non immigrant	1823	89.3												
Mother's smoking status:														
Mother smokes	521	25.6	520	26.1	471	24.3	-	-	283	20.3	-	-	238	18.6
Mother does not smoke	1516	74.4	1469	73.9	1465	75.7	-	-	1114	79.7	-	-	1043	81.4
Mother's BMI:														
Normal weight	1424	70.9												
Overweight	380	18.9												
Obese	204	10.2												
Family structure:														
Intact family	1617	79.3	1544	77.5	1481	76.1	1444	74.5	1272	72.6	1042	70.0	1048	68.9
Single parent family	193	9.5	230	11.5	240	12.3	257	13.3	256	14.6	232	15.6	246	16.2
Reconstituted family	231	11.3	219	11.0	255	11.6	238	12.3	225	12.8	214	14.4	228	15.0

(continued on next page)

Table 7 *(continued)*

	Age 1.5		Age 2.5		Age 3.5		Age 4.5		Age 5		Age 6		Age 7	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Living location during cycle 1:														
Urban	1290	64.2												
Rural	720	35.8												

Figure 2: Proportion of children in each food group consumption category by age.



consumers appears to increase over time. Following the exact opposite trend, the proportion of children in the highest milk and alternative category (four or more times per day) appeared to decrease over time. Finally, the proportion of children who ate high amounts of meat and alternatives (four or more times per day) is low, but appears to increase after 5 years of age.

5.1.2 Physician Consultations

The mean number of physician consultations in the past year decreased over time (Table 8). This sample contained a relatively high amount of overdispersion, however, this was accounted for in the model.[85]

Table 8: Mean, standard deviation and variance of the number of physician consultations in the past year.

Cycle	Age of child (years)	Sample size	Mean physician consultations	Standard deviation	Variance
2	1.5	2044	5.97	5.10	26.04
3	2.5	1995	3.90	4.41	19.45
4	3.5	1950	3.43	4.09	16.75
5	4	1861	2.97	3.20	10.27
6	5	1759	2.86	3.94	15.55
7	6	1316	2.74	2.94	8.66
8	7	1322	2.21	2.29	5.23

The mean number of consultation by type and child's age is presented in Figure 3. Consultations with general physicians are the most common type of visit for all years, followed by consultations with pediatricians.

A histogram of all physician consultation observations is displayed in Figure 4, which shows that the highest frequency of consultations is four.

Figure 3: Mean number of physician consultations in the past year by type and age.

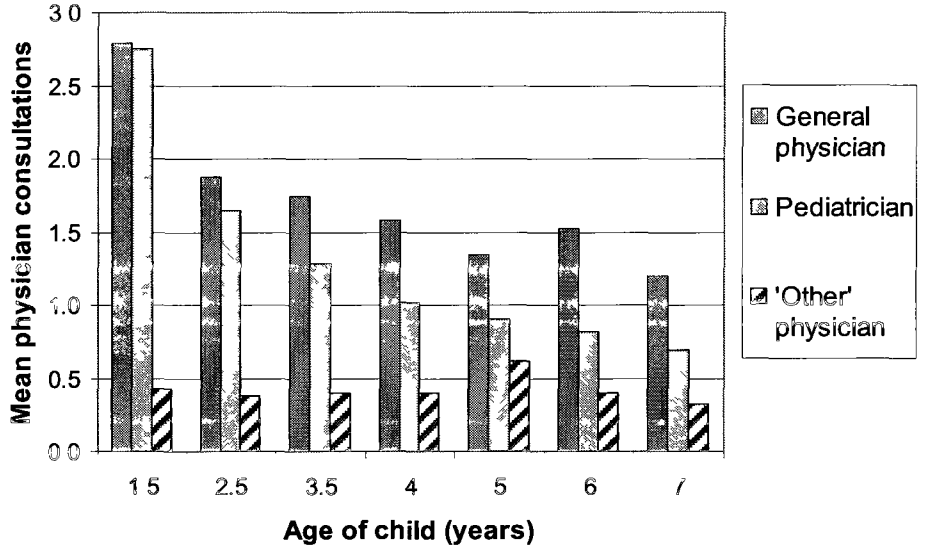
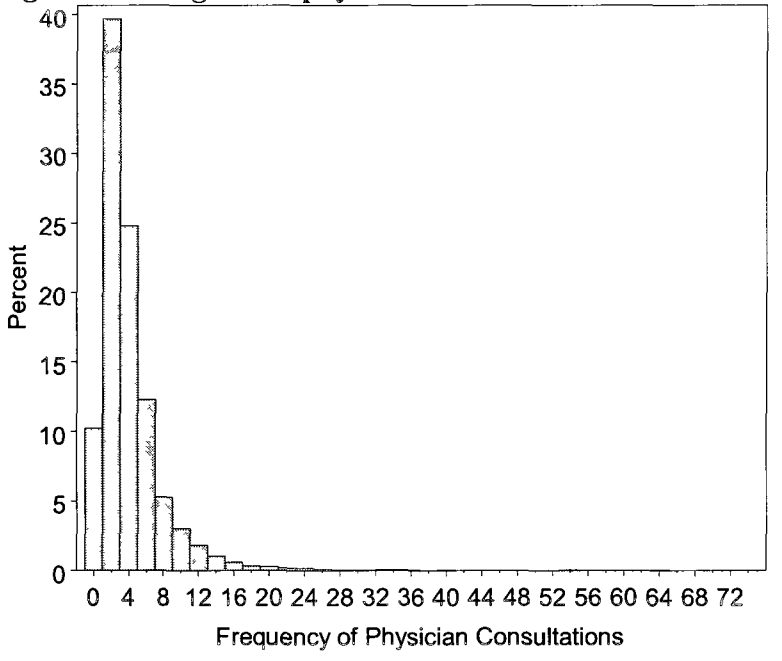


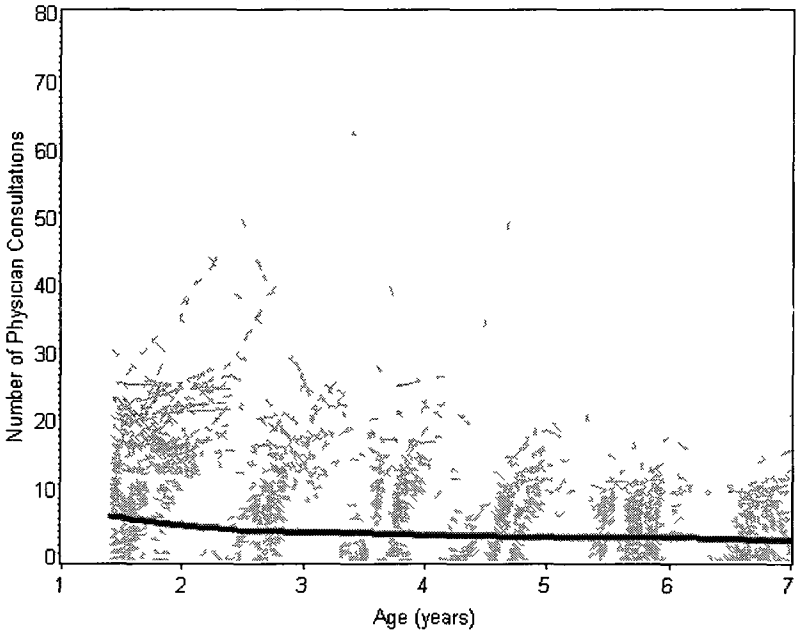
Figure 4: Histogram of physician consultations.



Time plots are used to depict longitudinal trends over time. Each thin line in Figure 5 depicts one child. The thick line is the mean, which shows that younger children have a

slightly higher mean number of physician consultations in the past year followed by a slowly decreasing trend.

Figure 5: Time plot of physician consultation for all children.



5.1.3 Infections

The pattern of the mean number of infections in the past three months over the seven years of the study can be seen in Table 9. The mean number of infections starts at 1.44 in cycle 2 then decreases to 1.15 in cycle 3. This is followed by a rise until the children start to attend school at age 5 (cycle 6), where it peaks then decreases again until the end of the study.

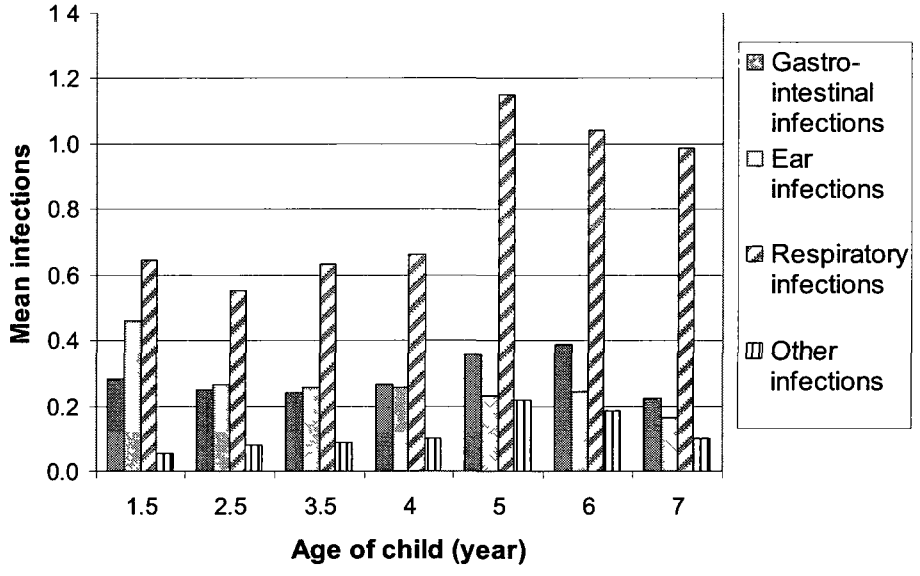
Types of infections are depicted in Figure 6 by the age of the child. The most common type of infection is consistently respiratory infections, which peaks when the children are in their first year of school at the age of 5. Ear infections appear to be most

Table 9: Mean, standard deviation and variance of number of infections in the past three months.

Cycle	Age of child (years)	Sample size	Mean infections	Standard deviation	Variance
2	1.5	2045	1.44	1.61	2.60
3	2.5	1997	1.15	1.30	1.70
4	3.5	1950	1.21	1.32	1.74
5	4	1944	1.29	1.30	1.70
6	5	1759	1.96	1.52	2.30
7	6	1492	1.86	1.56	2.44
8	7	1528	1.48	1.23	1.51

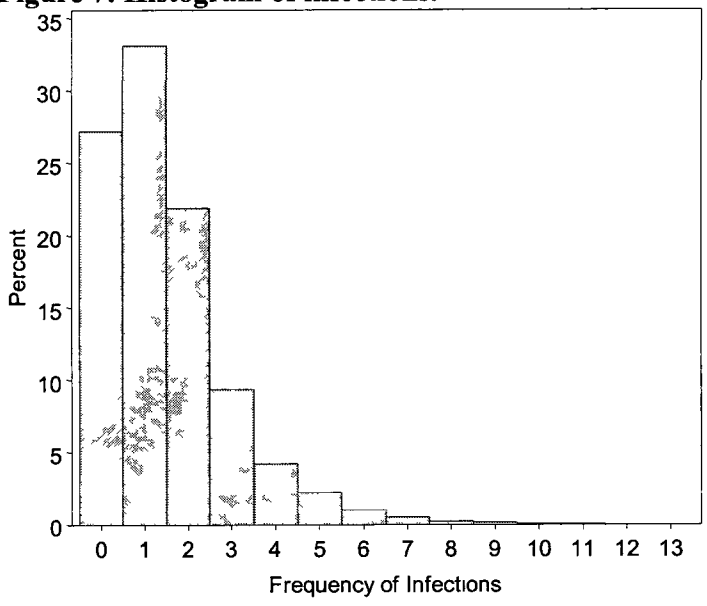
common in younger children and are less frequent as the child ages. Gastro-intestinal infections peak when the children are 6 years old.

Figure 6: Mean number of infections in the past three months by type and age.



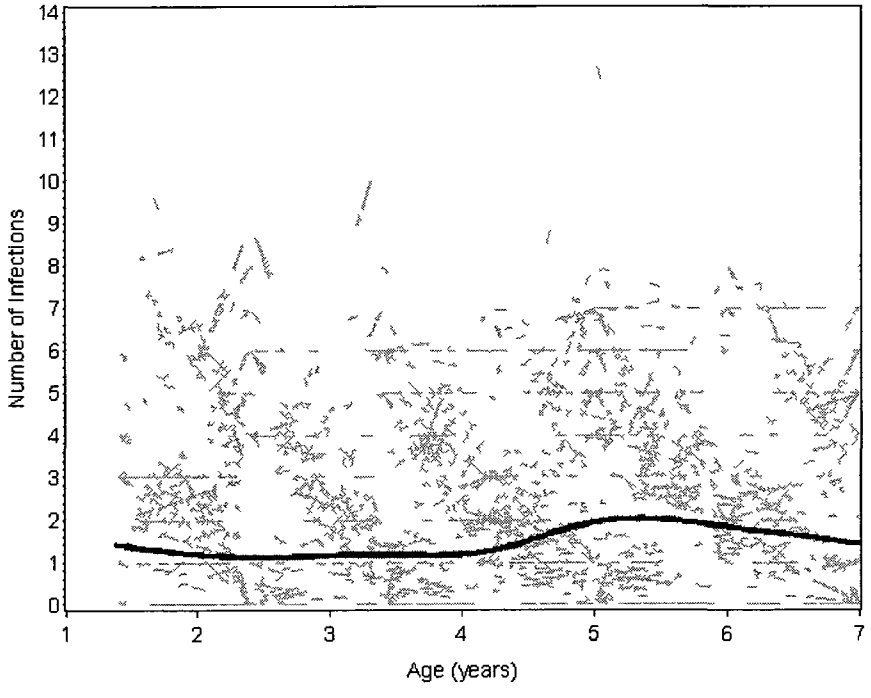
A histogram of all infection observations is displayed in Figure 7 shows that the highest frequency of infections is one.

Figure 7: Histogram of infections.



The time plot in Figure 8 illustrates that the trend of mean infections over time is not linear and peaks when the children commence school at age 5

Figure 8: Time plot of infection for all children.



5.1.4 Antibiotic Treatments

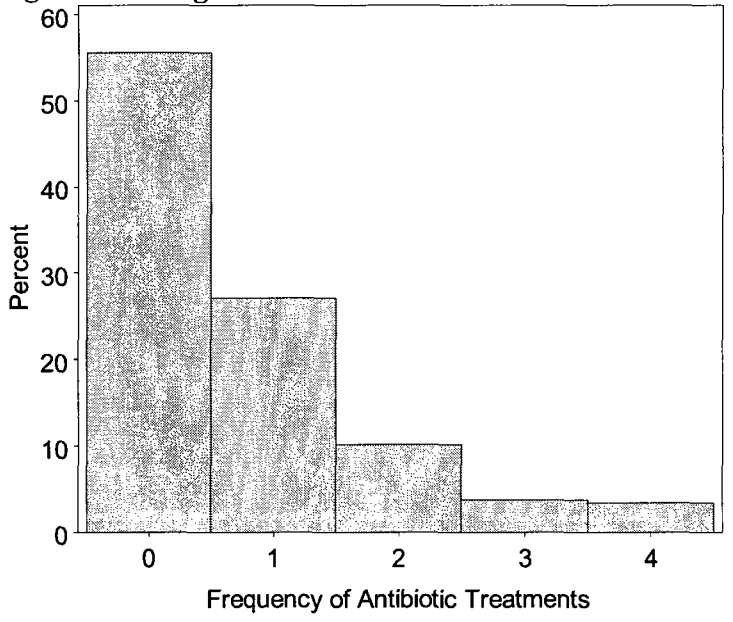
The number of antibiotic treatments in the past six months decreases over time with the exception of a rise occurring when the children are 4 years old (Table 10).

Table 10: Mean, standard deviation and variance of number of antibiotics treatments in the past six months.

Cycle	Age of child (years)	Sample size	Mean antibiotics treatments	Standard deviation	Variance
2	1.5	2045	1.10	1.25	1.56
3	2.5	1994	0.80	1.06	1.12
4	3.5	1950	0.69	0.97	0.95
5	4	1941	0.71	0.96	0.93
6	5	1757	0.65	0.92	0.84
7	6	1492	0.57	0.91	0.82
8	7	1527	0.40	0.75	0.56

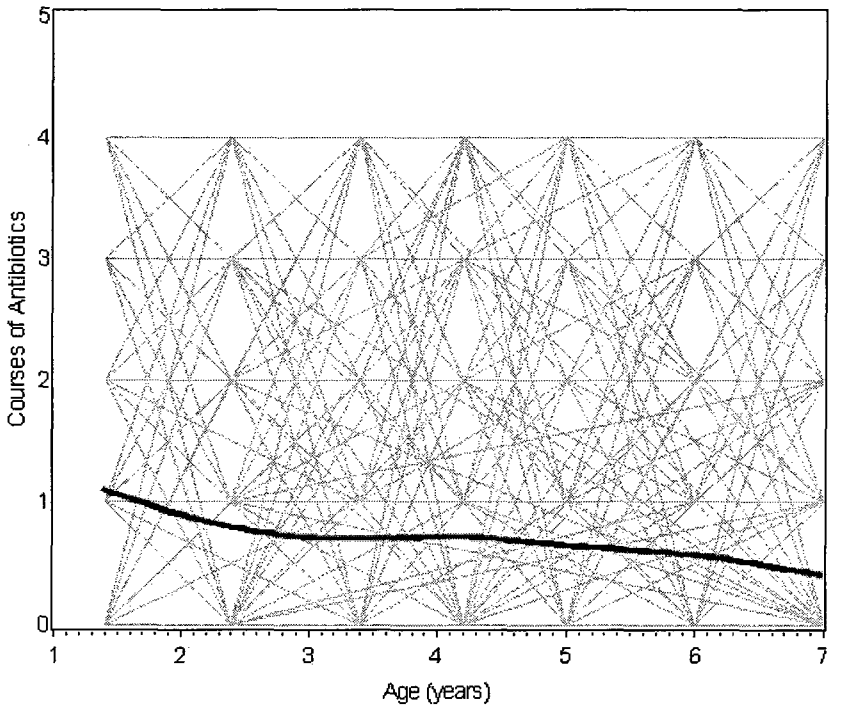
The histogram of the number of courses of antibiotics for all years is displayed in Figure 9, which shows that the child taking no antibiotics is most common.

Figure 9: Histogram of antibiotic treatments.



The time plot for antibiotic treatments is shown in Figure 10. Although this time plot is rather unusual due to the variable having a count with a range of 0 to 4, it depicts the mean line which decreases until approximately 3 years of age, rise slightly at 4 years of age, then continues to decrease.

Figure 10: Time plot of antibiotic treatments for all children.



5.2 Missing Data Analysis

5.2.1 Multivariable Missing Outcome Analysis using Logistic Regression

The following variables were found to be significantly associated with all missing health outcomes: low milk and alternative consumption (less than two times per day compared to between two and less than four times per day), sex, and high SES (see Table 11). Additionally, high meat and alternatives consumption (two or more times per day

compared to between one or more and two times per day), mother's immigrant status, and mother's BMI was significantly associated with missing physician consultations in cycle 8. Finally, maternal obesity (compared to the mother being normal weight) in cycle 2 was associated with missing infection data in cycle 8.

Table 11: Multivariable logistic regression of missing and non-missing cycle 8 indicators as the dependent variable.

Variables[†]	Missing physician consultations in cycle 8 OR (95% CI)	Missing infection in cycle 8 OR (95% CI)	Missing antibiotic treatment in cycle 8 OR (95% CI)
Grain products (low vs. med)	1.03 (0.79, 1.33)	0.94 (0.70, 1.25)	0.93 (0.70, 1.25)
Grain products (high vs. med)	1.01 (0.78, 1.32)	0.92 (0.69, 1.23)	0.94 (0.70, 1.25)
Vegetable (low vs. med)	1.12 (0.76, 1.65)	0.86 (0.56, 1.34)	0.89 (0.57, 1.37)
Vegetable (high vs. med)	1.07 (0.82, 1.40)	1.05 (0.80, 1.41)	1.06 (0.79, 1.42)
Fruit (low vs. med)	1.05 (0.75, 1.47)	1.06 (0.73, 1.53)	1.08 (0.74, 1.56)
Fruit (high vs. med)	0.83 (0.64, 1.08)	0.85 (0.64, 1.14)	0.86 (0.64, 1.14)
Milk and alternatives (low vs. med)	**2.67 (1.17, 6.08)***	**3.32 (1.47, 7.51)***	**3.29 (1.46, 7.45)***
Milk and alternatives (high vs. med)	0.92 (0.74, 1.15)	0.91 (0.71, 1.16)	0.90 (0.70, 1.15)
Meat and alternatives (low vs. med)	0.85 (0.68, 1.07)	0.80 (0.62, 1.04)	0.80 (0.62, 1.03)
Meat and alternatives (high vs. med)	**1.60 (1.07, 2.41)***	1.52 (0.99, 2.33)	1.51 (0.98, 2.31)
Sex (male vs female)	**1.39 (1.12, 1.73)***	**1.54 (1.21 1.95)***	**1.52 (1.20, 1.94)**
SES (low vs. med)	1.10 (0.82, 1.46)	0.81 (0.59, 1.11)	0.82 (0.60, 1.13)
SES (high vs. med)	**0.70 (0.53, 0.91)***	**0.65 (0.49, 0.87)***	**0.65 (0.49, 0.88)***
Mother immigrant vs non immigrant	**1.45 (1.01, 2.10)***	1.30 (0.87, 1.93)	1.33 (0.90, 1.98)
Mother's BMI (overweight vs normal weight)	**0.66 (0.45, 0.96)***	0.85 (0.57, 1.28)	0.85 (0.57, 1.27)
Mother's BMI (obese vs normal weight)	1.22 (0.93, 1.61)	**1.35 (1.00, 1.81)***	1.34 (0.99, 1.80)

[†] Only main exposures and covariates with statistical significance are displayed.
* $p < 0.05$

The results may allude to biased results in the study. However, note that these logistic regression models only included baseline covariates and considered no other year. All variables that were shown to be associated with missingness were included in the final models to account for potential bias due to missingness regardless of whether or not they were significant predictors.

5.2.2 Longitudinal Dropout Analysis

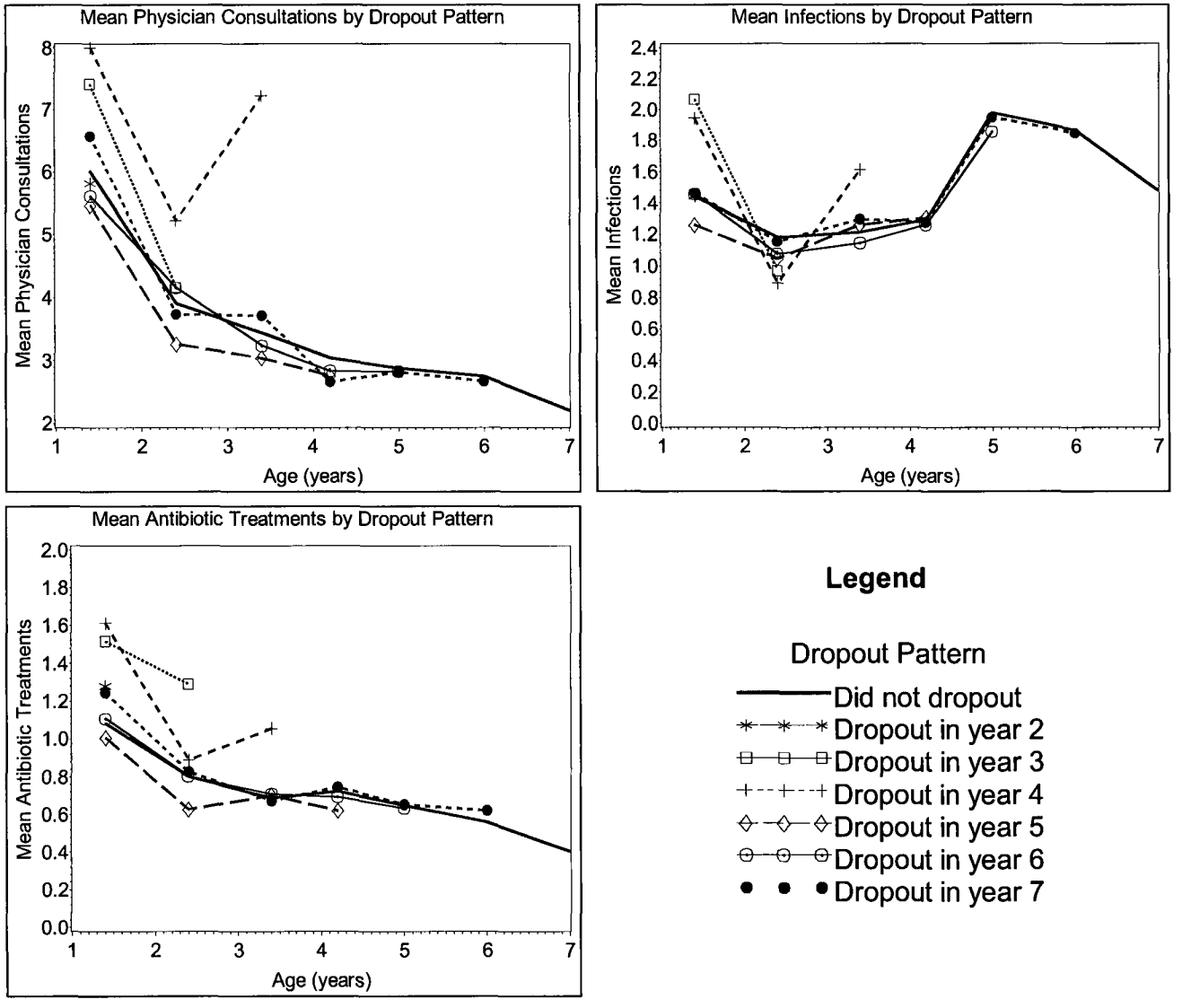
Table 12 depicts the frequencies and respective proportions of the children who dropped out in each year of the study. The most common dropout pattern was for those who were missing the last two years of data (13.51%), followed by those missing the last three years (7.49%). The graphical representation of dropout patterns by each health outcome is demonstrated in Figure 11. Both the physician consultation and infection outcomes possess similar dropout patterns. For both the physician consultation and infection graphs, dropping out in year 4 may be associated with higher mean physician consultations and mean infections. Also for both outcomes, those who dropped out in year 3 appeared to have higher mean physician consultations and infections in year 1 (and continued to year 2 for the infection graph). For the antibiotic treatment outcome, it appears as though dropping out in year 3 or 4 was associated with higher mean antibiotic treatments in year 1. Other than the cases mentioned above, the mean outcomes of children who dropped out appeared to be similar to those who stayed in the study.

Table 12: Frequencies and proportions of children who dropped out in various years of the study.

Dropout Pattern	Frequency	Proportion (%)
Did not drop out	1435	69.8
Dropped out in year 2	43	2.1
Dropped out in year 3	31	1.5
Dropped out in year 4	18	0.9
Dropped out in year 5	154	7.5
Dropped out in year 6	278	13.5
Dropped out in year 7	98	4.8

Although when looking at the dropout graphs it appears as though those who dropped out were significantly different than those who remained in the study, there were no significant differences found when performing the longitudinal dropout pattern analysis with outcomes as the dependent variable.

Figure 11: Mean outcomes by dropout pattern.



5.3 Relationship of Health Outcomes with Diet

In longitudinal studies, time plots are useful to depict and assess longitudinal relationships. Figure 12 shows that a quadratic trend may be adequate to model the mean number of physician consultations over time. Judging from the appearance of the graphs, there does not appear to be any overall significant differences in the crude relationships between physician consultation and any food group, however, differences by year may emerge for the high milk and alternatives and fruit food groups (Figure 12).

Figure 13 depicts the crude relationships between mean infections and each food group over time. The graphs indicate that a cubic or arbitrary trend may be adequate to model the mean number of infections over time. The following crude trends may be present: low milk and alternatives consumption being associated with a higher risk of infection, high milk and alternatives with lower risks of infection, and high fruit with lower overall risk of infection (Figure 13). For simplicity in the interpretation of the results, it was decided to treat time as categorical for the outcome.

Figure 14 depicts the relationships between the number of mean antibiotic treatments and each food group over time. Judging from the graphs, the trend may not be adequately modeled by a simple parametric curve, and therefore an arbitrary trend was fitted by treating time as categorical. The most obvious trend appears to be high vegetable consumption being associated with a lower risk of antibiotic treatment (Figure 14).

Figure 12: Time plots of all food groups by mean physician consultations.

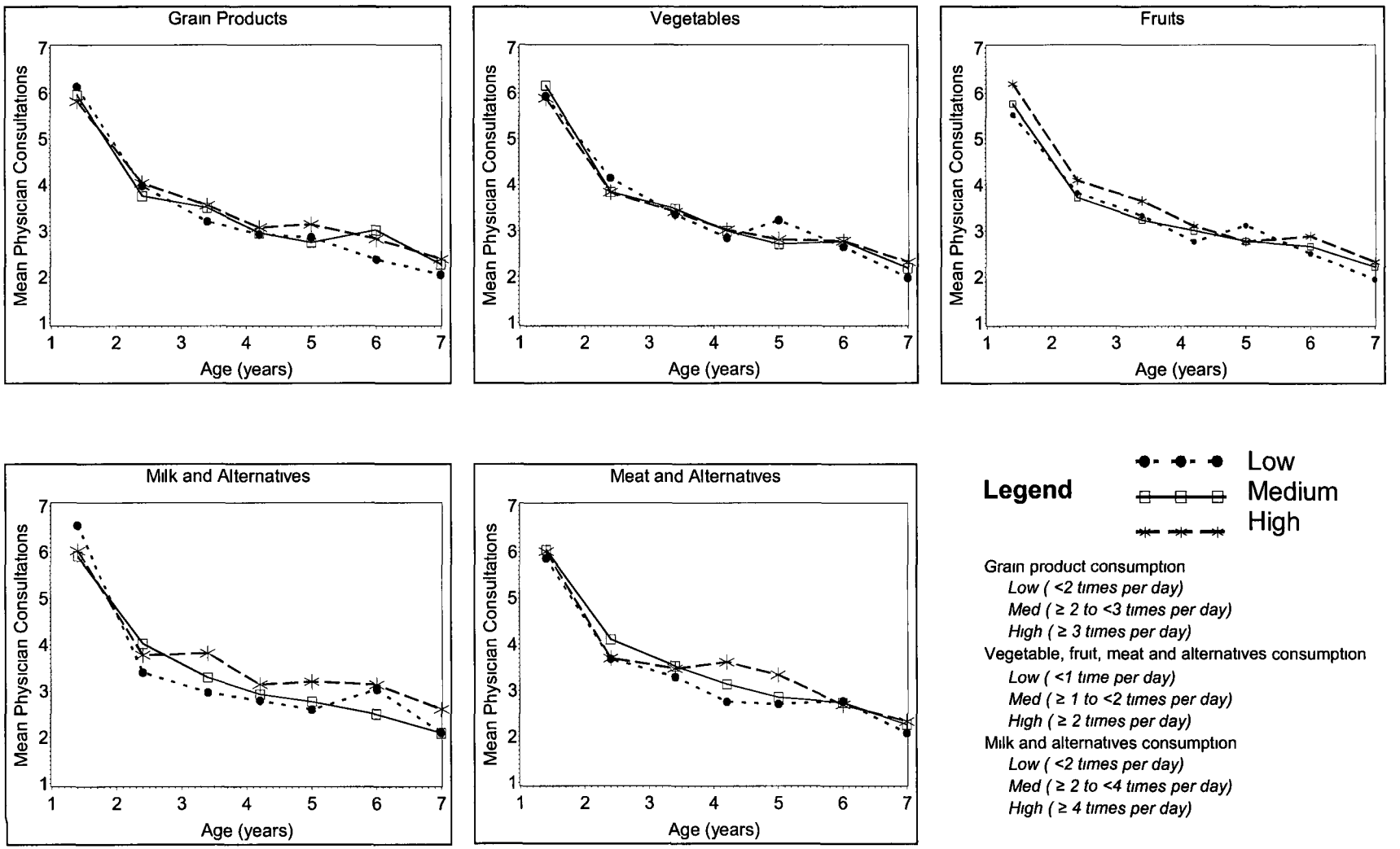


Figure 13: Time plots of all food groups by mean infections.

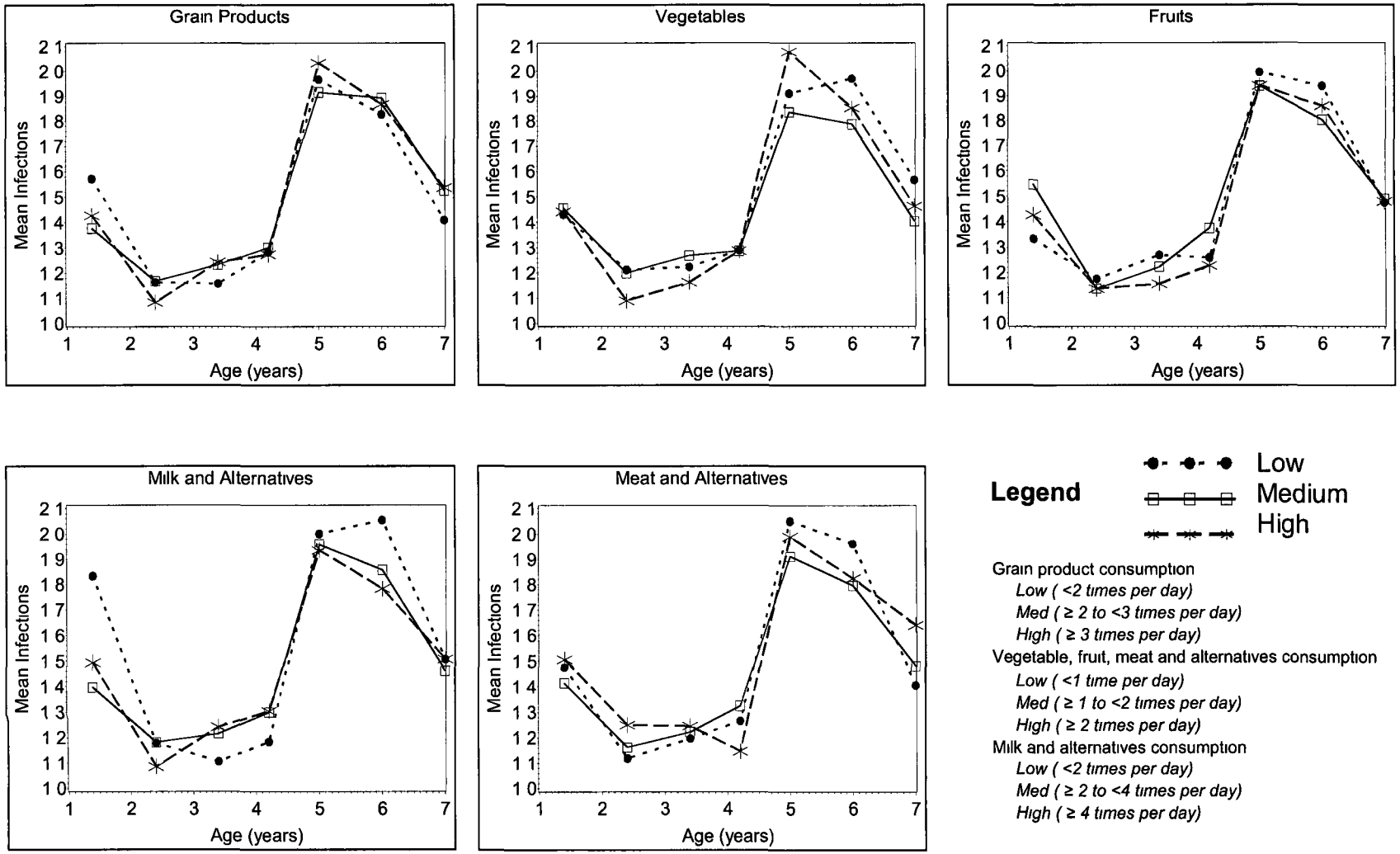
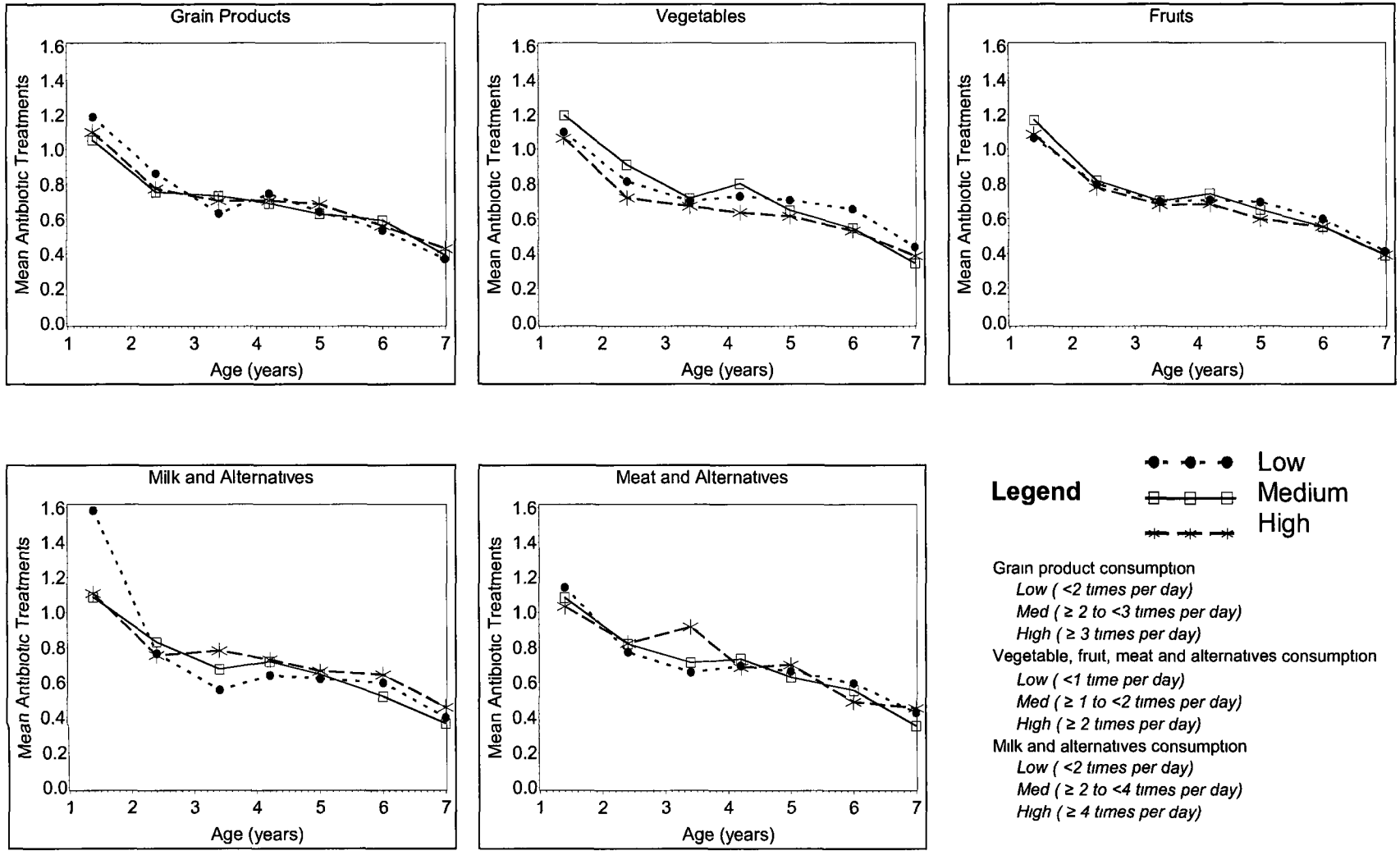


Figure 14: Time plots of all food groups by mean antibiotic treatments.



5.4 Bivariable and Multivariable Analysis

Bivariable and multivariable analyses were performed and are presented together by health outcome.

Both mother's smoking during pregnancy and mother's current smoking status were found to be protective in the models, which is counter intuitive, and also contradicts the descriptive analyses graphs (not shown) where the variables were not protective. When each variable was assessed independently in a longitudinal model, the variable was not protective, then upon the addition other important predictors the variable exhibited harmful effects. Thus, the inclusion of these two variables led to spurious effect estimates and were therefore removed from the analysis.

5.4.1 Physician Consultation Model

Of the 13 potential variables, the following were retained in the reduced models due to their association with missingness, if deemed important as outlined in the literature (see section 2.1), or if behaved as a potential confounders: sex, child's BMI, SES, mother's immigrant status and mother's BMI. Recall that a quadratic trend was determined to be most appropriate to model the mean number of physician consultations over time (see section 5.3).

The model building process for physician consultation showed consistency (Table 13). The first two columns (model 1 and model 2) in Table 13 depict similar risk ratios for the unadjusted bivariable model and the full adjusted bivariable model, showing that the associations in model 1 are unlikely to be as a result of confounding by these covariates. The first two columns of multivariable modeling (model 4 and model 5) in Table 13 also show similar risk ratios for the unadjusted multivariable and the full adjusted multivariable models.

Table 13: Risk ratios and 95% confidence intervals of the food groups and their year interaction terms in the models used to build the final physician consultation model.

		Bivariable models**			Multivariable models‡		
		Model 1 Full unadjusted	**Model 2**† Full adjusted	**Model 3**◊ Reduced adjusted	**Model 4** Full unadjusted	**Model 5**† Full adjusted	**Model 6**◊ Reduced adjusted
Food Group	Category†	Risk ratio (95% CI)	Risk ratio (95% CI)	Risk ratio (95% CI)	Risk ratio (95% CI)	Risk ratio (95% CI)	Risk ratio (95% CI)
Grain products	Low	1.01 (0.93, 1.09)	1.03 (0.94, 1.12)	1.00 (0.95, 1.04)	1.02 (0.94, 1.10)	1.04 (0.95, 1.14)	1.00 (0.96, 1.05)
	High	1.00 (0.93, 1.07)	1.02 (0.94, 1.11)	1.04 (0.99, 1.09)	0.99 (0.92, 1.07)	1.02 (0.94, 1.10)	1.03 (0.98, 1.09)
Grain*year	Low grain*year	1.02 (0.95, 1.09)	1.01 (0.93, 1.09)	-	1.01 (0.94, 1.08)	1.00 (0.92, 1.08)	-
	High grain*year	1.06 (0.98, 1.14)	1.04 (0.96, 1.13)	-	1.06 (0.99, 1.15)	1.05 (0.97, 1.13)	-
Grain*year^2	Low grain*year^2	0.99 (0.98, 1.00)	0.99 (0.98, 1.01)	-	0.99 (0.98, 1.01)	1.00 (0.98, 1.01)	-
	High grain*year^2	0.99 (0.98, 1.00)	0.99 (0.98, 1.00)	-	0.99 (0.98, 1.00)	0.99 (0.98, 1.00)	-
Vegetable	Low	0.95 (0.85, 1.06)	0.93 (0.82, 1.05)	0.99 (0.93, 1.05)	0.97 (0.86, 1.09)	0.95 (0.84, 1.09)	1.01 (0.94, 1.08)
	High	1.00 (0.92, 1.07)	0.95 (0.87, 1.03)	0.97 (0.93, 1.01)	0.96 (0.89, 1.04)	0.93 (0.85, 1.01)	**0.95 (0.91,1.00)***
Vegetable*year	Low vegetable*year	1.07 (0.97, 1.18)	1.09 (0.98, 1.21)	-	1.06 (0.96, 1.18)	1.09 (0.97, 1.21)	-
	High vegetable*year	1.00 (0.94, 1.07)	1.02 (0.95, 1.10)	-	1.01 (0.94, 1.08)	1.02 (0.95, 1.10)	-
Vegetable*year^2	Low vegetable*year^2	0.99 (0.97, 1.00)	0.98 (0.97, 1.00)	-	0.99 (0.97, 1.00)	0.98 (0.97, 1.00)	-
	High vegetable*year^2	1.00 (0.99, 1.01)	1.00 (0.99, 1.01)	-	1.00 (0.99, 1.01)	1.00 (0.99, 1.01)	-
Fruit	Low	0.93 (0.84, 1.03)	0.91 (0.82, 1.02)	0.97 (0.92, 1.03)	0.93 (0.84, 1.03)	0.92 (0.82, 1.03)	0.97 (0.92, 1.03)
	High	**1.08 (1.01,1.17)***	1.03 (0.95, 1.12)	1.04 (0.99, 1.09)	**1.09 (1.01,1.18)***	1.05 (0.97, 1.14)	1.04 (0.99, 1.09)
Fruit*year	Low fruit*year	1.05 (0.96, 1.15)	1.07 (0.97, 1.18)	-	1.04 (0.95, 1.14)	1.05 (0.95, 1.16)	-
	High fruit*year	0.99 (0.93, 1.06)	1.01 (0.94, 1.09)	-	0.99 (0.92, 1.06)	1.00 (0.93, 1.08)	-
Fruit*year^2	Low fruit*year^2	0.99 (0.98, 1.00)	0.99 (0.97, 1.00)	-	0.99 (0.98, 1.01)	0.99 (0.98, 1.01)	-
	High fruit*year^2	1.00 (0.99, 1.01)	1.00 (0.98, 1.01)	-	1.00 (0.99, 1.01)	1.00 (0.98, 1.01)	-
Milk & alternatives	Low	0.89 (0.71, 1.12)	0.85 (0.69, 1.06)	0.89 (0.78, 1.01)	0.92 (0.73, 1.15)	0.88 (0.71, 1.09)	0.89 (0.78, 1.01)
	High	1.01 (0.94, 1.07)	1.01 (0.94, 1.08)	0.99 (0.93, 1.05)	0.99 (0.93, 1.06)	1.00 (0.93, 1.07)	0.98 (0.92, 1.04)
Milk*year	Low milk*year	1.05 (0.90, 1.22)	1.06 (0.91, 1.24)	1.02 (0.99, 1.06)	1.02 (0.88, 1.20)	1.04 (0.89, 1.21)	1.03 (0.99, 1.06)
	High milk*year	1.00 (0.94, 1.07)	0.99 (0.92, 1.07)	**1.03 (1.01,1.05)***	1.01 (0.94, 1.08)	1.00 (0.92, 1.07)	**1.03 (1.01,1.05)***
Milk*year^2	Low milk*year^2	0.99 (0.97, 1.02)	0.99 (0.97, 1.02)	-	1.00 (0.98, 1.02)	1.00 (0.98, 1.02)	-
	High milk*year^2	1.01 (0.99, 1.02)	1.01 (0.99, 1.02)	-	1.00 (0.99, 1.02)	1.01 (0.99, 1.02)	-
Meat & alternatives	Low	0.97 (0.91, 1.04)	0.96 (0.89, 1.04)	**0.96 (0.92,1.00)***	0.98 (0.92, 1.05)	0.97 (0.90, 1.04)	**0.95 (0.91,0.99)***
	High	0.90 (0.78, 1.02)	0.91 (0.78, 1.05)	1.01 (0.93, 1.09)	0.90 (0.79, 1.03)	0.91 (0.79, 1.04)	0.99 (0.91, 1.07)
Meat*year	Low meat*year	0.98 (0.92, 1.04)	0.98 (0.91, 1.04)	-	0.97 (0.91, 1.03)	0.97 (0.91, 1.04)	-
	High meat*year	1.11 (0.97, 1.27)	1.10 (0.95, 1.28)	-	1.10 (0.97, 1.26)	1.10 (0.95, 1.27)	-
Meat*year^2	Low meat*year^2	1.00 (0.99, 1.01)	1.00 (0.99, 1.02)	-	1.01 (1.00, 1.02)	1.01 (0.99, 1.02)	-
	High meat*year^2	0.99 (0.97, 1.01	0.99 (0.96, 1.01)	-	0.99 (0.97, 1.01)	0.99 (0.96, 1.01)	-

† Each food group reference category is 'medium' consumption.

‡ Food group variables considered simultaneously in the same model.

◊ Results following backwards elimination.

** Each food group variable considered individually in the model.

† Adjusted for all covariates.

These consistencies show that the addition of the covariates does not drastically influence the food group estimates obtained. Risk ratios from model 3 (the reduced adjusted bivariable model) are similar to those in model 6 (the reduced adjusted multivariable model or final model) showing that the associations of each individual food group variable with the outcome are similar when adjusting for the other food group variables.

The final reduced adjusted multivariable model with all its covariates (model 6) is presented in Table 14. There was a quadratic rate of change in physician consultations by year, but this rate of change did not differ significantly among high, low or medium levels of

Table 14: Risk ratios, 95% confidence intervals, and p-values for the final model for physician consultation (model 6).

Characteristic	Category	Adjusted risk ratio	95% confidence intervals	P-value ($p \leq 0.05$)
Year		**0.72**	**0.69, 0.74**	**<.0001***
Year^2		**1.03**	**1.02, 1.03**	**<.0001***
Grain products	Low (< 2 times/day)†	1.00	0.96, 1.05	0.8724
	High (≥ 3 times/day)†	1.03	0.98, 1.09	0.1893
Vegetable	Low (< 1 time/day)‡	1.01	0.94, 1.08	0.8172
	High (≥ 2 times/day)‡	**0.95**	**0.91, 1.00**	**0.0320***
Fruit	Low (< 1 time/day)‡	0.97	0.92, 1.03	0.3574
	High (≥ 2 times/day)‡	1.04	0.99, 1.09	0.0883
Milk & alternatives	Low (< 2 times/day)†	0.89	0.78, 1.01	0.0721
	High (≥ 4 times/day)†	0.98	0.92, 1.04	0.5641
Meat & alternatives	Low (< 1 time/day)‡	**0.95**	**0.91, 0.99**	**0.0264***
	High (≥ 2 times/day)‡	0.99	0.91, 1.07	0.7338
Sex	Male (ref=female)	1.05	0.99, 1.12	0.0853
Chronic condition	Yes (ref=no)	**1.49**	**1.37, 1.61**	**<.0001***
Child's BMI	Obese (ref=normal weight)	1.01	0.95, 1.08	0.7069
	Overweight (ref=normal weight)	0.99	0.94, 1.04	0.5993
Childcare	Child in daycare (ref=at home)	**1.29**	**1.21, 1.37**	**<.0001***
	Care outside of home (ref=at home)	**1.08**	**1.02, 1.14**	**0.0052***
Breastfed	Exclusively ≥ 3 months (ref=no)	**0.94**	**0.88, 1.00**	**0.0541***
SES (tertiles)	Lowest third (ref=mid)	1.00	0.93, 1.07	0.9925
	Highest third (ref=mid)	1.02	0.96, 1.09	0.5049
Mother's age (years)	Less than 25 (ref = ≥ 35)	**1.15**	**1.03, 1.28**	**0.0147***
	25 to less than 30 (ref = ≥ 35)	**1.11**	**1.02, 1.21**	**0.0203***
	30 to less than 35 (ref = ≥ 35)	1.08	0.99, 1.17	0.0952
Mother's BMI	Obese (ref=normal weight)	**0.91**	**0.83, 0.99**	**0.0334***
	Overweight (ref=normal weight)	1.04	0.95, 1.15	0.3721
Mother immigrant	Yes (ref=no)	1.03	0.95, 1.10	0.4861
Milk*year	Low milk*year (ref= med milk*year)	1.03	0.99, 1.06	0.1437
	High milk*year (ref= med milk*year)	**1.03**	**1.01, 1.05**	**0.0073***

† ref = medium (≥ 2 to < 3 times/day)
‡ ref = medium (≥ 1 to < 2 times/day)
† ref = medium (≥ 2 to < 4 times/day)

food consumption, except for milk. Consuming a high amount of vegetables was associated with a 5% reduction in physician consultation compared to children who consumed a medium amount. Consuming low amounts of meat and alternatives was associated with a 5% reduction in physician consultation compared to children who consumed a medium amount. The association between milk and alternatives consumption varied by year as indicated by the significant milk and year interaction term. Therefore, risk ratios for high milk and alternatives were calculated for each year (Table 15). Children between the ages of 4 to 7 who consumed high amounts of milk had an increase in physician consultations compared to those who had medium consumption.

Table 15: Risk ratios and 95% confidence intervals for high milk and alternative consumption vs. med milk and alternative consumption by age.

Age	Risk ratio	95% confidence interval
1.5	0.98	0.92, 1.04
2.5	1.01	0.96, 1.06
3.5	1.04	0.99, 1.09
4	**1.07**	**1.02, 1.13***
5	**1.11**	**1.04, 1.18***
6	**1.14**	**1.05, 1.24***
7	**1.17**	**1.06, 1.30***

Children with a chronic condition as well as children who were in daycare or cared for outside the home had risk of higher physician consultation. An inverse relationship was seen between the mother's age and the risk of physician consultation as the younger the mother's age, the higher the risk of physician consultation. Exclusive breastfeeding for at least three months and mother's obesity were associated with a reduced risk of physician consultation.

5.4.2 Infection Model

Of the potential variables, sex, child's BMI, SES, and mother's BMI were retained in the reduced models due to their association with outcome missingness, deemed as important in the literature (see section 2.1), or behaved as a potential confounder. Recall from section 5.3 that an arbitrary trend was determined to be most appropriate to model the mean number of infections over time (see section 5.3); hence p-values of the year interactions are presented instead of risk ratios as in the previous section.

Consistency in risk ratios was seen throughout the model building process for infection as can be seen in Table 16. The first two columns (model 1 and model 2) in Table 16 depict mostly similar risk ratios for the full unadjusted bivariable model and the full adjusted bivariable model. The first two columns of multivariable modeling (model 4 and model 5) in Table 16 also show rather similar risk ratios for the full unadjusted multivariable and the full adjusted multivariable models. The consistencies depicted above indicate that the addition of the covariates does not drastically influence the food group estimates obtained. Model 3 (the reduced adjusted bivariable model) is similar model 6 (the reduced adjusted multivariable model or the final model) which means that the bivariable and multivariable results are comparable.

The final reduced adjusted multivariable model with all its covariates (model 6) is presented in Table 17. Lower risks of infection were seen each year except for an increase in the risk when children were 5 and 6 years old. Consuming high amounts of fruits was associated with a 5% reduction in infection risk compared to children who ate medium amounts of fruit. Having a chronic condition, being in daycare or in any care outside the home, or having an obese mother was associated with an increased risk of infection.

Table 16: Risk ratios and 95% confidence intervals of the food groups and their year interaction terms in the models used to build the final infection model.

		Bivariable models**			**Multivariable models‡**		
		Model 1 Full unadjusted	**Model 2†** Full adjusted	**Model 3◊** Reduced adjusted	**Model 4** Full unadjusted	**Model 5†** Full adjusted	**Model 6◊** Reduced adjusted
Food Group	Category†	Risk ratio (95% CI)	Risk ratio (95% CI)	Risk ratio (95% CI)	Risk ratio (95% CI)	Risk ratio (95% CI)	Risk ratio (95% CI)
Grain products	Low (<2 times/day)†	**1.14 (1.02,1.27)***	**1.14 (1.01,1.29)***	0.99 (0.95, 1.03)	**1.15 (1.03,1.28)***	**1.15 (1.02,1.30)***	0.99 (0.95, 1.04)
	High (≥ 3 times/day)	1.03 (0.92, 1.16)	1.06 (0.93, 1.21)	1.03 (0.98, 1.08)	1.02 (0.91, 1.15)	1.06 (0.94, 1.21)	1.04 (0.99, 1.10)
Grain*year	P-value from Type 3 GEE	p = 0.2556	p = 0.2488	-	p = 0.2935	p = 0.2882	-
Vegetable	Low (<1 time/day)	0.99 (0.83, 1.17)	0.93 (0.75, 1.14)	1.02 (0.97, 1.08)	1.00 (0.84, 1.19)	0.95 (0.77, 1.17)	1.03 (0.97, 1.09)
	High (≥ 2 times/day)	0.99 (0.89, 1.10)	0.94 (0.84, 1.05)	0.99 (0.95, 1.03)	1.00 (0.89, 1.12)	0.97 (0.85, 1.10)	1.00 (0.95, 1.04)
Vegetable*year	P-value from Type 3 GEE	**p = 0.0450***	p = 0.3055	-	**p = 0.0284***	p = 0.1203	-
Fruit	Low (<1 time/day)	**0.86 (0.75,0.99)***	0.87 (0.74, 1.02)	0.97 (0.92, 1.02)	**0.86 (0.74,0.99)***	0.87 (0.74, 1.03)	0.96 (0.91, 1.01)
	High (≥ 2 times/day)	0.93 (0.84, 1.04)	0.89 (0.80, 1.00)	**0.95 (0.91,0.99)***	0.93 (0.84, 1.04)	0.89 (0.79, 1.01)	**0.95 (0.90,0.99)***
Fruit*year	P-value from Type 3 GEE	p = 0.3102	p = 0.3664	-	p = 0.2049	p = 0.1685	-
Milk &	Low (<2 times/day)	1.25 (0.81, 1.92)	1.32 (0.81, 2.13)	1.02 (0.96, 1.09)	1.28 (0.83, 1.97)	1.35 (0.83, 2.17)	1.02 (0.95, 1.10)
alternatives	High (≥ 4 times/day)	1.05 (0.96, 1.15)	1.06 (0.95, 1.18)	0.99 (0.95, 1.03)	1.05 (0.96, 1.16)	1.07 (0.96, 1.20)	0.99 (0.95, 1.04)
Milk*year	P-value from Type 3 GEE	p = 0.6809	p = 0.2774	-	p = 0.6872	p = 0.2382	-
Meat &	Low (<1 time/day)	1.04 (0.94, 1.14)	1.00 (0.90, 1.12)	1.00 (0.96, 1.04)	1.04 (0.95, 1.15)	1.01 (0.90, 1.12)	0.99 (0.95, 1.03)
alternatives	High (≥ 2 times/day)	1.04 (0.86, 1.26)	0.99 (0.79, 1.25)	1.01 (0.94, 1.08)	1.03 (0.85, 1.25)	0.98 (0.78, 1.23)	1.00 (0.93, 1.08)
Meat*year	P-value from Type 3 GEE	p = 0.5110	p = 0.3386	-	p = 0.5045	p = 0.4058	-

† Each food group reference category is 'medium' consumption.

‡ Food group variables considered simultaneously in the same model.

◊ Results following backwards elimination.

** Each food group variable considered individually in the model.

† Adjusted for all covariates.

Exclusive breastfeeding for at least three months, mother being an immigrant, and high SES were associated with a decreased risk of infection.

Table 17: Risk ratios, 95% confidence intervals, and p-values for the final model for infection (model 6).

Characteristic	Category	Adjusted risk ratio	95% confidence intervals	P-value ($p \leq 0.05$)
Year	Year 1 (ref=year 0)	**0.77**	**0.72, 0.82**	**<.0001***
	Year 2 (ref=year 0)	**0.78**	**0.73, 0.84**	**<.0001***
	Year 3 (ref=year 0)	**0.83**	**0.77, 0.89**	**<.0001***
	Year 4 (ref=year 0)	**1.27**	**1.18, 1.36**	**<.0001***
	Year 5 (ref=year 0)	**1.18**	**1.10, 1.28**	**<.0001***
	Year 6 (ref=year 0)	**0.91**	**0.84, 0.99**	**0.0237***
Grain products	Low (< 2 times/day)†	0.99	0.95, 1.04	0.7267
	High (≥ 3 times/day) †	1.04	0.99, 1.10	0.0960
Vegetable	Low (< 1 time/day)‡	1.03	0.97, 1.09	0.3229
	High (≥ 2 times/day)‡	1.00	0.95, 1.04	0.9364
Fruit	Low (< 1 time/day)‡	0.96	0.91, 1.01	0.1411
	High (≥ 2 times/day)‡	**0.95**	**0.90, 0.99**	**0.0166***
Milk & alternatives	Low (< 2 times/day)†	1.02	0.95, 1.10	0.5377
	High (≥ 4 times/day)†	0.99	0.95, 1.04	0.7647
Meat & alternatives	Low (< 1 time/day)‡	0.99	0.95, 1.03	0.6270
	High (≥ 2 times/day)‡	1.00	0.93, 1.08	0.9214
Sex	Male (ref=female)	0.99	0.94, 1.04	0.7240
Chronic condition	Yes (ref=no)	**1.38**	**1.30, 1.47**	**<.0001***
BMI	Obese (ref=normal weight)	0.94	0.87, 1.02	0.1378
	Overweight (ref=normal weight)	0.96	0.91, 1.02	0.2069
Childcare	Child in daycare (ref=at home)	**1.28**	**1.21, 1.35**	**<.0001***
	Care outside of home (ref=at home)	**1.11**	**1.06, 1.17**	**<.0001***
Breastfed	Exclusively ≥ 3 months (ref=no)	**0.95**	**0.90, 1.00**	**0.0419***
SES (tertiles)	Lowest third (ref=mid)	1.04	0.98, 1.10	0.1876
	Highest third (ref=mid)	**0.95**	**0.90, 1.00**	**0.0415***
Immigrant mother	Yes (ref=no)	**0.80**	**0.73, 0.89**	**<.0001***
Mother's BMI	Obese (ref=normal weight)	**1.12**	**1.04, 1.21**	**0.0027***
	Overweight (ref=normal weight)	1.03	0.96, 1.10	0.4033
Living location	Urban (ref=rural)	1.05	0.99, 1.10	0.0843

† ref = medium (≥ 2 to < 3 times/day)
‡ ref = medium (≥ 1 to < 2 times/day)
† ref = medium (≥ 2 to < 4 times/day)

5.4.3 Antibiotic Treatment Model

Of the 13 potential variables, sex, child's BMI and SES were retained in the reduced models due to their association with outcome missingness, deemed as important in the literature (see section 2.1), or behaved as a confounder. Recall from section 5.3 that an

arbitrary trend was determined to be most appropriate to model the mean number of antibiotic treatments over time (see section 5.3), hence p-values of the year interactions are presented instead of risk ratios as in the previous section.

For the antibiotic model, the bivariable results were very similar to the multivariable results (Table 18). Consistency in the risk ratio was seen throughout the model building process as can be seen in Table 18. The first two columns (model 1 and model 2) in Table 18 depict mostly similar risk ratios for the full unadjusted bivariable model and the full adjusted bivariable model with the exception of a reduced risk ratio for low milk consumption. The consistencies depicted indicate that the addition of the covariates does not drastically influence the food group estimates obtained with the exception of low milk consumption. Model 3 (the reduced adjusted bivariable model) is similar to model 6 (the reduced adjusted multivariable model or the final model) which means the bivariable and multivariable results were comparable as effects were neither lost nor amplified.

The final reduced adjusted multivariable model with all its covariates (model 6) can be seen in Table 19. The rate of antibiotic treatments by year decreases in time. The only food group significantly associated with antibiotic treatments was vegetables, where consuming high amounts of vegetables was associated with a 9% lower risk of antibiotic treatment than medium vegetable consumption. Children with a chronic condition, in daycare or cared for outside the home, and children in the lowest SES category, all had a higher risk of antibiotic treatment. Children who were exclusively breastfed for at least three months had a lower risk of antibiotic treatment.

Table 18: Risk ratios and 95% confidence intervals of the food groups and their year interaction terms in the models used to build the final antibiotic treatment model.

		Bivariable models**			**Multivariable models‡**		
		Model 1 Full unadjusted	**Model 2†** Full adjusted	**Model 3◊** Reduced adjusted	**Model 4** Full unadjusted	**Model 5†** Full adjusted	**Model 6◊** Reduced adjusted
Food Group	Category†	Risk ratio (95% CI)	Risk ratio (95% CI)	Risk ratio (95% CI)	Risk ratio (95% CI)	Risk ratio (95% CI)	Risk ratio (95% CI)
Grain products	Low (<2 times/day)	**1.13 (1.01, 1.26)***	**1.14 (1.01, 1.29)***	1.01 (0.96, 1.07)	**1.13 (1.01, 1.27)**	**1.15 (1.02, 1.30)***	1.01 (0.96, 1.07)
	High (≥ 3 times/day)	1.06 (0.95, 1.19)	1.08 (0.95, 1.22)	1.02 (0.96, 1.09)	1.06 (0.95, 1.19)	1.09 (0.96, 1.24)	1.03 (0.96, 1.09)
Grain*year	P-value from Type 3 GEE	p = 0.1762	p = 0.3354	-	p = 0.2685	p = 0.3224	-
Vegetable	Low (<1 time/day)	0.92 (0.78, 1.09)	0.87 (0.72, 1.06)	0.95 (0.88, 1.02)	0.93 (0.78, 1.10)	0.88 (0.72, 1.08)	0.96 (0.89, 1.04)
	High (≥ 2 times/day)	**0.90 (0.80, 1.00)***	**0.88 (0.78, 0.99)***	**0.91 (0.86, 0.96)***	0.90 (0.81, 1.01)	0.90 (0.80, 1.03)	**0.91 (0.86, 0.96)***
Vegetable*year	P-value from Type 3 GEE	p = 0.3118	p = 0.3581	-	p = 0.2568	p = 0.1821	-
Fruit	Low (<1 time/day)	0.92 (0.79, 1.06)	0.87 (0.74, 1.03)	0.97 (0.92, 1.04)	0.90 (0.78, 1.05)	0.87 (0.73, 1.03)	0.94 (0.91, 1.04)
	High (≥ 2 times/day)	0.95 (0.85, 1.06)	0.90 (0.80, 1.02)	0.96 (0.91, 1.02)	0.97 (0.87, 1.09)	0.92 (0.81, 1.04)	0.97 (0.91, 1.04)
Fruit*year	P-value from Type 3 GEE	p = 0.9239	p = 0.6642	-	p = 0.6217	p = 0.4207	-
Milk &	Low (<2 times/day)	**1.42 (1.06, 1.91)**	1.24 (0.86, 1.81)	0.94 (0.84, 1.03)	**1.42 (1.06, 1.91)***	1.25 (0.87, 1.81)	0.94 (0.84, 1.04)
alternatives	High (≥ 4 times/day)	1.01 (0.92, 1.11)	1.00 (0.90, 1.11)	1.01 (0.95, 1.07)	1.02 (0.93, 1.13)	1.02 (092, 1.13)	1.02 (0.96, 1.08)
Milk*year	P-value from Type 3 GEE	p = 0.1891	p = 0.1302	-	p = 0.2632	p = 0.2414	-
Meat &	Low (<1 time/day)	1.04 (0.94, 1.14)	1.03 (0.93, 1.15)	0.98 (0.93, 1.04)	1.04 (0.94, 1.15)	1.03 (0.93, 1.15)	0.98 (0.93, 1.03)
alternatives	High (≥ 2 times/day)	0.93 (0.77, 1.11)	0.92 (0.74, 1.13)	0.98 (0.88, 1.09)	0.92 (0.76, 1.11)	0.91 (0.73, 1.12)	0.98 (088, 1.09)
Meat*year	P-value from Type 3 GEE	p = 0.3941	p = 0.5051	-	p = 0.4101	p = 0.5668	-

† Each food group reference category is 'medium' consumption.

** Each food group variable considered individually in the model.

‡ Food group variables considered simultaneously in the same model.

† Adjusted for all covariates.

◊ Results following backwards elimination.

Table 19: Risk ratios, 95% confidence intervals, and p-values for the final antibiotic treatment model (model 6).

Characteristic	Category	Adjusted risk ratio	95% confidence intervals	P-value ($p \leq 0.05$)
Year	Year 1 (ref=year 0)	**0.68**	**0.64, 0.74**	**<.0001***
	Year 2 (ref=year 0)	**0.57**	**0.53, 0.62**	**<.0001***
	Year 3 (ref=year 0)	**0.57**	**0.53, 0.62**	**<.0001***
	Year 4 (ref=year 0)	**0.49**	**0.44, 0.54**	**<.0001***
	Year 5 (ref=year 0)	**0.43**	**0.39, 0.49**	**<.0001***
	Year 6 (ref=year 0)	**0.31**	**0.27, 0.35**	**<.0001***
Grain products	Low (< 2 times/day)†	1.01	0.96, 1.07	0.6427
	High (≥ 3 times/day)†	1.03	0.96, 1.09	0.4165
Vegetable	Low (< 1 time/day)‡	0.96	0.89, 1.04	0.3211
	High (≥ 2 times/day)‡	**0.91**	**0.86, 0.96**	**0.0017***
Fruit	Low (< 1 time/day)‡	0.97	0.91, 1.04	0.4388
	High (≥ 2 times/day)‡	0.97	0.91, 1.04	0.3715
Milk & alternatives	Low (< 2 times/day)†	0.94	0.84, 1.04	0.2011
	High (≥ 4 times/day)†	1.02	0.96, 1.08	0.5582
Meat & alternatives	Low (< 1 time/day)‡	0.98	0.93, 1.03	0.4706
	High (≥ 2 times/day)‡	0.98	0.88, 1.09	0.7390
Sex	Male (ref=female)	1.00	0.94, 1.07	0.9339
Chronic condition	Yes (ref=no)	**1.54**	**1.42, 1.68**	**<.0001***
BMI	Obese (ref=normal weight)	0.97	0.89, 1.07	0.5919
	Overweight (ref=normal weight)	0.95	0.89, 1.02	0.1892
Childcare	Child in daycare (ref=at home)	**1.61**	**1.50, 1.74**	**<.0001***
	Care outside of home (ref=at home)	**1.16**	**1.08, 1.24**	**<.0001***
Breastfed	Exclusively ≥ 3 months (ref=no)	**0.82**	**0.76, 0.89**	**<.0001***
SES (tertiles)	Lowest third (ref=mid)	**1.09**	**1.01, 1.18**	**0.0312***
	Highest third (ref=mid)	0.97	0.90, 1.04	0.4287

† ref = medium (≥ 2 to < 3 times/day)
‡ ref = medium (≥ 1 to < 2 times/day)
† ref = medium (≥ 2 to < 4 times/day)

5.5 Model Assessment

5.5.1 Multiple Imputation for Missing Data

The number of imputed values for each variable in the study is reported in Table 24, Appendix B. Child's BMI in cycle 6 and 7 had the highest imputed values of 66% and 81% respectively. Some variables such as sex and breastfed exclusively for at least three months were complete, therefore did not require imputation.

The resulting models for physician consultation, infection and antibiotic treatment can be found from Table 25 to Table 27 in Appendix C. The physician consultation

imputed model (Table 25) had some additional significant effects in the imputed model: high grain product consumption, fruit consumption, child's sex, high SES, mother's age being between 30 and 35, and mother being an immigrant. The largest absolute difference in risk ratios between the model and the imputed model occurred for the low milk and alternatives variable where the risk ratio changed from 0.89 in the non-imputed model to 0.94 in the imputed model. The infection imputed model's (Table 26) additional significant effects were child being obese, mother being obese, and living location. There were two effects that were no longer significant; time in cycle 8 and high fruit consumption. The largest absolute difference in risk ratios occurred for the daycare variable in the 'child being in daycare' category where the risk ratio changed from 1.28 to 1.22 in the imputed model. The antibiotic treatment imputed model (Table 27) was almost identical to the imputed antibiotic model whereby the same variables were significant in both models. The largest absolute difference in risk ratios was the daycare variable (in the cared for in daycare category) where the risk ratio changed from 1.61 to 1.49.

The results of the multiple imputation models were not substantially different from the non-imputed results as all the estimates were in the same direction. However, the imputed models generally had more significant terms than the non-imputed models. The increased significance is most likely due to the imputed models having more power as their sample sizes were considerably larger. Overall, it is expected that the non imputed models are a conservative estimate of the effect as the imputed models have smaller standard errors which in turn caused the imputed model to have more significant variables.

5.5.2 Model Assessment using Survey Weights

The unweighted and weighted analysis of the final model for physician consultation is depicted in Table 20. This model remained mostly unchanged after weighting. Sex in the weighted model appeared to be significant based on its confidence intervals but its p-value was 0.0594.

Similarly, when applying weights to the infection model, there were few changes (Table 21). The following variables were no longer significant in the model: year six, exclusive breastfeeding for at least three months, and being from high SES families.

Table 20: Unweighted versus weighted analysis of final physician consultation model.

		Unweighted		Weighted	
Characteristic	Category	Risk ratio	95% Confidence intervals	Risk ratio	95% Confidence intervals
Year		**0.72**	**0.69, 0.74**	**0.71**	**0.69, 0.74**
Year 2		**1.03**	**1.02, 1.03**	**1.03**	**1.02, 1.03**
Grain products	Low (< 2 times/day) †	1.00	0.96, 1.05	0.99	0.95, 1.05
	High (≥ 3 times/day) †	1.03	0.98, 1.08	1.04	0.98, 1.09
Vegetable	Low (< 1 time/day)‡	1.01	0.95, 1.08	1.01	0.95, 1.08
	High (≥ 2 times/day)‡	0.96	0.91, 1.00	0.95	0.91, 1.00
Fruit	Low (< 1 time/day)‡	0.97	0.91, 1.03	0.98	0.91, 1.04
	High (≥ 2 times/day)‡	1.04	0.99, 1.09	1.04	0.99, 1.09
Milk & alternatives	Low (< 2 times/day)†	0.89	0.78, 1.01	0.87	0.75, 1.00
	High (≥ 4 times/day)†	0.98	0.92, 1.04	1.00	0.94, 1.07
Meat & alternatives	Low (< 1 time/day)‡	**0.95**	**0.91, 0.99**	**0.95**	**0.91, 0.99**
	High (≥ 2 times/day)‡	0.99	0.91, 1.07	0.97	0.90, 1.05
Sex	Male (ref=female)	1.05	0.99, 1.12	1.06	1.00, 1.13
Chronic condition	Yes (ref=no)	**1.48**	**1.36, 1.60**	**1.46**	**1.35, 1.58**
BMI	Obese (ref=normal weight)	1.01	0.95, 1.08	1.00	0.94, 1.07
	Overweight (ref=normal weight)	0.98	0.93, 1.03	0.99	0.94, 1.05
Childcare	Child in daycare (ref=at home)	**1.29**	**1.21, 1.38**	**1.30**	**1.21, 1.39**
	Care outside home (ref=at home)	**1.07**	**1.02, 1.13**	**1.07**	**1.01, 1.13**
Breastfed	Exclusively ≥ 3 months (ref=no)	**0.93**	**0.88, 0.99**	**0.95**	**0.89, 1.01**
SES (tertiles)	Lowest third (ref=mid)	1.01	0.94, 1.07	1.03	0.96, 1.10
	Highest third (ref=mid)	1.03	0.96, 1.09	1.05	0.98, 1.11
Mother's age (years)	Less than 25 (ref = ≥ 35)	**1.16**	**1.03, 1.29**	**1.16**	**1.02, 1.32**
	25 to less than 30 (ref = ≥ 35)	**1.11**	**1.01, 1.21**	**1.09**	**0.99, 1.21**
	30 to less than 35 (ref = ≥ 35)	1.08	0.99, 1.17	1.05	0.95, 1.16
Immigrant mother	Yes (ref=no)	0.89	0.82, 0.98	0.90	0.81, 0.99
Milk*year	Low milk*year†	1.03	0.99, 1.06	1.02	0.99, 1.06
	High milk*year†	**1.03**	**1.01, 1.05**	**1.02**	**1.00, 1.04**

† ref = medium (≥ 2 to < 3 times/day)
‡ ref = medium (≥ 1 to < 2 times/day)
† ref = medium (≥ 2 to < 4 times/day)

Living area became significant after weighting. All of these changes occurred to variables that were borderline significant or insignificant. The lower confidence interval for high grain product consumption did increase and was rounded which appears significant, but the p-value still remained non-significant (p-value=0.0697).

Table 21: Unweighted versus weighted analysis of the final infection model.

		Unweighted		Weighted	
Characteristic	Category	Risk ratio	95% Confidence intervals	Risk ratio	95% Confidence intervals
Year	Year 1 (ref=year 0)	**0.77**	**0.72, 0.82**	**0.77**	**0.72, 0.82**
	Year 2 (ref=year 0)	**0.78**	**0.73, 0.84**	**0.79**	**0.73, 0.85**
	Year 3 (ref=year 0)	**0.83**	**0.77, 0.89**	**0.84**	**0.78, 0.90**
	Year 4 (ref=year 0)	**1.27**	**1.18, 1.36**	**1.30**	**1.21, 1.40**
	Year 5 (ref=year 0)	**1.18**	**1.10, 1.28**	**1.21**	**1.12, 1.31**
	Year 6 (ref=year 0)	**0.91**	**0.84, 0.99**	0.95	0.87, 1.02
Grain products	Low (< 2 times/day)†	0.99	0.95, 1.04	1.00	0.95, 1.04
	High (≥ 3 times/day)†	1.04	0.99, 1.10	1.05	1.00, 1.10
Vegetable	Low (< 1 time/day)‡	1.03	0.97, 1.09	1.03	0.97, 1.09
	High (≥ 2 times/day)‡	1.00	0.95, 1.04	0.98	0.94, 1.03
Fruit	Low (< 1 time/day)‡	0.96	0.91, 1.01	0.96	0.91, 1.01
	High (≥ 2 times/day)‡	**0.95**	**0.90, 0.99**	**0.95**	**0.90, 1.00**
Milk & alternatives	Low (< 2 times/day)†	1.02	0.95, 1.10	1.02	0.94, 1.10
	High (≥ 4 times/day)†	0.99	0.95, 1.04	0.99	0.95, 1.04
Meat & alternatives	Low (< 1 time/day)‡	0.99	0.95, 1.03	0.99	0.95, 1.03
	High (≥ 2 times/day)‡	1.00	0.93, 1.08	0.99	0.92, 1.06
Sex	Male (ref=female)	0.99	0.94, 1.04	1.00	0.95, 1.05
Chronic condition	Yes (ref=no)	**1.38**	**1.30, 1.47**	**1.38**	**1.29, 1.47**
BMI	Obese (ref=normal weight)	0.94	0.87, 1.02	0.93	0.86, 1.01
	Overweight (ref=normal weight)	0.96	0.91, 1.02	0.99	0.94, 1.05
Childcare	Child in daycare (ref=at home)	**1.28**	**1.21, 1.35**	**1.28**	**1.21, 1.36**
	Care outside home (ref=at home)	**1.11**	**1.06, 1.17**	**1.11**	**1.06, 1.17**
Breastfed	Exclusively ≥ 3 months (ref=no)	**0.95**	**0.90, 1.00**	0.95	0.90, 1.01
SES (tertiles)	Lowest third (ref=mid)	1.04	0.98, 1.10	1.04	0.98, 1.10
	Highest third (ref=mid)	**0.95**	**0.90, 1.00**	0.95	0.90, 1.01
Immigrant mother	Yes (ref=no)	**0.80**	**0.73, 0.89**	**0.80**	**0.73, 0.89**
Mother's BMI	Obese (ref=normal weight)	**1.12**	**1.04, 1.21**	**1.12**	**1.03, 1.21**
	Overweight (ref=normal weight)	1.03	0.96, 1.10	1.04	0.97, 1.11
Living location	Urban (ref=rural)	1.05	0.99, 1.10	**1.06**	**1.01, 1.12**

† ref = medium (≥ 2 to < 3 times/day)
‡ ref = medium (≥ 1 to < 2 times/day)
† ref = medium (≥ 2 to < 4 times/day)

The antibiotic treatment model also underwent few changes following weighting as can be seen in Table 22. Low SES became non-significant after adding the weights, and high SES became significant. High grain product consumption became significant after weighting with a p-value of 0.0449.

Table 22: Unweighted versus weighted analysis of antibiotic treatment final model.

		Unweighted		Weighted	
Characteristic	Category	Risk ratio	95% Confidence intervals	Risk ratio	95% Confidence intervals
Year	Year 1 (ref=year 0)	**0.68**	**0.64, 0.74**	**0.67**	**0.62, 0.71**
	Year 2 (ref=year 0)	**0.57**	**0.53, 0.62**	**0.56**	**0.51, 0.60**
	Year 3 (ref=year 0)	**0.57**	**0.53, 0.62**	**0.56**	**0.52, 0.60**
	Year 4 (ref=year 0)	**0.49**	**0.44, 0.54**	**0.49**	**0.45, 0.54**
	Year 5 (ref=year 0)	**0.43**	**0.39, 0.49**	**0.44**	**0.40, 0.49**
	Year 6 (ref=year 0)	**0.31**	**0.27, 0.35**	**0.32**	**0.29, 0.36**
Grain products	Low (< 2 times/day)†	1.01	0.96, 1.07	1.03	0.97, 1.09
	High (≥ 3 times/day)†	1.03	0.96, 1.09	**1.06**	**1.00, 1.13**
Vegetable	Low (< 1 time/day)‡	0.96	0.89, 1.04	0.98	0.91, 1.06
	High (≥ 2 times/day)‡	**0.91**	**0.86, 0.96**	**0.90**	**0.85, 0.96**
Fruit	Low (< 1 time/day)‡	0.97	0.91, 1.04	0.94	0.88, 1.01
	High (≥ 2 times/day)‡	0.97	0.91, 1.04	0.99	0.93, 1.05
Milk & alternatives	Low (< 2 times/day)†	0.94	0.84, 1.04	0.93	0.84, 1.03
	High (≥ 4 times/day)†	1.02	0.96, 1.08	1.02	0.97, 1.07
Meat & alternatives	Low (< 1 time/day)‡	0.98	0.93, 1.03	1.01	0.96, 1.07
	High (≥ 2 times/day)‡	0.98	0.88, 1.09	0.97	0.88, 1.07
Child's sex	Male (ref=female)	1.00	0.94, 1.07	1.04	0.98, 1.12
Chronic condition	Yes (ref=no)	**1.54**	**1.42, 1.68**	**1.49**	**1.37, 1.62**
BMI	Obese (ref=normal weight)	0.97	0.89, 1.07	0.94	0.86, 1.03
	Overweight (ref=normal weight)	0.95	0.89, 1.02	1.04	0.98, 1.11
Childcare	Child in daycare (ref=at home)	**1.61**	**1.50, 1.74**	**1.64**	**1.53, 1.76**
	Care outside home (ref=at home)	**1.16**	**1.08, 1.24**	**1.18**	**1.11, 1.26**
Breastfed	Exclusively ≥ 3 months (ref=no)	**0.82**	**0.76, 0.89**	**0.84**	**0.78, 0.91**
SES (tertile)	Lowest third (ref=mid)	**1.09**	**1.01, 1.18**	1.04	0.97, 1.12
	Highest third (ref=mid)	0.97	0.90, 1.04	**0.92**	**0.86, 0.99**

† ref = ≥ 2 to < 3 times/day
‡ ref = ≥ 1 to < 2 times/day
† ref = ≥ 2 to < 4 times/day

Overall, all the food groups that were significant in the unweighted models were also significant in the weighted models. Only one food group became significant after weighting: in the weighted antibiotic treatment model, high grain consumption became significant with a 6% increase in risk of antibiotic treatment compared to medium grain

consumption. There were some changes in significance when comparing covariates in the weighted and unweighted models but most of these changes occurred to covariates that were close to significance.

5.5.3 Residuals and Cook's D

Model assessment using graphical means was conducted for each health outcome. Histograms of the residuals, residuals graphed against the adjusted predicted means, and the residuals plotted against all covariates indicated a good model fit for all models. The residuals versus the adjusted predicted means for all the models are depicted in Figure 15 to Figure 17, and showed no residual trends as the mean line is near zero.

Figure 15: Pearson residual versus the adjusted predicted mean for the physician consultation model.

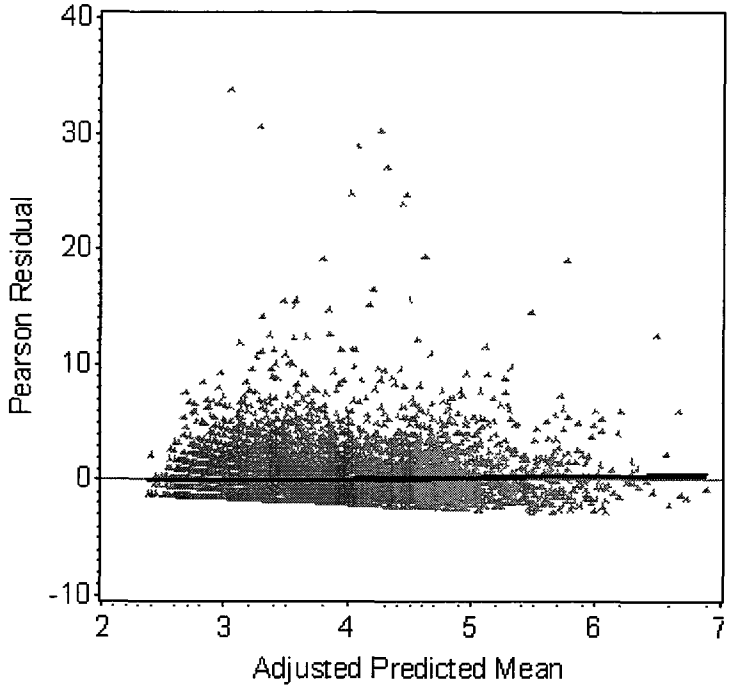


Figure 16: Pearson residual versus the adjusted predicted mean for the infection model.

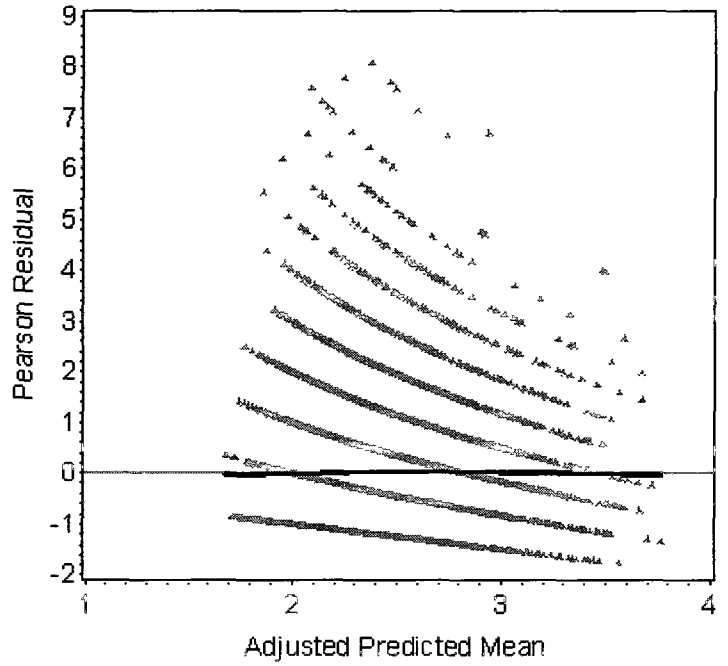
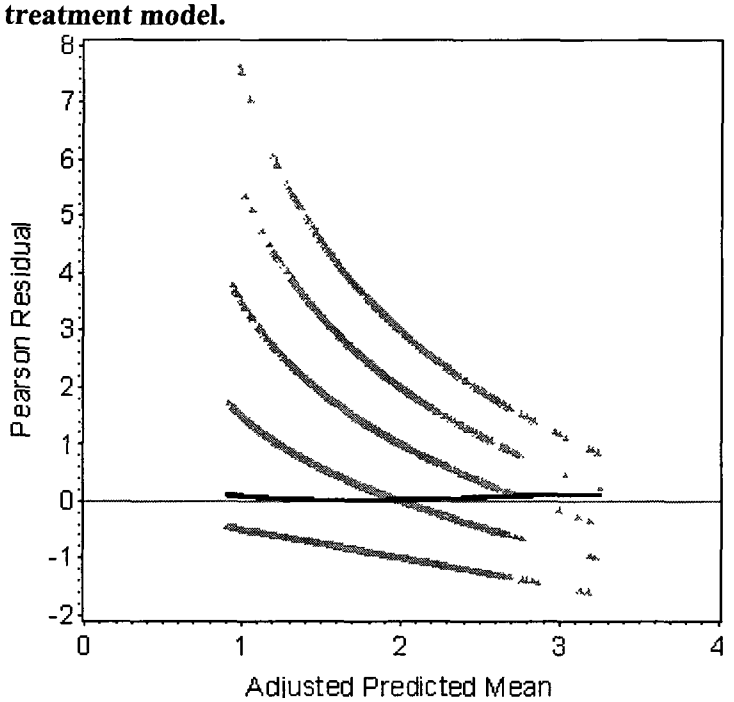


Figure 17: Pearson residual versus the adjusted predicted mean for the antibiotic treatment model.



All potential outliers identified by the histograms of the residuals and all the cluster outliers identified by Cook's cluster D were individually assessed. All extreme observations had high counts of the health outcomes and were usually identified as having a chronic disease. None of the observations identified warranted the child's removal from the study.

Chapter 6: Discussion

6.1 Summary of Findings

Table 23 summarizes the direction of the relationships of the food groups and the covariates with each health outcome in the final multivariable models. Note that three of the covariates showed similar directional relationships in all of the models: chronic condition and the child being in daycare or cared for outside of the home were both positively associated, and breastfeeding exclusively for at least three months was negatively associated with all outcomes.

6.2 Interpretation

Physician consultations decreased steadily over time. After controlling for child and parental characteristics, dietary intake of some food groups over time were found to be significantly associated with physician consultations. Children who consumed high amounts of vegetables or low amounts of meats and alternatives had a 5% reduction in physician consultations compared to those who had medium consumption. Furthermore, children between the ages of 4 to 7 who consumed high amounts of milk had an increase in physician consultations compared to those who had medium consumption. The child's consumption of grains, and fruits over time were not associated with physician consultations.

Although there was not a relationship between the frequency of consumption of all food groups found and physician consultations as hypothesized, the results are in accordance with the hypothesis for some food groups as high vegetable and low meat

consumption were associated with a decreased risk of consultations. The finding for 'low' meat (less than

Table 23: Summary table of relationship direction between all food groups and covariates and each health outcome.

Characteristic	Category	Physician consultation	Infection	Antibiotic treatment
Year (Continuous)		↓	Ø	Ø
Year^2 (Continuous)		↑	Ø	Ø
Year (Categorical)	Year 1 (ref=year 0)	Ø	↓	↓
	Year 2 (ref=year 0)	Ø	↓	↓
	Year 3 (ref=year 0)	Ø	↓	↓
	Year 4 (ref=year 0)	Ø	↑	↓
	Year 5 (ref=year 0)	Ø	↑	↓
	Year 6 (ref=year 0)	Ø	↓	↓
Grain products	Low (< 2 times/day)†	-	-	-
	High (≥ 3 times/day)†	-	-	-
Vegetable	Low (< 1 time/day)‡	-	-	-
	High (≥ 2 times/day)‡	↓	-	↓
Fruit	Low (< 1 time/day)‡	-	-	-
	High (≥ 2 times/day)‡	-	↓	-
Milk & alternatives	Low (< 2 times/day)†	-	-	-
	High (≥ 4 times/day)†	-	-	-
Meat & alternatives	Low (<1 time/day)‡	↓	-	-
	High (≥ 2 times/day)‡	-	-	-
Sex	Male (ref=female)	-	-	-
Chronic condition	Yes (ref=no)	↑	↑	↑
BMI	Obese (ref=normal weight)	-	-	-
	Overweight (ref=normal weight)	-	-	-
Childcare	Child in daycare (ref=at home)	↑	↑	↑
	Care outside of home (ref=at home)	↑	↑	↑
Breastfed	Exclusively ≥ 3 months (ref=no)	↓	↓	↓
SES (tertiles)	Lowest third (ref=mid)	-	-	↑
	Highest third (ref=mid)	-	↓	-
Mother's age (years)	Less than 25 (ref = ≥ 35)	↑	Ø	Ø
	25 to less than 30 (ref = ≥ 35)	↑	Ø	Ø
	30 to less than 35 (ref = ≥ 35)	-	Ø	Ø
Immigrant mother	Yes (ref=no)	-	↓	Ø
Mother's BMI	Obese (ref=normal weight)	↓	↑	Ø
	Overweight (ref=normal weight)	-	-	Ø
Living location	Urban (ref=rural)		-	Ø
Milk*year	Low milk (< 2 times/day)†*year	-	Ø	Ø
	High milk (≥ 4 times/day)†*year	↑	Ø	Ø

↑ significant positive relationship
↓ significant negative relationship
- no significant relationship
Ø indicates that the variable was not included in the model

† ref = medium (≥ 2 to < 3 times/day)
‡ ref = medium (≥ 1 to < 2 times/day)
† ref = medium (≥ 2 to < 4 times/day)

one time per day) is most closely related to Canada's Food Guide recommendations of one serving of meat and alternatives per day for children.[1]

These results are interesting as no published studies were found concerning the association between food group consumption and physician consultations.

General infections varied by year and peaked when children were five years old, the age when they first attended school in Québec. Our study found that children who consumed high amounts of fruits over time had a 5% reduction of general infections compared to children who consumed medium amounts of fruits. Only the fruit food group was in accordance with our hypothesis as it was the only food group to be associated with the frequency of infections.

The literature concerning dietary intake and infection is limited and inconsistent with other studies and with this study's finding. He et al.[57] showed a relationship between high consumption of both root vegetables and milk and decreased odds of having a cold with a cough. He et al.[57] also found and association between high consumption of milk and decreased odds of cold with phlegm. Although He et al.[57] did not have similar results to this study, there were similarities with the study by Hirota et al.[58] on Japanese children. The study Hirota et al.[58] found that frequent intake of vegetables or fruits other than green/yellow vegetables was associated with a decreased risk of mild influenza and the frequent intake of milk products was associated with a decreased risk of severe influenza. This finding may be the most consistent with our study's finding of fruit being protective against infection as Hirota et al's[58] food category of 'vegetables or fruit other than green/yellow vegetables' is meant to represent foods high in vitamin C (like fruit) and was protective against infections.

It is interesting that He et al.[57] and Hirota et al.[58] found that high milk or milk product consumption was protective against infection and our study did not. He et al.[57]

attributed this association to be due to vitamin A intake, but also notes that it may be a chance finding. Hirota et al.[58] states that the association between milk and severe influenza is difficult to interpret and hypothesized that the lifestyle associated with milk product consumption might be associated with the decrease in illness rather than the milk products themselves. Both of these studies did not control for SES or family income and used proxy measures such as parental education and 'room space per capita', therefore the explanation of the relationship found with milk may be reasonable.

There were also similarities between our study and the study by Li et al.'s[59] concerning vegetable and fruit intake in pregnant women. Li et al.[59] found pregnant women who consumed the highest quartile of vegetables and fruit had reduced risk of upper respiratory tract infections. This finding is similar to our study's as fruits were associated with infection, however, our study found no association between vegetable consumption and infection.[59]

The literature concerning infections is challenging to compare to our study as they have different outcomes. Our study considers general infections (including gastro-intestinal infections, ear infections, respiratory infections, and 'other' infections), He et al.'s[57] study considered only colds with cough/phlegm, Hirota et al.'s[58] study considered only mild and severe influenza, and Li et al.'s[59] study considers only upper respiratory tract infections. Technically, our study combined all these outcomes. Furthermore, the study by Hirota et al.[58] categorized the severity of illness and in our study, the severity of the infection was not considered. Moreover, these studies differed by means of study population, food group categories, covariates, and study design.

The number of antibiotic treatments decreased steadily over time with the slight exception of a rise occurring when the children were 4 years old. It was found that children who consumed high amounts of vegetables had a 9% decrease in the risk of antibiotic treatments compared to children who had medium consumption. Only the vegetable food group was in accordance with our hypothesis as it was the only food group to be associated with the frequency of infections.

The literature supporting this finding is limited as only one study was found. This study was a small observational study which involved providing fresh fruit two times per day to Australian Aboriginal children.[60] The average antibiotic prescription per month decreased from seven to one after six months of the fruit intervention.[60] Our study, although shows a trend that the fruit food group influences the number of antibiotic treatments, is difficult to compare to the Australian study as it does not contain information on the other food groups that the children were consuming and differs drastically in methods and sample.[60]

This section outlines the proposed mechanism to explain the relationships found in our study between vegetable and/or fruit consumption and lower risks of some adverse health outcomes. Vegetables and fruits are known to provide vitamins, nutrients and bioactive compounds which may stimulate the immune system as well as have antibacterial and antiviral effects.[15] In a review by Lampe[15] concerning the mechanisms of action of vegetables and fruits in health it was stated "nutrients and other constituents of fruits and vegetables have the potential to affect almost all aspects of the immune system". It was shown in our study's literature review that deficiencies in Vitamin A, C, folate, zinc, and selenium can negatively impact the immune system and cause an

increased risk of infections. All these vitamins and nutrients can be obtained from vegetables and fruits, which may explain why vegetables and fruits were found to be protective in our study.

Vegetables and fruits are usually seen grouped together in studies and food guides. However, our study suggests that vegetables and fruits each have a unique effect on the health outcomes of children. Vegetables and fruits have many similar essential nutrients, but also vary. For instance, vegetables tend to have more carotenoids and folate and fruit typically have more vitamin C.[59] This difference between fruits and vegetables may be associated with different effects on the immune system, which may provide some evidence to consider fruits and vegetables separately.

Low consumption of meat and alternatives was found to be associated with decreased physician consultations. This finding is consistent with the recommended servings of meat and alternatives, which is one serving per day. A possible explanation of why low consumption may be protective is that children who consume low amounts of meat and alternatives may be consuming more vegetables and fruits compared with high meat consumers. The Mediterranean Diet, which consists of eating mainly plant food (vegetables, fruits, legumes and grains) and low amounts of meat and alternatives, has been linked with many health benefits such as reduced overall mortality, cardiovascular diseases, mortality from cancer, and incidence of Parkinson's and Alzheimer's disease.[100]

There are some factors that should be noted about the health outcomes in our study. Antibiotics should be prescribed only for bacterial infections that the child's immune system is not capable of combating on its own; however, it is known that antibiotics are often inappropriately prescribed for viral infections.[53,101] Physician

consultations would be expected to be correlated with antibiotic treatments as a physician consultation is required to obtain an antibiotic prescription. Perhaps this is why both of these health outcomes had similar associations concerning the vegetable food group.

6.3 Limitations

There are some limitations of this study. Due to the nature of performing secondary analysis on existing data, certain variables which may be associated with the exposures or outcomes could not be taken into account. For example, additional factors that may influence children's number of physician consultations include the characteristics of the physician themselves and wait times.[29] Factors that may influence the number of general infections a child incurs that were not accounted for in this study include season (but for only 3 years), child's vitamin intake, exposure to second hand smoke, family size and home heating methods.[31,32,39,45,102-104] Factors that may influence antibiotic prescription that were not accounted for include physician characteristics and their trends such as the physician's awareness of antibiotic resistant bacteria.[35,41,105] Although the QLSCD does not have information on the variables mentioned above, the most important covariates outlined in the literature were available.

The QLSCD utilizes mainly self reported data. Although self reported data are not as accurate as measured data for many constructs, it is common in large studies such as the Canadian Community Health Survey and has been proven to be very valuable.[106]

There are limitations to calculating BMI using height and weight, as methods such as skin fold thickness may be more accurate in children. However, in this study the Cole method was used to calculate BMI, which is widely used and internationally recognized as a close approximation to adult overweight and obesity cut points.[80]

Since this study focused mainly on food group consumption, it lacks insight into the quality of foods consumed (such as refined grain versus whole grain), dietary patterns (such as eating balanced meals including foods from all food groups) and also the variety of foods consumed within each food group (such as eating a wide variety of fruits and not just one type).

The main food group measure was done by means of a Food Frequency Questionnaire, which does not account for actual amounts, servings, caloric values or nutrient intake. However, as noted in section 4.3.2, the frequencies of the foods consumed was associated with the number of servings when compared to a 24 hour recall survey that was conducted on a subset of children in cycle 5. The classification of the food groups was arbitrarily imposed as low, medium, and high because these categories were data driven to allow for ample category size during analysis. Therefore, each food group category was based on the distribution in the sample and do not reflect Canada's Food Guide serving sizes.

The necessary classification of the last category of the infection and the antibiotic treatment outcome variable as 'four' when truly it represents 'four or more' is a limitation of this study. However, it is not expected to have substantial impact on the results because there were very few children in the 'four or more' category. Furthermore, if the variable was treated as dichotomous there would be a loss of information on higher frequencies of infections and antibiotic treatments. If treated as ordinal, the variable would require collapsing categories as there would be small cell counts, especially in later years when children had fewer infections and antibiotic treatments. Furthermore, interpreting ordinal analyses is complex, especially so for this study as the independent

variables were time-varying and categorical. Consequently, the infection and antibiotic treatment variables were analyzed as counts.

There is also the possibility of misclassification. Misclassification could have arisen from the mothers having difficulties in recalling the correct information as some recall periods were as large as a year. Misclassification is also possible if the mother classifies the health outcome incorrectly. For instance, when the mothers are reporting physician consultations, they may not think of specifying that their child visited another type of physician such as a homeopathic doctor.

6.4 Strengths

Although there were limitations to this study, there were also numerous advantages. The study's data collection methodology was thoroughly planned and well executed which led to a partial non-response rate that was below 5%.[63] This indicated little unit nonresponse.[63] Also, incentives and other participant retainment methods were employed.[63] Furthermore, this study had a large sample size, which yielded high statistical power and results that could be generalizable to the population.

This study considered many parent and child characteristics which controlled for potential confounding, as well as identified characteristics that were important in the relationship between dietary intake and the health outcomes. Furthermore, the important covariates in the final models resulted in expected associations with the outcomes and also were in the predicted directions which provided validation of the results.

The dietary measures of food groups used in this study provide a realistic approach to assessing diet quality, as nutrients are not eaten in isolation and have

complex relationships with other nutrients.[107] Therefore, considering diet in terms of food groups may be more generalizable.

There are many advantages that are attributed to this study being longitudinal. Longitudinal studies are capable of determining effects over time, which allows for the direct measurement of change.[85,87] Within-individual changes can be accounted for in longitudinal studies, whereas in cross sectional studies only "between-individuals differences" can be studied.[85]

The thorough measures taken concerning missing data in this study provided strength in validating the results. Multiple imputation is considered to be the best manner in which to deal with missing data.[83] Furthermore, the type of missing data was considered carefully as different types can lead to bias and require specific methodological considerations. In this study, data could not be treated as 'missing completely at random' because dropouts were significantly associated with several known variables, and also there were relationships found between missingness and the value of some variables.[83,85] Data 'not missing at random' means the reasons for missingness are related to characteristics that are unobserved.[85] Although it is impossible to rule out 'not missing at random', it is less likely as it was found that the outcomes were longitudinally unassociated with different patterns of dropping out of the study. For these reasons, we have assumed that the data were likely 'missing at random'.[83,85]

In longitudinal studies, each child acts as their own control which eliminates confounding from within-subject differences.[85] Consequently, even factors that have not been measured which may influence the effects are eliminated when each child's

response is compared year to year.[85] According to Fitzmaurice[85], these eliminations allow for a "very precise estimation of change".

The methods used in this study allowed for the retention of children even if the child's data was not complete for the study's entire duration. For instance, multivariable repeated measure ANOVA would require the removal of all individuals with missing responses at any time point.[85] GEE allowed for the retention of these children, even if they were missing some years of data or dropped out of the study.[85]

6.5 Proposed Future Research

This study provided valuable information on how children's eating habits can influence their health over time. However, these results prompt new questions to arise. For instance, it is unknown whether or not the quality of food consumed influences the health outcomes in a similar manner, such as consuming mainly whole wheat grain products versus refined grains. Also, it is unknown whether or not factors such as the variety of food consumed in each food group influence health outcomes. Additional studies to validate the proposed mechanisms of association between food groups and health outcomes are also required. As this is the first longitudinal study of its kind, further studies are needed to validate the results and to determine whether or not these results can be replicated in other populations.

6.6 Conclusion

Nutrition is an important part of every child's growth and development. In children, following the nutritional recommendations of "Eating Well with Canada's Food Guide" (2007) ensures that adequate energy, macronutrients (carbohydrates, lipids and

proteins) vitamins, minerals and other nutrients are consumed.[1] However, it is well known that many Canadian children are not eating as recommended.[3-9] Furthermore, it has been shown that the dietary habits formed during childhood are likely carried throughout life.[14,108] Therefore, forming good dietary habits in children is important, not just to reduce the risks of the health outcomes as discussed in this study, but also to prevent other conditions that are associated with poor diet quality in adults such as overweight and obesity, cardiovascular diseases, cancer, and all-cause mortality.[10-12]

Since we have shown that good dietary intake does indeed have beneficial effects on health outcomes such as frequency of physician consultations, infections, and antibiotic treatments, this research could provide a foundation for policies and also strengthen the evidence for children to eat according to Canada's Food Guide as recommended. Furthermore, educational public health programs promoting proper nutrition in children could be created or strengthened, which could in turn decrease health care burdens.

Reference List

(1) Health Canada [Internet]. Eating Well with Canada's Food Guide. 2007; Available at: http://www.hc-sc.gc.ca/fn-an/food-guide-aliment/index-eng.php. Accessed 10/26, 2009.

(2) Arvaniti F, Panagiotakos DB. Healthy indexes in public health practice and research: a review. Crit.Rev.Food Sci.Nutr. 2008 Apr;48(4):317-327.

(3) L'alimentation des jeunes québécois : un premier tour de table - Enquête sur la santé dans les collectivités canadiennes - Nutrition (2004). 2008.

(4) Hanning RM, Woodruff SJ, Lambraki I, Jessup L, Driezen P, Murphy CC. Nutrient intakes and food consumption patterns among Ontario students in grades six, seven, and eight. Can.J.Public Health 2007 Jan-Feb;98(1):12-16.

(5) Moffat T, Galloway T. Food consumption patterns in elementary school children. Can.J.Diet.Pract.Res. 2008 Fall;69(3):152-154.

(6) St John M, Durant M, Campagna PD, Rehman LA, Thompson AM, Wadsworth LA, et al. Overweight Nova Scotia children and youth: the roles of household income and adherence to Canada's Food Guide to Healthy Eating. Can.J.Public Health 2008 Jul-Aug;99(4):301-306.

(7) Sylvestre M-, O'Loughlin J, Gray-Donald K, Hanley J, Paradis G. Association between fruit and vegetable consumption in mothers and children in low-income, urban neighborhoods. Health Education and Behavior 2007;34(5):723-734.

(8) Taylor JP, Timmons V, Larsen R, Walton F, Bryanton J, Critchley K, et al. Nutritional concerns in aboriginal children are similar to those in non-aboriginal children in Prince Edward Island, Canada. J.Am.Diet.Assoc. 2007 Jun;107(6):951-955.

(9) Garriguet D. Canadians' eating habits. Health Rep. 2007;18(2):17-32.

(10) Kant AK. Dietary patterns and health outcomes. Journal of the American Dietetic Association 2004;104(4):615-635.

(11) Van Duyn MA, Pivonka E. Overview of the health benefits of fruit and vegetable consumption for the dietetics professional: selected literature. J.Am.Diet.Assoc. 2000 Dec;100(12):1511-1521.

(12) Togo P, Osler M, Sorensen TI, Heitmann BL. Food intake patterns and body mass index in observational studies. Int.J.Obes.Relat.Metab.Disord. 2001 Dec;25(12):1741-1751.

(13) World Health Organization [Internet]. Nutrition. 2010; Available at: http://www.who.int/topics/nutrition/en/. Accessed 04/25, 2010.

(14) The Report of the Dietary Guidelines Advisory Committee on Dietary Guidelines for Americans, 2005. 2005.

(15) Lampe JW. Health effects of vegetables and fruit: Assessing mechanisms of action in human experimental studies. American Journal of Clinical Nutrition 1999;70(3 SUPPL.):475S-490S.

(16) Phillips S, Jacobs Starkey L, Gray-Donald K. Food habits of Canadians: food sources of nutrients for the adolescent sample. Can.J.Diet.Pract.Res. 2004 Summer;65(2):81-84.

(17) Briefel RR, Johnson CL. Secular trends in dietary intake in the United States. Annu.Rev.Nutr. 2004;24:401-431.

(18) Bloem MW, Wedel M, Egger RJ, Speek AJ, Schrijver J, Saowakontha S, et al. Mild vitamin A deficiency and risk of respiratory tract diseases and diarrhea in preschool and school children in northeastern Thailand. Am.J.Epidemiol. 1990 Feb;131(2):332-339.

(19) Barreto ML, Santos LM, Assis AM, Araujo MP, Farenzena GG, Santos PA, et al. Effect of vitamin A supplementation on diarrhoea and acute lower-respiratory-tract infections in young children in Brazil. Lancet 1994 Jul 23;344(8917):228-231.

(20) Dibley MJ, Sadjimin T, Kjolhede CL, Moulton LH. Vitamin A supplementation fails to reduce incidence of acute respiratory illness and diarrhea in preschool-age Indonesian children. J.Nutr. 1996 Feb;126(2):434-442.

(21) Vijayaraghavan K, Radhaiah G, Prakasam BS, Sarma KV, Reddy V. Effect of massive dose vitamin A on morbidity and mortality in Indian children. Lancet 1990 Dec 1;336(8727):1342-1345.

(22) Sommer A, Tarwotjo I, Katz J. Increased risk of xerophthalmia following diarrhea and respiratory disease. Am.J.Clin.Nutr. 1987 May;45(5):977-980.

(23) Ekweagwu E, Agwu AE, Madukwe E. The role of micronutrients in child health: A review of the literature. Afr.J.Biotechnol. 2008;7(21):3804-3810.

(24) Villamor E, Fawzi WW. Effects of vitamin a supplementation on immune responses and correlation with clinical outcomes. Clin.Microbiol.Rev. 2005 Jul;18(3):446-464.

(25) Douglas RM, Hemila H, Chalker E, Treacy B. Vitamin C for preventing and treating the common cold. Cochrane Database Syst.Rev. 2007 Jul 18;(3)(3):CD000980.

(26) Dhur A, Galan P, Hercberg S. Folate status and the immune system. Prog.Food Nutr.Sci. 1991;15(1-2):43-60.

(27) Bunout D, Barrera G, Leiva L, Gattas V, de la Maza MP, Haschke F, et al. Effect of a nutritional supplementation on bone health in Chilean elderly subjects with femoral osteoporosis. J.Am.Coll.Nutr. 2006 Jun;25(3):170-177.

(28) Rasmussen M, Krolner R, Klepp KI, Lytle L, Brug J, Bere E, et al. Determinants of fruit and vegetable consumption among children and adolescents: a review of the literature. Part I: Quantitative studies. Int.J.Behav.Nutr.Phys.Act. 2006 Aug 11;3:22.

(29) Janicke DM, Finney JW. Determinants of children's primary health care use. J.Clin.Psychol.Med.Settings 2000;7(1):29-39.

(30) Twaij M. Urinary tract infection in children: A review of its pathogenesis and risk factors. Journal of The Royal Society for the Promotion of Health 2000;120(4):220-226.

(31) Witorsch RJ, Witorsch P. Environmental tobacco smoke and respiratory health in children: A critical review and analysis of the literature from 1969 to 1998. Indoor Built Environ. 2000;9(5):246-264.

(32) Dhooge IJM. Risk factors for the development of otitis media. Current Allergy and Asthma Reports 2003;3(4):321-325.

(33) Falagas ME, Mourtzoukou EG, Vardakas KZ. Sex differences in the incidence and severity of respiratory tract infections. Respir.Med. 2007 Sep;101(9):1845-1863.

(34) Sedberry-Ross S, Pohl HG. Urinary tract infections in children. Curr.Urol.Rep. 2008 Mar;9(2):165-171.

(35) Rossignoli A, Clavenna A, Bonati M. Antibiotic prescription and prevalence rate in the outpatient paediatric population: analysis of surveys published during 2000-2005. Eur.J.Clin.Pharmacol. 2007 Dec;63(12):1099-1106.

(36) Lu N, Samuels ME. Increased health care utilization associated with child day care among health maintenance organization and Medicaid enrollees. Ambul.Child Health 2001;7(3-4):219-230.

(37) Brownell M, Kozyrkyj A, Roos N, Friesen D, Mayer T, Sullivan K. Health service utilization by Manitoba children. Can.J.Public Health 2002 Nov-Dec;93 Suppl 2:S57-62.

(38) Nikièma B, Zunzunegui MV, Séguin L, Gauvin L, Potvin L. Poverty and cumulative hospitalization in infancy and early childhood in the Québec birth cohort: A puzzling pattern of association. Matern.Child Health J. 2008;12(4):534-544.

(39) Taussig LM, Wright AL, Holberg CJ, Halonen M, Morgan WJ, Martinez FD. Tucson Children's Respiratory Study: 1980 to present. J.Allergy Clin.Immunol. 2003;111(4):661-675.

(40) Nyquist AC, Gonzales R, Steiner JF, Sande MA. Antibiotic prescribing for children with colds, upper respiratory tract infections, and bronchitis. JAMA 1998 Mar 18;279(11):875-877.

(41) Marra F, Patrick DM, Chong M, Bowie WR. Antibiotic use among children in British Columbia, Canada. J.Antimicrob.Chemother. 2006 Oct;58(4):830-839.

(42) Summerbell CD, Waters E, Edmunds LD, Kelly S, Brown T, Campbell KJ. Interventions for preventing obesity in children. Cochrane Database Syst.Rev. 2005 Jul 20;(3)(3):CD001871.

(43) Taveras EM, Berkey CS, Rifas-Shiman SL, Ludwig DS, Rockett HRH, Field AE, et al. Association of consumption of fried food away from home with body mass index and diet quality in older children and adolescents. Pediatrics 2005;116(4).

(44) Strock GA, Cottrell ER, Abang AE, Buschbacher RM, Hannon TS. Childhood obesity: A simple equation with complex variables. J.Long-Term Eff.Med.Implants 2005;15(1):15-32.

(45) Jedrychowski W, Maugeri U, Flak E, Mroz E, Bianchi I. Predisposition to acute respiratory infections among overweight preadolescent children: an epidemiologic study in Poland. Public Health 1998 May;112(3):189-195.

(46) Wintergerst ES, Maggini S, Hornig DH. Contribution of selected vitamins and trace elements to immune function. Ann.Nutr.Metab. 2007;51(4):301-323.

(47) Veugelers PJ, Fitzgerald AL, Johnston E. Dietary intake and risk factors for poor diet quality among children in Nova Scotia. Can.J.Public Health 2005 May-Jun;96(3):212-216.

(48) Patrick H, Nicklas TA. A review of family and social determinants of children's eating patterns and diet quality. J.Am.Coll.Nutr. 2005 Apr;24(2):83-92.

(49) Taylor JP, Evers S, McKenna M. Determinants of healthy eating in children and youth. Can.J.Public Health 2005 Jul-Aug;96 Suppl 3:S20-6, S22-9.

(50) Hjern A, Haglund B, Rosen M. Socioeconomic differences in use of medical care and antibiotics among schoolchildren in Sweden. Eur.J.Public Health 2001 Sep;11(3):280-283.

(51) Gorman BK, Braverman J. Family structure differences in health care utilization among U.S. children. Soc.Sci.Med. 2008 Dec;67(11):1766-1775.

(52) Thrane N, Olesen C, Schonheyder HC, Sorensen HT. Socioeconomic factors and prescription of antibiotics in 0- to 2-year-old Danish children. J.Antimicrob.Chemother. 2003 Mar;51(3):683-689.

(53) Kozyrskyj AL, Dahl ME, Chateau DG, Mazowita GB, Klassen TP, Law BJ. Evidence-based prescribing of antibiotics for children: role of socioeconomic status and physician characteristics. CMAJ 2004 Jul 20;171(2):139-145.

(54) Janicke DM, Finney JW. Children's Primary Health Care Services: Social-Cognitive Factors Related to Utilization. J.Pediatr.Psychol. 2003;28(8):547-557.

(55) Mangione-Smith R, Elliott MN, Stivers T, McDonald L, Heritage J, McGlynn EA. Racial/ethnic variation in parent expectations for antibiotics: implications for public health campaigns. Pediatrics 2004 May;113(5):e385-94.

(56) Heck KE, Parker JD. Family structure, socioeconomic status, and access to health care for children. Health Serv.Res. 2002 Feb;37(1):173-186.

(57) He QQ, Wong TW, Du L, Lin GZ, Gao Y, Jiang ZQ, et al. Nutrition and children's respiratory health in Guangzhou, China. Public Health 2008 Dec;122(12):1425-1432.

(58) Hirota Y, Takeshita S, Ide S, Kataoka K, Ohkubo A, Fukuyoshi S, et al. Various factors associated with the manifestation of influenza-like illness. Int.J.Epidemiol. 1992 Jun;21(3):574-582.

(59) Li L, Werler MM. Fruit and vegetable intake and risk of upper respiratory tract infection in pregnant women. Public Health Nutr. 2010 Feb;13(2):276-282.

(60) Jones R, Smith F. Are there health benefits from improving basic nutrition in a remote Aboriginal community? Aust.Fam.Physician 2006;35(6):453-454.

(61) Pai M, McCulloch M, Gorman JD, Pai N, Enanoria W, Kennedy G, et al. Systematic reviews and meta-analyses: an illustrated, step-by-step guide. Natl.Med.J.India 2004 Mar-Apr;17(2):86-95.

(62) Dwan K, Altman DG, Arnaiz JA, Bloom J, Chan A-, Cronin E, et al. Systematic review of the empirical evidence of study publication bias and outcome reporting bias. PLoS ONE 2008;3(8).

(63) Jetté M, Des Groseilliers L. "Survey Description and Methodology" in Longitudinal Study of Child Development in Québec (ELDEQ 1998-2002). 2000;Vol. 1, No. 1.

(64) Institut de la statistique Québec [Internet]. Québec Longitudinal Study of Child Development. 2006; Available at: http://www.jesuisjeserai.stat.gouv.qc.ca.proxy.bib.uottawa.ca/default_an.htm. Accessed June/10, 2009.

(65) Guide de l'utilisateur de la banque de données du volet 2006 (Enfants de 92 à 104 mois). Étude longitudinale du développement des enfants du Québec - Phase II (ÉLDEQ

2003-2010. Institut de la statistique Québec. 2007; Available at: http://www.jesuisjeserai.stat.gouv.qc.ca/doc_tech_an.htm.

(66) Thibault JK, Jetté M, Gingras L. "Concepts, Definitions and Operational Aspects, Part I - QLSCD: Overview of the Study Instruments for the 1999 and 2000 Rounds", in Québec Longitudinal Study of Child Development (QLSCD 1998-2002) - From Birth to 29 Month, Volume. 2, No. 12. 2003.

(67) ÉLDEQ 1998-2002 : retour sur l'étude, Description et méthodologie de la phase I de l'ÉLDE. Institut de la statistique du Québec, Direction Santé Québec. 2004; Available at: http://www.jesuisjeserai.stat.gouv.qc.ca.proxy.bib.uottawa.ca/pdf/doc_tech/E4-E5Retour_sur_ELDEQ.pdf. Accessed 05/14, 2009.

(68) Guide de l'utilisateur de la banque de données du volet 2005 (Enfants de 80 à 92 mois). Étude longitudinale du développement des enfants du Québec - Phase II (ÉLDEQ 2003-2010). 2006; Available at: http://www.jesuisjeserai.stat.gouv.qc.ca/doc_tech_an.htm.

(69) Guide de l'utilisateur de la banque de données du volet 2003 (Enfants de 56 à 68 mois). Étude longitudinale du développement des enfants du Québec - Phase II (ÉLDEQ 2003-2010). 2004; Available at: http://www.jesuisjeserai.stat.gouv.qc.ca/doc_tech_an.htm. Accessed 04/11, 2009.

(70) Plante N, Courtemanche R. Pondération des données du volet 2005. Institut de la statistique du Québec. 2006; Available at: http://www.jesuisjeserai.stat.gouv.qc.ca/doc_tech_an.htm. Accessed 04/11, 2009.

(71) Plante N, Fontaine C, Courtemanche R. Pondération des données du volet 2004. Institut de la statistique du Québec. 2005; Available at: http://www.jesuisjeserai.stat.gouv.qc.ca.proxy.bib.uottawa.ca/doc_tech_an.htm. Accessed 04/11, 2009.

(72) Plante N, Tremblay M, Courtemanche R. Pondération des données du volet 2003 : poids longitudinaux et transversaux. Institut de la statistique du Québec. 2004; Available at: http://www.jesuisjeserai.stat.gouv.qc.ca.proxy.bib.uottawa.ca/doc_tech_an.htm. Accessed 04/11, 2009.

(73) Plante N, Tremblay M, Courtemanche R. Pondération longitudinale : volets 1998 à 2001 (ÉLDEQ 1998-2002). Institut de la statistique du Québec. Available at: http://www.jesuisjeserai.stat.gouv.qc.ca.proxy.bib.uottawa.ca/doc_tech_an.htm. Accessed 04/11, 2009.

(74) "Étude longitudinale du développement des enfants du Québec (ÉLDEQ), Pondération longitudinale : volets 1998 à 2002". 2003. Institut de la statistique du Quebec. Available at: http://www.jesuisjeserai.stat.gouv.qc.ca/pdf/doc_tech/E4-E5PonderationE1-E5.pdf. Accessed 09/17, 2009.

(75) Thibault J, Jetté M, Desrosiers H. "Concepts, Definitions and Operational Aspects, Part I - Design of Phase I of the ELDEQ, Instruments and Procedures" in Longitudinal Study of Child Development in Québec, Québec, Institut de la statisitque du Québec, Vol. 1, No. 12. 2001.

(76) Canadian Health Services Research Foundation. Discussion Paper: Pharmacare in Canada. 2002; Available at: http://www.chsrf.ca/. Accessed 04/02, 2010.

(77) Signorello LB, Buchowski MS, Cai Q, Munro HM, Hargreaves MK, Blot WJ. Biochemical validation of food frequency questionnaire-estimated carotenoid, α-tocopherol, and folate intakes among African Americans and non-hispanic whites in the southern community cohort study. Am.J.Epidemiol. 2010;171(4):488-497.

(78) Steffen LM, Jacobs Jr. DR, Murtaugh MA, Moran A, Steinberger J, Hong C-, et al. Whole grain intake is associated with lower body mass and greater insulin sensitivity among adolescents. Am.J.Epidemiol. 2003;158(3):243-250.

(79) Tatone-Tokuda F. "Examining the association between sleep duration, diet and body mass index in Québec children." MS thesis U of Ottawa, Ottawa. 2008.

(80) Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition for child overweight and obesity worldwide: international survey. BMJ 2000 May 6;320(7244):1240-1243.

(81) Exclusive Breastfeeding Duration - 2004 Health Canada Recommendations. 2004;2009(June, 2009).

(82) Health Canada [Internet]. Food and Nutrition "Body Mass Index (BMI) Nomogram". 2003; Available at: http://www.hc-sc.gc.ca/fn-an/nutrition/weights-poids/guide-ld-adult/bmi_chart_java-graph_imc_java-eng.php. Accessed August/20, 2009.

(83) Allison PA. Missing Data. Sage University Papers series on Quantitative Applications in the Social Sciences. Thousand Oaks, CA: Sage Publications, Inc.; 2001.

(84) Aschengrau A, Seage G, R. Essentials of Epidemiology in Public Health. 2nd ed. United States of America: Jones and Baetlett Publishers, Inc.; 2008.

(85) Fitzmaurice GM, Laird NM, Ware JH. Applied Longitudinal Analysis. Hoboken, New Jersey: Wiley-Interscience, A John Wiley & Sons, Inc.; 2004.

(86) Weiss RE. Modeling Longitudinal Data. New York, NY: Springer Science + Business Media Inc.; 2005.

(87) Song P. Correlated Data Analysis: Modeling, Analytic, and Applications. New York: Springer Science+Business Media; 2007.

(88) Stokes M, Davis C, Koch G. Categorical Data Analysis Using the SAS System. 2nd ed.; 2000.

(89) Twisk J, de Vente W. Attrition in longitudinal studies. How to deal with missing data. J.Clin.Epidemiol. 2002 Apr;55(4):329-337.

(90) Siddiqui O, Ali MW. A comparison of the random-effects pattern mixture model with last-observation-carried-forward (LOCF) analysis in longitudinal clinical trials with dropouts. J.Biopharm.Stat. 1998 Nov;8(4):545-563.

(91) Saha C, Jones MP. Bias in the last observation carried forward method under informative dropout. J.Stat.Plann.Inference 2009;139(2):246-255.

(92) Molenberghs G, Verbeke G. Models for Discrete Longitudinal Data. : Springer Series in Statistics; 2005.

(93) Hardin JW, Hirsch S. Generalized Estimating Equations. Boca Ranton, Florida: CRC Press LLC; 2000.

(94) SAS Institute Inc. SAS/STAT® 9.2 User's Guide. : Cary, NC: SAS Institute Inc.; 2008.

(95) Xie F, Paik MC. Multiple imputation methods for the missing covariates in generalized estimating equation. Biometrics 1997 Dec;53(4):1538-1546.

(96) SAS Online Doc, Version. SAS Institute Inc. Cary, NC, USA. 1999; Available at: http://www.uc.edu/sashtml/.

(97) Ake CF. Rounding after Multiple Imputation with Non-binary Categorical Covariates. Sugi 30 Focus Session .

(98) Lohr SL. Sampling: Design and Analysis. Pacific Grove, CA: Duxbury Press; 1999.

(99) Korn EL, Graubard BI. Analysis of Large Health Surveys: Accounting for the Sampling Design. J. R. Statist. Soc. A 1995;158(2):263.

(100) Sofi F, Cesari F, Abbate R, Gensini GF, Casini A. Adherence to Mediterranean diet and health status: meta-analysis. BMJ 2008;337.

(101) Cadieux G, Tamblyn R, Dauphinee D, Libman M. Predictors of inappropriate antibiotic prescribing among primary care physicians. CMAJ 2007 Oct 9;177(8):877-883.

(102) Lubianca Neto JF, Hemb L, Brunelli E Silva D. Systematic literature review of modifiable risk factors for recurrent acute otitis media in childhood. J.Pediatr. 2006;82(2):87-96.

(103) Purssell E. Upper respiratory tract infection in infants from a nutritional perspective. J.Fam.Health.Care. 2009;19(5):164-168.

(104) Stephen AI, Avenell A. A systematic review of multivitamin and multimineral supplementation for infection. J.Hum.Nutr.Diet. 2006 Jun;19(3):179-190.

(105) Paluck E, Katzenstein D, Frankish CJ, Herbert CP, Milner R, Speert D, et al. Prescribing practices and attitudes toward giving children antibiotics. Can.Fam.Physician 2001 Mar;47:521-527.

(106) Canadian Community Health Survey (CCHS) –Annual Component User Guide - 2008 Microdata File. Statistic Canada. 2009; Available at: http://www.statcan.gc.ca/imdb-bmdi/document/3226_D7_T9_V5-eng.pdf. Accessed 01/06, 2010.

(107) Hu FB. Dietary pattern analysis: a new direction in nutritional epidemiology. Curr.Opin.Lipidol. 2002 Feb;13(1):3-9.

(108) Nicklaus S, Boggio V, Chabanet C, Issanchou S. A prospective study of food variety seeking in childhood, adolescence and early adult life. Appetite 2005 Jun;44(3):289-297.

Appendix A: MEDLINE Search Strategy

1. exp Food/
2. diet/ or nutritional requirements/ or nutritional status/
3. Food Habits/
4. Child Nutrition Disorders/
5. 1 or 2 or 3 or 4
6. Anti-Bacterial Agents/
7. exp Infection/
8. gastroenteritis/ or dysentery/ or esophagitis/ or gastritis/
9. exp Otitis Media/
10. 7 or 8 or 9
11. Office Visits/
12. ((Doctor$ or physician$ or pediatric$ or medic$) adj (visit$ or consultation$ or appointment$ or medic$)).tw.
13. 11 or 12
14. hospitalization/ or "length of stay"/ or patient readmission/
15. 5 and 6
16. 5 and 10
17. 5 and 13
18. 5 and 14
19. 15 or 16 or 17 or 18
20. limit 19 to ("all child (0 to 18 years)" and last 10 years)-

Appendix B: Imputation Values

Table 24: Number of observed and imputed values for all variables by child's age.

	Age 1.5		Age 2.5		Age 3.5		Age 4.5		Age 5		Age 6		Age 7	
	Observed	Imputed	Observed	Imputed	Observed	Imputed	Observed	Imputed	Observed	Imputed	Observed	Imputed	Observed	Imputed
Dependent Variables														
Physician consultations	2044	13	1995	62	1950	107	1861	196	1759	298	1316	741	1322	735
Infections	2045	12	1997	60	1950	107	1941	116	1759	298	1492	565	1528	529
Antibiotic treatments	2045	12	1994	63	1950	107	1944	113	1757	300	1492	565	1527	530
Independent Variables														
Grain products	2044	13	1996	61	1949	108	1944	113	1758	299	1492	565	1322	735
Vegetables	2043	14	1995	62	1947	110	1943	114	1756	301	1492	565	1314	743
Fruits	2041	16	1996	61	1949	108	1942	115	1756	301	1492	565	1312	745
Milk and alternatives	2044	13	1996	61	1949	108	1944	113	1758	299	1492	565	1322	735
Meat and alternatives	2044	13	1996	61	1949	108	1944	113	1758	299	1492	565	1322	735
Child/Birth Characteristics														
Sex	2057	0												
Chronic condition	2045	12	-	-	1950	107	-	-	1759	298	1492	565	-	-
BMI	1771	286	1867	190	1832	225	1865	192	1237	820	1134	923	1474	583
Child care	2045	12	1995	62	1950	107	1941	116	1698	359	1492	565	-	-
Breastfed exclusively for ≥ 3 months	2057	0												
Parental/Family Characteristics														
SES	2029	28	1974	83	1931	126	-	-	1747	310	1482	575	1521	536
Age group of mother	2042	15												
Mother's immigrant status	2041	16												
Mother's BMI	2008	49												
Family location	2010	47												

Appendix C: Multiple Imputation Models

Table 25: Risk ratios, 95% confidence intervals, and p-values for the imputed physician consultation model.

Characteristic	Category	Risk ratio	95% confidence intervals	P-value ($p \leq 0.05$)
Year		**0.73**	**0.72, 0.74**	**<.0001**
Year 2		**1.03**	**1.02, 1.03**	**<.0001**
Grain products	Low (< 2 times/day)*	0.99	0.97, 1.01	0.3571
	High (≥ 3 times/day)*	**1.03**	**1.01, 1.06**	**0.0098**
Vegetable	Low (< 1 time/day)‡	1.01	0.98, 1.03	0.6902
	High (≥ 2 times/day)‡	**0.96**	**0.94, 0.99**	**0.0018**
Fruit	Low (< 1 time/day)‡	**0.97**	**0.94, 0.99**	**0.0208**
	High (≥ 2 times/day)‡	**1.05**	**1.03, 1.07**	**<.0001**
Milk & alternatives	Low (< 2 times/day)†	0.94	0.87, 1.02	0.1151
	High (≥ 4 times/day)†	0.98	0.95, 1.01	0.1862
Meat & alternatives	Low (< 1 time/day)‡	**0.96**	**0.94, 0.98**	**0.0011**
	High (≥ 2 times/day)‡	0.99	0.94, 1.03	0.5402
Sex	Male (ref=female)	**1.06**	**1.04, 1.08**	**<.0001**
Chronic condition	Yes (ref=no)	**1.41**	**1.37, 1.46**	**<.0001**
Child's BMI	Obese (ref=normal weight)	1.02	0.98, 1.06	0.4101
	Overweight (ref=normal weight)	0.98	0.94, 1.01	0.1671
Childcare	Child in daycare (ref=at home)	**1.24**	**1.20, 1.27**	**<.0001**
	Care outside of home (ref=at home)	**1.06**	**1.03, 1.09**	**<.0001**
Breastfed	Exclusively ≥ 3 months (ref=no)	**0.95**	**0.92, 0.97**	**<.0001**
SES (tertiles)	Lowest third (ref=mid)	1.00	0.97, 1.03	0.9673
	Highest third (ref=mid)	**1.04**	**1.02, 1.07**	**0.0017**
Mother's age (years)	Less than 25 (ref = ≥ 35)	**1.20**	**1.16, 1.24**	**<.0001**
	25 to less than 30 (ref = ≥ 35)	**1.10**	**1.06, 1.13**	**<.0001**
	30 to less than 35 (ref = ≥ 35)	**1.07**	**1.04, 1.10**	**<.0001**
Mother's BMI	Obese (ref=normal weight)	**0.91**	**0.87, 0.94**	**<.0001**
	Overweight (ref=normal weight)	1.03	1.00, 1.07	0.0907
Mother immigrant	Yes (ref=no)	**1.05**	**1.02, 1.08**	**0.0032**
Milk*year	Low milk*year†	1.01	0.99, 1.03	0.4438
	High milk*year†	**1.03**	**1.01, 1.04**	**<.0001**

* ref = ≥ 2 to < 3 times/day
‡ ref = ≥ 1 to < 2 times/day
† ref = ≥ 2 to < 4 times/day

Table 26: Risk ratios, 95% confidence intervals, and p-values for the imputed infection model.

Characteristic	Category	Risk ratio	95% confidence intervals	P-value ($p \leq 0.05$)
Year	Year 1 (ref=year 0)	**0.78**	**0.73, 0.82**	**<.0001**
	Year 2 (ref=year 0)	**0.81**	**0.76, 0.86**	**<.0001**
	Year 3 (ref=year 0)	**0.85**	**0.81, 0.90**	**<.0001**
	Year 4 (ref=year 0)	**1.27**	**1.21, 1.34**	**<.0001**
	Year 5 (ref=year 0)	**1.21**	**1.14, 1.29**	**<.0001**
	Year 6 (ref=year 0)	0.94	0.89, 1.00	0.0506
Grain products	Low (< 2 times/day)*	0.99	0.95, 1.02	0.3833
	High (≥ 3 times/day)*	1.02	0.99, 1.06	0.2426
Vegetable	Low (< 1 time/day)‡	1.03	0.99, 1.08	0.1845
	High (≥ 2 times/day)‡	0.99	0.96, 1.03	0.7435
Fruit	Low (< 1 time/day)‡	0.98	0.94, 1.02	0.3895
	High (≥ 2 times/day)‡	0.97	0.94, 1.01	0.1363
Milk & alternatives	Low (< 2 times/day)†	1.01	0.96, 1.07	0.6169
	High (≥ 4 times/day)†	0.99	0.96, 1.02	0.5682
Meat & alternatives	Low (< 1 time/day)‡	1.00	0.97, 1.03	0.8856
	High (≥ 2 times/day)‡	1.03	0.97, 1.09	0.3507
Sex	Male (ref=female)	0.98	0.95, 1.01	0.2797
Chronic condition	Yes (ref=no)	**1.33**	**1.28, 1.39**	**<.0001**
BMI	Obese (ref=normal weight)	0.95	0.89, 1.01	0.0841
	Overweight (ref=normal weight)	**0.95**	**0.91, 1.00**	**0.0395**
Childcare	Child in daycare (ref=at home)	**1.22**	**1.17, 1.27**	**<.0001**
	Care outside of home (ref=at home)	**1.09**	**1.05, 1.14**	**<.0001**
Breastfed	Exclusively ≥ 3 months (ref=no)	**0.95**	**0.92, 0.99**	**0.0094**
SES (tertiles)	Lowest third (ref=mid)	1.03	0.99, 1.06	0.1714
	Highest third (ref=mid)	**0.94**	**0.90, 0.97**	**0.0009**
Immigrant mother	Yes (ref=no)	**0.81**	**0.77, 0.86**	**<.0001**
Mother's BMI	Obese (ref=normal weight)	**1.12**	**1.07, 1.18**	**<.0001**
	Overweight (ref=normal weight)	**1.05**	**1.01, 1.09**	**0.0125**
Living location	Urban (ref=rural)	**1.05**	**1.01, 1.09**	**0.0184**

* ref = ≥ 2 to < 3 times/day
‡ ref = ≥ 1 to < 2 times/day
† ref = ≥ 2 to < 4 times/day

Table 27: Risk ratios, 95% confidence intervals, and p-values for the imputed antibiotic treatment model.

Characteristic	Category	Risk ratio	95% confidence intervals	P-value ($p \leq 0.05$)
Year	Year 1 (ref=year 0)	**0.69**	**0.65, 0.74**	**<.0001**
	Year 2 (ref=year 0)	**0.59**	**0.55, 0.64**	**<.0001**
	Year 3 (ref=year 0)	**0.59**	**0.55, 0.64**	**<.0001**
	Year 4 (ref=year 0)	**0.52**	**0.48, 0.57**	**<.0001**
	Year 5 (ref=year 0)	**0.45**	**0.42, 0.50**	**<.0001**
	Year 6 (ref=year 0)	**0.33**	**0.30, 0.36**	**<.0001**
Grain products	Low (< 2 times/day)*	1.01	0.96, 1.06	0.6311
	High (≥ 3 times/day)*	1.03	0.98, 1.09	0.2150
Vegetable	Low (< 1 time/day)‡	0.98	0.92, 1.04	0.5378
	High (≥ 2 times/day)‡	**0.91**	**0.86, 0.96**	**0.0003**
Fruit	Low (< 1 time/day)‡	0.97	0.92, 1.03	0.3464
	High (≥ 2 times/day)‡	0.96	0.92, 1.02	0.1698
Milk & alternatives	Low (< 2 times/day)†	0.93	0.86, 1.00	0.0513
	High (≥ 4 times/day)†	1.02	0.97, 1.07	0.3958
Meat & alternatives	Low (< 1 time/day)‡	0.98	0.94, 1.02	0.3234
	High (≥ 2 times/day)‡	1.03	0.95, 1.12	0.4198
Sex	Male (ref=female)	1.00	0.96, 1.04	0.8334
Chronic condition	Yes (ref=no)	**1.51**	**1.43, 1.60**	**<.0001**
BMI	Obese (ref=normal weight)	0.97	0.89, 1.04	0.3808
	Overweight (ref=normal weight)	0.95	0.90, 1.01	0.0780
Childcare	Child in daycare (ref=at home)	**1.49**	**1.40, 1.58**	**<.0001**
	Care outside of home (ref=at home)	**1.13**	**1.07, 1.19**	**<.0001**
Breastfed	Exclusively ≥ 3 months (ref=no)	**0.83**	**0.79, 0.87**	**<.0001**
SES (tertiles)	Lowest third (ref=mid)	**1.08**	**1.02, 1.14**	**0.0055**
	Highest third (ref=mid)	0.98	0.93, 1.03	0.4749

* ref = ≥ 2 to < 3 times/day
‡ ref = ≥ 1 to < 2 times/day
† ref = ≥ 2 to < 4 times/day