The Burden of Biopsy-Proven Pediatric Celiac Disease in Ontario, Canada: Derivation of Health Administrative Data Algorithms and Determination of Health Services Utilization

Jason Chan

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

School of Epidemiology, Public Health and Preventive Medicine

Faculty of Medicine

University of Ottawa

Abstract

The Burden of Biopsy-Proven Pediatric Celiac Disease in Ontario, Canada: Derivation of Health Administrative Data Algorithms and Determination of Health Services Utilization

Introduction: The main objective of this thesis is to develop an algorithm to accurately identify cases of biopsy-proven Celiac Disease (CD) in children aged 6 months-14 years old from Ontario health administrative data.

Method: CD cases diagnosed in 2005-2011 were identified from CHEO, and linked to the health administrative data to serve as reference for algorithms derivation. Algorithms based on outpatient physician visits for CD plus endoscopy billing code were constructed and tested.

Results: The best algorithm selected based on performance from derivation study and clinical expertise consisted of an OHIP-based endoscopy billing claim followed by 1 or more adult or pediatric gastroenterologist encounters after the endoscopic procedure. The sensitivity, specificity, PPV, and NPV for the algorithm were 70.4%, >99.9%, 53.3% and >99.9% respectively.

Conclusion: Study results suggest that the currently available Ontario health administrative data is not suitable for identifying incident pediatric CD cases.

Acknowledgements

I would like to offer my gratitude to the following individuals, who provided me with great advices and support throughout my thesis:

- Dr. Eric I Benchimol, my thesis supervisor, who guided me every step along the research process. His passion for clinical research is infectious; through him I have gained a deeper appreciation for my research career, and I aspire to become a successful researcher like him one day.

- Dr. Steven Hawken, Dr. David Mack and Dr. Doug Manuel, my thesis advisory committee members. They provided insightful clinical and methodological advices that helped improved the quality of the study, and show me great support when I ran into different obstacles during the research process

- Ms. Nassim Mojaverian, my ICES analyst, who helped me create the datasets and troubleshot technical difficulties related to the analysis. Through her teaching I became more confident in research data analysis.

- My classmates from the MSc. Program, who have provided me with great support and encouragement for the past two years. I am grateful to have met you all!

- The staffs and patients at the Ottawa Environmental Health Clinic, who inspired me to become an epidemiologist contributing to evidence-based medicine.

- My family in Hong Kong, who provided me with unconditional support throughout my academic and professional career

- Special thanks to K.K., S.C., V.H., N.M. and T.D. who provided me with the greatest moral support!

This project was funded by the CHEO Research Institute ICES Research Grant. I am grateful for the financial support from the Ontario Graduate Scholarship.

Table of Contents

List of Tables

List of Figures

Abbreviations

AGA	Anti-gliadin antibodies
CD	Celiac Disease
CHEO	Children's Hospital of Eastern Ontario
CIHI	Canadian Institute of Health Information
CMA	Census Metropolitan Area
CI	Confidence Interval
CPRD	Clinical Practice Research Datalink
DAD	Discharge Abstract Database
DGP	Deamidated gliadin peptides
EMA	Anti-endomysial antibodies
ESPGHAN	European Society for Pediatric Gastroenterology, Hepatology, and Nutrition
IgA	Immunoglobulin A
IgG	Immunoglobulin G
ICES	Institute of Clinical Evaluative Science
NASPGHAN	North American Society for Pediatric Gastroenterology, Hepatology, and Nutrition
NPV	Negative Predictive Value
OHIP	Ontario Health Insurance Plan
PPV	Positive Predictive Value
PY	Person-Year
RPDB	Registered Persons Database
SDS	Same Day Surgery (Database)
tTG	tissue transglutaminase antibody

Chapter 1 : Introduction

The goals of this chapter are:

1. **Introduce the main research questions and hypothesis**
2. **Provide a background on pediatric Celiac disease and its relevance to child health**
3. **Describe the epidemiology of pediatric Celiac disease from the current literature**
4. **Introduce health administrative data and its application in health research**

1.1 General Goal and Main Hypothesis:

The main goal of the thesis was to develop a case-identifying algorithm that can accurately capture biopsy-proven celiac disease (CD) in children in Ontario, Canada from population-based health administrative databases. We intended to use the derived algorithm to determine the incidence of pediatric Celiac disease (CD) in Ontario from 1995-2011 from the health administrative databases in Ontario.

• 1.1.1 Research Questions and Hypothesis:

The major research questions and their respective hypotheses are:

1. What is/are the best algorithm(s) to identify pediatric biopsy-proven CD cases from health administrative data?

    - *Hypothesis:* Health administrative data algorithms are comprised of different combinations of procedural and diagnostic codes. By developing an algorithm that best reflects the actual health service utilization pattern of pediatric CD patients, it is possible to accurately identify pediatric biopsy-proven CD from health administrative data.

2. What is the incidence of pediatric biopsy-proven CD in Ontario from 1995 to 2011? Is the trend in diagnosis increasing over time?

    - *Hypothesis:* Due to the increased availability of disease-screening tests and the increased awareness of the condition in the past decade, we hypothesized that there will be an increase in the incidence of biopsy-proven CD cases during the study time period (1995-2011).

1.2 Background of Pediatric Celiac Disease

- **1.2.1 Introduction**

Celiac disease is an autoimmune condition characterized by enteropathy as a result of exposure and immune response to gluten, a protein commonly found in wheat, rye, and barley.(1, 2) Once considered a rare disease, CD is now regarded as one of the most common autoimmune disorders, with an estimated prevalence of 1% worldwide.(3) CD can be diagnosed at any age following exposure to gluten. In both adults and children, the classical symptoms include diarrhea, abdominal distention, abdominal pain, and nutritional deficiencies; however, the condition has a more profound impact on growth and development in children, with failure to thrive, weight loss, linear growth stunting, and developmental delay being the most common symptoms among pediatric patients.(4, 5) Due to increased efforts to screen first degree relatives and at-risk populations, a rising proportion of children with CD are asymptomatic at the time of diagnosis.(6) In addition, symptoms that were once thought to be unrelated to CD are now recognized as the non-classical presentation of the disease. These symptoms include: vomiting, constipation, and anemia.(7) With recent advances in disease screening and detection, it was noted that some patients can have histological changes in their intestinal mucosa while experiencing no symptoms at all. These asymptomatic cases are referred to as “silent cases”.(4) The new discoveries in the presentation of the disease have led to the debate

of whether the observed rising incidence of CD around the globe is the result of true increase in the number of new cases or better disease detection and screening.(8)

- **1.2.2 Celiac Disease's Relevance to Child Health**

Children with undiagnosed or untreated CD are at risk for multiple developmental health deficits, including stunted growth, delayed puberty, and impaired bone health. They are also at risk for complications in later life, such as iron-deficiency anemia, osteoporosis, infertility, malignancy, and an increased risk for the development of other autoimmune disorders.(9) The only effective treatment is life-time adherence to a gluten-free diet. With timely diagnosis, a gluten-free diet can reverse the damage in the intestinal linings and other body systems caused by the disease. It has also been documented that pediatric CD patients who suffered from growth delays experienced catch-up growth once they were started on a gluten-free diet.(10) In addition, pediatric patients with reduced bone mineral density are able to make a complete recovery in bone mineralization after one year of a gluten-free diet.(11)

In terms of malignancy, a cohort study with an average follow-up time of 25 years has documented that CD patients have 1.4 times the risk of any malignancy when compared to the general population. However, with the exception of non-Hodgkin Lymphoma, CD patient's risk for malignancy approaches the general population's after 15 years on a gluten-free diet.(12) Another cohort study reported that patient who followed a strict gluten-free diet for three years had an even lower risk for non-Hodgkin lymphoma than the general population.(13) There is growing evidence suggesting that CD patients who were diagnosed in childhood did not have a higher risk for malignancy or mortality as compared to the adult patients, potentially due to early implementation of gluten-free diet to adverse the risk for malignancy.(4, 14, 15) These findings all signify the importance of timely diagnosis of CD among children so that the complications of the disease can be reversed or prevented.

- **1.2.3 Pathogenesis of Celiac Disease**

Celiac disease is characterized by the specific change in mucosal architectural changes within the duodenum and jejunum, such as the shortening of the villi linings (villous atrophy), the increase in intestinal gland proliferation (crypt hyperplasia), and the infiltration of lymphocytes into the mucosal cell layers known as the lamina propria and the epithelium.(4) These changes results from damage caused by antibodies to gluten, a protein component of wheat, rye and barley. Gluten contains gliadins, which are a class of alcohol-soluble peptides called prolamines. They are mainly composed of the amino acids glutamine and proline.(4) The gliadin peptide itself contains very few negatively charged amino acids, which makes them less likely to interact with antigen-presenting molecules.

On the surface of the brush border and within the subepithelial compartment of the intestinal mucosa, there are intracellular enzymes called tissue transglutaminases (tTG). They catalyze the deamidation of glutamine residues on gliadin peptides into glutamic acid, which is negatively charged.(16) In CD patients, their tTG is upregulated, and as a result, more negatively charged amino acids are formed in the gliadin peptide due to the deamidation of glutamine to glutamic acid.(17) The negatively charged amino acids on gliadin have greater affinity to bind to the HLA-DQ molecules, which are glycoproteins that are expressed on the surface of antigen presenting cells.(1) The expression of the genotypes HLA-DQ2 or HLA-DQ8 is particularly known to be associated with CD.(18)

The antigen presenting cells present the gliadin peptides to the T-cells in the lamina propria and trigger their activation.(17) The activated T-cells release different types of inflammatory cytokines, which can cause direct damage to the epithelial cells; the cytokines can also trigger the releases of other enzymes and cytokines that can cause a series of immune-mediated responses that lead to injury the epithelial cell layer, which result in the histological changes in CD patients' duodenal biopsy samples.(19) With a better understanding of the pathogenesis of

the disease, some of the components in the disease-causing pathway are now targeted for disease detection(20), as described in chapter 1.2.5.

- **1.2.4 Etiology of Celiac Disease**

There is a strong genetic predisposition to CD. Concordance of 75% has been reported among monozygotic twins.(21) The prevalence of CD among patients' first degree relative is up to 20%, significantly greater than the 1-2% prevalence reported among the general population.(22) Some of the genes associated with CD have been identified, with the most well-known example being the HLA-DQ genes located on chromosome 6.(18) The HLA-DQ genes code for glycoproteins that are involved in immunological processes, and the HLA-DQ2 protein are expressed in 90-95% of CD patients, while the HLA-DQ8 are expressed in the remaining patients.(23) However, 20-40% of the general population is carrying these genotypes without developing the disease.(24) Therefore, the expression of these genes alone is necessary but not sufficient for the development of CD.

CD is also associated with other conditions. Patients with type 1 diabetes mellitus, Down syndrome, or other autoimmune disorders have increased risk of CD. The prevalence of CD among type-1 diabetes and Down syndrome patients are estimated to be 3 -6% and 6-8% respectively.(22, 25) This high risk has led to clinical recommendations for the routine screening of CD among these at-risk populations.

Another etiological factor for CD that has gained wide attention in recent years is infant feeding practice. A study from Sweden, which tracked the incidence of CD among children under the age of 15 since 1973, documented a dramatic increase in CD incidence from 1985-1994 among children less than 2 years of age. This was followed by a substantial decline in the number of new CD cases from 1995-1997, with almost no cases of CD being diagnosed within the first year of life.(26, 27) The decline in cases among children under 2 years of age coincided with the introduction of a new national infant feeding recommendation that encouraged parents to

introduce gluten in smaller amounts starting from 4 months of age, and to continue breast-feeding during introduction of gluten. This observation sparked great interest in the roles of gluten introduction and breastfeeding in the development of CD. A recent meta-analysis of 5 case-control studies suggested that breastfeeding at the time of gluten introduction decreased the risk of CD when compared to formula feeding. In addition, introducing gluten at ≤4 months or ≥7 months was found to be associated with an increased risk for CD in observational studies.(28) However, a recent placebo-controlled randomized trial, which involved administrating gluten protein to infants at 4 months of age for 8 weeks continuously compared to placebo , reported that early introduction of gluten did not affect the risk of CD by 3 years of age.(29) In addition, the study also reported that breastfeeding practices did not influence the risk of CD.(29) The trial refuted the association between CD development and the timing of gluten introduction or breastfeeding practices seen in observational studies.

- **1.2.5 Diagnosis of Celiac Disease**

The diagnosis of CD typically consists of serological screening test followed by, and if positive, confirmation with intestinal biopsy. The four main types of serologic tests are: antibodies to tissue transglutaminase (tTG), anti-endomysial antibodies (EMA), anti-gliadin antibodies (AGA), and deamidated gliadin peptides (DGP) antibodies.(30) While both immunoglobulin A (IgA) and immunoglobulin G (IgG) based test are available for each serum antibody, IgA is the standard antibody measurement for suspected CD patients who are not IgA deficient.(31) The IgA-based tTG and the EMA are the most commonly used serologic tests, as they are highly sensitive and specific. For children, the IgA-tTG has a reported sensitivity and specificity that ranged from 90-100% and 94-100% respectively, while the IgA EMA has a sensitivity of 88-100% and a specificity of 90-100%.(32) However, the accuracy of the EMA is laboratory-dependent, and the test requires primate esophageal tissue. The IgA-tTG is much cheaper to perform, and it is an automated test that requires little involvement from a test operator, which makes its result more

reliable.(31, 33) Therefore, modern clinical practice guidelines suggest IgA-tTG as the primary screening test for clinical use.

The downside for all these IgA-based serologic tests is that they are susceptible to misclassification of patients as false negatives in the setting of IgA deficiency. IgA deficiency is 10-16 times more common among symptomatic CD patients than in the general population, and its prevalence among CD patients is estimated to be at 1.7-2.6%.(34, 35) Therefore, a patient's serum IgA titre should be measured when an IgA-based serologic test is used. When a patient suspected of having CD suffers from IgA deficiency, an IgG based serologic test can be used instead.(31) In that case, the IgG-based DGP is recommended for CD screening.

In recent years, with the identification of genes associated with CD, it is becoming possible to screen for the presence of these CD-affiliated genes to identify susceptible CD patients. As mentioned above, almost all of the CD patients either express the HLA-DQ2 or HLA-DQ8 proteins, with HLA-DQ2 being expressed in 90-95% of the CD patients. The HLA-DQ2 and HLA-DQ8 are encoded by the HLA-DQ genes located on chromosome 6; therefore, genetic screening involves screening for genetic variants of the HLA-DQ genes that will lead to the expression of the HLA-DQ2 or HLA-DQ8 protein.(18) The presence of the HLA-DQ2 and/or HLA-DQ8 genotypes are necessary but not sufficient for the development of the disease, as the genotypes are also carried by 20-40% of the population. Regardless, the genetic test can still be of value for ruling out the disease, as absence of these genetic markers would almost certainly rule out CD. Therefore, some of the newer clinical guidelines have suggested using HLA testing as the first line of test if it is available to the patients.(36) However, the test is not currently covered under the Ontario Health Insurance Plan (OHIP), and costs approximately $300, hindering its use for CD screening.

The North American Society for Pediatric Gastroenterology, Hepatology, and Nutrition (NASPGHAN) and other pediatric gastroenterology societies regard duodenal biopsy as the

gold standard for diagnosing CD.(1, 2, 36) The biopsy is usually performed to confirm the CD diagnosis after a patient has a positive serological screening test. The biopsy samples are obtained through an endoscopy during esophagogastroduodenoscopy. Due to the patchiness of mucosal involvement, multiple samples are recommended from the first and second parts of the duodenum.(1, 36) In order for the results to be accurate, the patient must maintain a gluten-containing diet for at least 3-6 months prior to biopsy. (1)

The most recognized biopsy evaluation criteria were put forth by Marsh (37), who graded the histological changes in the duodenal samples to describe the severity of the disease. The grading system is shown in Table 1.1 below. Each grade is associated with a specific set of histological changes. For example, a Marsh I grading is defined by normal mucosal and villous structure, with an increased number of intraepithelial lymphocytes. The Marsh IIIC, which represents the most severe form of the disease, is signified by totally villous atrophy with severe crypt hyperplasia, as well as lesions generated from the infiltration of the inflammatory cells into the mucosal cell layers.

Table 1-1 Marsh's histological classification for celiac disease

Marsh Classification	Increased intraepithelial lymphocytes	Crypt hyperplasia	Villous Atrophy
I	Yes	No	No
II	Yes	Yes	No
IIIA	Yes	Yes	Yes (partial)
IIIB	Yes	Yes	Yes (subtotal)
IIIC	Yes	Yes	Yes (total)

While a biopsy is strongly recommended by the NASPGHAN clinical guidelines for the confirmation of CD diagnosis, it is not always performed in clinical practice. A national study from the United States documented that only 11% of suspected CD patients underwent biopsy.(38) Another American study reported a biopsy rate of 39% in those with abnormal serology.(39) The procedure is inherently invasive, which can deter patients from agreeing to it. In addition, with recent advances in serologic testing, the question of whether CD can be diagnosed with non-invasive tests alone has been raised.(40)

In the recently updated CD diagnostic guideline by the European Society for Pediatric Gastroenterology, Hepatology, and Nutrition (ESPGHAN), recommendations stated if the patient shows symptoms or signs suggestive of celiac disease and the tTG titre is 10 times greater than the upper limit of normal, the pediatric gastroenterologist may consider performing additional laboratory tests (e.g. EMA or HLA testing) instead of biopsy.(36) However, several studies have cautioned about the risk of relying on serologic test results for CD diagnosis. While the IgA tTG and EMA tests were reported to have high sensitivity and high specificity, these parameters are not always constant across all histological manifestation of CD. For instance, the sensitivity of IgA tTG and EMA for partial villous atrophy (Marsh IIIb) can range from 30% to 89%, while the sensitivity for Marsh II grade lesion was less than 50%. (41) In addition, a study that examined the correlation between serologic test results and histology of the tissues from biopsy documented cases where patients with negative EMA and tTG results exhibited Marsh 3A or above histological changes from their biopsies.(42) Furthermore, while the specificity of the IgA tTG and EMA are in the 90% range, the low prevalence of CD among children (<1%) can result in a low positive predictive value (PPV) for the test, which increases the likelihood for false positive diagnosis.(31)

Diagnosing CD based on symptoms or serology alone can place patients at risk of being misdiagnosed, with its lifestyle, financial, and health implications. Children who are incompletely

investigated may be at risk for non-adherence to a strict gluten-free diet and therefore at risk of growth impairment, osteoporosis, infertility, and cancer. In addition, they can also increase the financial burden of the current health system, as they are being treated as having CD instead of their underlying health issue, which can result in unnecessary clinical visits and laboratory tests. Therefore, from the point of view of both the patient and the health system, it is crucial for children who are suspected of CD to be definitively diagnosed through biopsy.

- **1.2.6 Epidemiology of Pediatric Celiac Disease**

CD is estimated to have a prevalence of 1-3% in the general population, while the pediatric prevalence is estimated to be 0.3 to 1%.(7, 20, 43, 44) In recent years, many studies have been conducted in an attempt to examine the distribution of the disease among the pediatric populations in different part of the world, but primarily in Western countries.

According to recent studies, the incidence for pediatric CD is increasing. In Southeast Scotland, a retrospective cohort study measured the incidence of CD among children less than 16 years of age from 1990-2009. The incidence increased 6.4-fold from 1.8 to 11.7 per 100,000 person-years (PY) from 1990-1994 to 2005-2009. In Cardiff and the Vale of Glamorgan in UK, the pediatric incidence of CD was reported at 6.89 per100,000 PY during 2001-2005, which was comparable to the rate reported in Scotland from 2000-2004 (5.89 per 100,000 PY).(45)

In Estonia, a study examined the number of new CD cases in children up to 19 years of age from 1976 to 2010 through both retrospective (1976-1989) and prospective (1990-2010) data collection. They reported an 8.8-fold increase in Incidence rate from 0.21 per 100,000 PY in the retrospective period to 1.85 per 100,000 PY in the prospective period.(43) However, the observed increase in new cases could be overestimated in the study, as some cases may not have been captured through the retrospective data collection process.

In Sweden, a registry that tracked the incidence of CD in children below 15 years was created in 1973. Over the 36 years of follow-up time, they observed a 4% annual increase in CD

incidence. The incidence rate increased from 10 per 100,000 PY to 33 per 100,000 PY for 1973-1984 and 1985 to 1994 respectively. The incidence rate decreased to 15 per 100,000PY temporarily in 1995-1997, only to increase again to 31 and 42 per 100,000 PY in 1998-2003 and 2004-2009 respectively.(46) The most prominent finding was the dramatic change in incidence between 1985-1994 and 1995-1997 among the children under the age of two years. In 1985-1987, the incidence for children under two was 200-240 per 100,000 PY. However, by 1995 the incidence dropped abruptly to 51 per 100,000 PY.(26) The rapid decline in CD new cases in 1995 coincided with the release of a new infant feeding guideline that recommended introducing gluten at 4 months of age with continued breastfeeding, leading to the hypothesis that CD is associated with infant feeding practices.

Spain had the highest pediatric incidence rate reported in Europe, with a rate of 54 per 100,000 PY in 2006-2007 amongst children under the age of 15.(3)

In North America, the only recent epidemiology study of pediatric CD comes from Alberta, Canada, which described an incidence of 2 per 100,000 PY in children <18 from 1990 to 1997. However, after the introduction of IgA-EMA in late 1990s, the incidence increased by 3.7-fold to 7.3 per 100,000 PY in 2000-2006. The study raised the question of whether the increased incidence amongst children was the result of better disease detection or increased awareness.

In terms of prevalence, a screening study in Sweden among 12 years old that were born during 1984 to 1996 reported a prevalence of 3%.(47) Another study from UK that used the Clinical Practice Research Datalink (CPRD) to capture patients with a CD diagnostic code estimated a point prevalence of 0.03% for children under 5, and 0.13% for children under 17 in June 2011.(48) In Denmark, a study that involved capturing the biopsy-confirmed CD cases through the country's National Registry of Pathology reported an increase in prevalence from 0.04% in 2000 to 0.08% in 2010 among children under 17 years old.(20) In Norway, a study that involved

using the Norwegian Patient Register, which is a health administrative database that documented all activities from hospitals and outpatient clinics, reported a prevalence of 0.38% in children under the age of 12 from 1999 to 2011.(25) Compared with the 1% prevalence rate as estimated for the general population, these study findings seem to suggest that CD is less prevalent among children than in adults. But this is an expected finding, as CD is typically a non-life threatening chronic condition that can be diagnosed at any age, thus the prevalence of the disease will gradually increase with age.

In summary, CD incidence has increased in children, and recent studies reporting incidence ranging from 2 to 54 cases per 100,000 PY. The characteristics of pediatric patients in the past decade have also changed significantly. Across all of the incidence studies, an increase in the age of diagnosis has been observed. The mean age of diagnosis for CD used to be under the age of two; in the past decade the average diagnosis age has risen to between 4 and 7 years.(5, 43, 46) In addition, the proportions of patients with the non-classical or silent presentation of the disease are increasing, and it has been estimated that 25% of the patients that are currently diagnosed with CD are asymptomatic.(7) These changes in patient characteristics are likely due to better disease detection through serologic testing and the increased awareness of recommendations to test the at-risk populations, such as the first relatives of CD patients and the patients with conditions that are associated with CD (e.g. diabetes, Down syndrome).(33) For example, in Alberta, the age of diagnosis has increased after serologic tests were implemented, and 1 in 4 children that were diagnosed with CD were screened because he/she had conditions associated with CD.(5) Serologic tests allowed clinicians to identify patients with mild or absent symptoms. While the benefit of treating asymptomatic CD is still the subject of debate, it has been reported that patients who had asymptomatic CD experienced improvements in symptoms that were not recognized before

diagnosis; their vitamin and minerals absorption and bone mineral density also improved significantly after starting a gluten-free diet.(49, 50)

- **1.2.7 Using Health Administrative Data Research to Study Chronic Disease**

Health administrative data are defined as “information passively collected, often by government and health care providers, for the purpose of managing the health care of patients”.(51) In Canada, examples of health administrative data include the Ontario Health Insurance Plan (OHIP) physician billings database and the Discharge Abstract Database (DAD) from the Canadian Institute for Health Information (CIHI). There are several qualities that make health administrative data attractive for research. As they are routinely collected as part of the health care delivery or administrative process, research using these data is less costly than conducting primary data collection. In addition, they usually have wide population coverage, which can potentially increase the representativeness of study patients and generalizability of study findings. In Canada, provincially-collected health administrative data are population-based, comprising all legal residents of the province with a valid health card. Health administrative data have been used in a large variety of epidemiological research; some of these applications include: estimating the incidence and prevalence of disease in a population, measuring the health outcomes of patients with certain exposures or conditions, longitudinal surveillance of a disease cohort, and health systems and services research.(52-56)

While health administrative data offer many great advantages for epidemiological studies, they are not created for research. Therefore, in order to determine the appropriateness of a database for research, one has to be aware of how the health administrative database is created in order to infer the qualities and the limitations of the data. For example, physician billing data comprises the reported services provided to the patient during an encounter. The inputted data could be collected and reported by physicians themselves, their administrative staff, or a billing agent. Inaccurate reporting, up-coding for improved reimbursement, data entry errors, or other

problems will result in data inaccuracy or incompleteness, which can introduce sources of bias to the results generated from the data.

Using administrative data to identity patients with chronic diseases also poses unique challenges. Very often, more than one clinical visit is required for a patient to be diagnosed with a chronic disease. This feature should be reflected in how one ascertains cases from a health administrative database. Rather than searching for patients that have any clinical encounters related to the disease of interest in the health administrative database, a case-identifying algorithm that represents multiple healthcare contacts for the condition of interest maybe more accurate in classifying a patient's disease status. It is also important to evaluate the performance of such algorithms in identifying the disease at a given population, as the accuracy of the algorithm can vary depending on the characteristics of the study population. For example, it has been documented that the performance of an algorithm can vary greatly across different age groups.(52, 54)

The most common way of evaluating the performance of an algorithm is to create an external reference cohort in which true disease status is known, and to apply the algorithm to the health administrative base to verify if the patients' disease status identified from the health administrative data matches the reference cohort. The sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) are calculated for the algorithms and the one with the best performance can then be applied for case-ascertainment. The use of validated algorithms can minimize the risk for misclassification bias, which can improve the internal validity and the accuracy of the study conclusion.(57)

Despite the fact that health administrative data are used in thousands of studies published every year, relatively few of these studies evaluated the performance of codes and algorithms that were used. Prior to June 2009, only 271 health administrative data studies were reported to

include a data validation process. The reporting qualities of these validation studies varied greatly, with only 36.9% of them reporting four or more markers of diagnostic accuracy (e.g. sensitivity, specificity, PPV, NPV, likelihood ratios, c-statistics, etc.).(58) For the pediatric population, only 37 studies evaluated the performances of the diagnostic codes or algorithms that were used in the study.(59) These figures are alarming as they suggest that a significant proportion of existing health administrative data research could be vulnerable to misclassification bias, could have drawn erroneous conclusions, and may have resulted in poor decision-making in health practice and policy.

In the current study, Ontario health administrative data housed at the Institute of Clinical Evaluative Science (ICES) was used to identify the cases of CD among pediatric patients aged 6 months to 14 years old in Ontario. To derive the algorithm for identifying CD cases in the target population, both true positive and true negatives cohorts were created using the hospital records from the Children's Hospital of Eastern Ontario (CHEO) and non-CD residents of the city of Ottawa respectively. We hypothesized that CD could be accurately identified based on an algorithm of procedures and codes in Ontario health administrative data, and a pediatric CD cohort created to conduct longitudinal surveillance of the disease.

Chapter 2 : Study Design

The goals of this chapter are:

- **Reviewing the study objectives**
- **Introducing the health administrative databases hosted in the Institute for Clinical Evaluative Science (ICES) as a source of research data**
- **Describe the creation of the true positive and true negative reference cohorts for case-identifying algorithm derivation**
- **Explaining the process of case-identifying algorithms development and testing**
- **Describe the application of case-identifying algorithm in tabulating the incidence rate of biopsy-proven CD in Ontario**

2.1 Study Objectives

The current study had 2 main objectives:

1. To develop an algorithm to accurately identify cases of biopsy-proven CD in pediatric patients aged 6 months-14 years old from Ontario health administrative data housed in ICES

2. To estimate the incidence of biopsy-proven CD among pediatric patients in Ontario using the derived case-identifying algorithm

In this chapter, the methods that were used to complete each study objective will be described in detail.

2.2 Ontario Health Administrative Data at the Institute for Clinical Evaluative Science

The Institute for Clinical Evaluative Science (ICES) was founded in April 1992, in order to improve the health of Ontario residents by conducting population-based health services research using health administrative data collected routinely when providing health care service.(60) With health services records of almost 13 million people, ICES is one of the world's

largest sources of individual-level health data. The patients' records are anonymized to protect privacy, and health records are linkable across databases using the ICES unique identification number (IKN).The IKN is an encrypted and scrambled identifier based on Ontario health card number. These qualities made ICES data an ideal data source for conducting epidemiological research.

ICES hosts a large variety of health administrative data, such as physician billing claims, prescription claims, hospital discharge records, and emergency and ambulatory care visits records. It also houses data from Canadian census, disease registries and various health surveys. With linkage across databases possible, researchers can to study the associations between different population characteristics and clinical health outcomes.

To conduct this derivation study and identify biopsy-proven CD among children aged 6 months to 14 years old, the following databases housed in ICES were used (61):

1. Ontario Health Insurance Health Plan (OHIP)

   The OHIP database included most claims billed by the primary care physicians and specialists in the Province of Ontario. For the physicians who are not paid through fee-for-service, they are required to perform shadow billings for the services that they provided. In the dataset, each row represents a single service that is provided by the physician for a patient. The major variables include the patient and physician identifiers, the diagnostic code, and the fee-code for the service (e.g. outpatient visit, laboratory testing, procedure etc.) that was provided.

2. CIHI Discharge Abstract Data (DAD)

   The Ontario portion of the CIHI DAD database documents all acute hospitalizations in Canada.(62) Each row of the data represents a single admission from a patient, and it is first created by the professional medical coders at the hospitals based on the information from the patient's chart. The coders would complete a discharge abstract based on the

instruction from the DAD Abstracting Manual and the hospitals would forward the abstracts to the Canadian Institute for Health Information (CIHI) on a monthly basis. The major variables include the patient identifier, demographic data (e.g. gender, date of birth, postal code etc.), admission and discharge date, and diagnostic codes documenting the main and secondary diagnoses of the patient.

3. Same Day Surgery Database (SDS)

   The SDS database documents the same-day surgery or procedure stay from a patient in each of its data row. It is derived from the National Ambulatory Care Reporting System (NACRS), which collects information on patients' hospitalizations and visits to community based ambulatory care. The key variables include patient identifiers, demographic information, diagnostic codes and procedure codes for the services that were provided during the patient's visit.

4. Registered Persons Database (RPDB)

   The RPDB contains information on the demographics, residency and the OHIP eligibility by year for individuals who have ever received an OHIP card. The data was collected and provided by the Ministry of Health directly. Each data row represents an individual who has ever held an OHIP card.

5. 2006 Ontario Census Area Profile

   The 2006 census questionnaire was estimated to contain the basic demographic information (Age, gender and marital status) on 100% of the Canadian population in 2006. It will be used as the standard population to standardize the CD incidence rates tabulated in the study to the age and sex structure of the Canadian population.

2.3 Creation of True Positive and True Negative Reference Cohorts

Traditionally, case-identifying algorithms utilize true-positive and true-negative reference standard populations to evaluate their performances in discriminating between true positive and true negative cases. The creation of the reference standard is most frequently done using chart

review to ensure accurate classification of cases.(58) While the process is time-consuming, it is essential that studies using administrative data use accurate case and control definitions to minimize misclassification bias.(57)

- **2.3.1 Construction of the True-Positive Reference Cohort for Algorithm Derivation**

Since CHEO is the sole pediatric health care centre and the only pediatric endoscopy centre in the city of Ottawa, all of the pediatric patients who underwent a biopsy for CD would have undergone endoscopy at CHEO. Therefore, the cases identified at CHEO were used as the true positive reference standard for the derivation of the biopsy-proven CD algorithm in Ottawa.

The inclusion criteria for the true positive cases identification include: CD patients aged 6 months to 14 years old, who were residing in the Census Metropolitan Area (CMA) of Ottawa during 2005-2011. They must have had a valid OHIP number for data linkage and a complete medical chart for review. They must also have continuous OHIP eligibility from 2005 to 2011 as determined using the RPDP; this is to ensure that all of their medical services related to CD would be captured by the OHIP billing database during 2005-2011 to minimize the potential for missing data resulting in errors during the algorithm derivation process.

Three different strategies were used to ensure that all of the biopsy-proven CD cases at CHEO are captured. The first method involved using the ICD-9 code for CD (579.0) to search for the potential cases within the electronic record and the shadow-billing databases at CHEO. Physicians at CHEO are predominantly paid a salary under an alternative funding plan; however, they must submit “shadow-billing” records for administrative purposes. Secondly, an electronic search was performed at CHEO’s pathology database. CD cases diagnosed during the study period were identified from the database with the Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) for celiac disease. Potential cases were identified using the SNOMED CT code for duodenal biopsy associated with CD (“T64300 D6218”). Thirdly,

since newly diagnosed CD patients at CHEO were always sent for dietary counselling regarding a gluten-free diet with a CHEO-based dietician, we requested dietician clinic lists for CD patients seen during 2005-2011. These three methods combined to ensure completeness of the list of all patients with suspected biopsy-proven CD diagnosed at CHEO and in Ottawa from 2005-2011.

Once all of the potential true cases were identified using the three methods described above, their medical charts were reviewed to confirm celiac disease diagnosis. The designated reviewers (Mr. Jason Chan and Dr. Eric Benchimol) determined whether patients had CD based on diagnostic guidelines developed by the North American Society for Pediatric Gastroenterology, Hepatology and Nutrition (NASPGHAN).(1) True biopsy-proven cases were defined as those whose histology findings received a Marsh classification of II or above.(40) Those who received a Marsh I classification with abnormal levels of tTG antibodies were considered as “suspected celiac disease” cases and they were excluded from the true positive and true negative reference cohorts. Duplicate chart reviews by both reviewers were conducted in 10% of charts; the Kappa agreement statistic between the two reviewers was calculated. The standardized data collection form used for the chart review is attached in Appendix I. Table 2.1 summarized the inclusion and exclusion criteria for the construction of the Ottawa true positive reference cohort:

Table 2-1 Inclusion and exclusion criteria for the construction of the true positive reference cohort for pediatric celiac disease at CHEO

Inclusion	Exclusion
CD diagnosed at age 6 months-14 years old	Diagnosed at age<6 months or age>=15 years old
CD diagnosed between 2005-2011, with duodenal histology meeting Marsh II criteria and above	Positive tTG with duodenal histology meeting Marsh I criteria or below (called “suspected CD”)
Resided in the Ottawa Metropolitan Area between 2005-2011	Resided outside of the Ottawa Metropolitan Area between 2005-2011
Continuous OHIP eligibility from 2005-2011	No valid health card or no continuous OHIP eligibility from 2005-2011

Once all true positive biopsy-proven CD cases were confirmed with chart review, the data extracted from the CHEO health records were encrypted and transferred to ICES for linkage to the selected health administrative databases. Similarly, those confirmed as not having CD were transferred to ICES for linkage and inclusion in the true negative reference standard cohort.

- **2.3.2 Construction of the True-Negative Reference Cohort for Algorithm Derivation**

The true-negative reference cohort consisted of the following groups of patients:

a) All patients determined not to have biopsy-proven CD based on the above chart review at CHEO and living in Ottawa from 2005-2011.

b) All other children 6 months to 14 years living in the CMA of Ottawa from 2005-2011 (who were not seen at CHEO or diagnosed with CD based on chart review, and therefore presumed to not have biopsy-proven CD), identified from Ontario health administrative data.

Those included in the true-negative population also met with the same qualification criteria detailed in Table 2.1. The strategy of including all disease negative patients within a given jurisdiction allowed for a similar CD prevalence representative of the population, and therefore

accurate determination of negative and positive predictive values (57). This strategy has been used in previous algorithm derivation and validation studies in Ontario.(53, 54)

2.4 Algorithm Derivation for Biopsy-Proven CD

As mentioned in the introduction, it usually takes more than one clinic encounter for a chronic condition to be diagnosed. If one were to ascertain disease cases from the health administrative databases based on the presence of a single health claim with the diagnostic code for the condition of interest, it can generate many false positive cases, leading to an overestimation of the number of disease cases. For example, if a patient was investigated for the condition of interest and proven not to have the disease, he/ she would still have health claims related to the condition of interest in the administrative data. Similarly, mistakes in service billing can occur where a health claim can be falsely billed with the diagnostic code for the condition of interest. If the case definition “the presence of a single health claim for the condition of interest” is used to ascertain disease cases from the health administrative databases, the two scenarios above would misclassify the patients as true cases when they did not have the disease.

To accurately identify disease cases from health administrative databases, the case definition should reflect the multiple-encounters nature of the disease investigation and follow up processes. A case-identifying algorithm consists of a combination of health care encounters with different diagnostic and procedural codes attached to them, and is a better representation of the service utilization pattern for the patients who truly have the disease. For a patient to be diagnosed with biopsy-proven CD, they should have procedural claims documenting the esophagogastroduodenoscopy (upper endoscopy) in addition to the health claims related to the disease investigation and follow up process. Therefore, the biopsy-proven CD case-identifying algorithm was planned to consist of 2 major components: (1) The presence of a health claim for an endoscopic procedure and (2) the numbers of other health claims related to CD within a given time interval. By varying the definition for each of the algorithm components different

algorithms were generated and tested (e.g. the presence of an OHIP claim with the fee code for endoscopy plus 1/2/3/4 inpatient/outpatient health claims for CD in 1/2/3/4 years).

The health claims related to the endoscopic procedures were identified from the OHIP billing database using the procedural codes for esophagogastroduodenoscopy. Alternatively, since the endoscopy for CD investigation was commonly a day-surgery procedure, a majority of the endoscopic procedures were recorded in the Same Day Surgery (SDS) database as well. Therefore, the endoscopy component of the algorithms could be varied and was tested based on the source of the endoscopy health claims. Currently, there are no studies that reported on the accuracy of the procedural codes for endoscopy. However, this should not pose major problems in the algorithms' validity, as the two components of the algorithms would be tested as a whole against the reference cohort. Therefore, even if the endoscopy billing codes were inaccurate, it would be reflected in the performances of the algorithms.

The second component of the algorithm is the numbers of health claims related to CD. These health claims could represent clinical visits in outpatient settings or hospitalizations due to CD. The CD-related outpatient encounters could be identified from the OHIP billing database by searching for claims that contains the abbreviated ICD-9 diagnostic code for CD (579). In addition, we further narrowed the claims to the ones that were billed by physicians with specific specialties. For example, it is likely that patients with CD would have clinical encounters with a gastroenterologist during the disease investigation or follow up process; therefore, we tested algorithms with only gastroenterologist-related outpatient encounters. While it is less common for children and youth to be admitted to the hospital for CD, the hospitalization related to CD was identified from the Discharge Abstract database (DAD) by searching for hospitalizations that contained the ICD-10 diagnostic code for CD (K900.0). Therefore, the second component of the algorithm also varied based on the number of CD related outpatient encounter as captured in the OHIP billing database AND/OR the number of hospitalization captured by the

DAD. The procedural and diagnostic codes used for the algorithms derivation are listed in table 2.2 below:

Table 2-2 Health administrative data codes used for CD algorithms derivation

Database	Variable	Code(s)
Component 1: Endoscopy procedural code		
OHIP	feecode	Z399, Z400, Z527, Z528, Z547, Z558, Z560, Z561, Z749, Z584
SDS	prcode1-20	0115, 0116, 5791, 5792, 5795
Component 2: inpatient/outpatient encounters related to CD		
OHIP	dxcode	579
DAD	dxcode 1-16	5790
	dx10code1-25	K900

The performance of the algorithms as case definition to ascertain CD cases within Ottawa health administrative data was tested against the Ottawa true positive and negative reference cohorts. The algorithms' abilities to correctly identify the true cases and non-cases were evaluated by calculating their sensitivity, specificity, PPV, and NPV. 95% confidence intervals (CI) for the algorithms' operating characteristics were calculated according to the efficient-score method corrected for continuity.(63)

The accuracy of the algorithms was also tested in different age group in order to account for any differences in health service utilization related to CD among children in different age groups.

We planned to accept the algorithm with the highest PPV while maintaining a good sensitivity

(optimally >80%) and to use this algorithm to identify all pediatric patients with biopsy-proven CD in Ontario during 1995-2011.

2.5 Estimating the Incidence of Biopsy-Proven CD in Ontario

We applied the selected algorithms with the best performance to the health administrative data to determine the incidence of pediatric CD from 1995 to 2011. To distinguish incident from the prevalent cases, a lookback period of 3 years was implemented for all of the potential CD cases captured by the selected algorithms. Cases captured by the algorithms accompanied by the presence of any CD-related administrative codes within the 3 year lookback period were classified as prevalent cases. This lookback period was based on clinical expertise, and previously validated pediatric inflammatory bowel disease algorithms.(64) Age- and sex-standardized incidence per 100,000 children were calculated for each year from 1995-2011 for the biopsy-proven CD cases. 95% CIs were determined based on the gamma distribution.(65) Trends in incidence over time were determined using Poisson regression, controlling for age and sex. Due to data privacy measures within ICES, only the birth year of the children were shown in the Ontario health administrative data. As a result, the age of the children was calculated and rounded down to the nearest integer such that an age of zero would represent children aged 6 months to less than 1 year old. Age and the year of diagnosis were entered as continuous variables into the models. The year of diagnosis was scaled relative to the year 1995 such that the intercepts of the models would be interpretable, representing the incidence of CD in 1995 for female patients aged 6 months to 1 years old. Over-dispersion of data was tested by evaluating the Pearson Chi-Square and the Deviance of the Poisson regression models. To adjust for over-dispersion (should the problem exists), a dispersion parameter would be introduced and modelled as a ratio of the Pearson Chi-square to its associated degree of freedom. Potential interactions between the 3 independent variables were examined. To test the linearity assumption between year and the incidence of CD, quadratic terms for year were added to the model to examine if they would contribute to the fitness of the model significantly.

Chapter 3 : Results

The goals of this chapter are:

- **review the results from true positive and true negative reference cohort creation**
- **discuss the findings from algorithms derivation**
- **examine the trends in pediatric CD diagnosis in Ontario using selected algorithms**

3.1 Results from true positive and true negative reference cohort creation

- **3.1.1 True Positive Reference Cohort Creation**

316 potential cases of pediatric CD cases between 2005 and 2011 were identified, of which 116 charts were excluded (see Figure 3.1). The remaining 200 patients' medical charts were fully reviewed and 125 true positive cases of incident celiac disease were identified. A total of 20 charts were reviewed by both the main chart reviewer (JC) and a gastroenterology clinician (EIB). The two reviewers achieved 100% agreement in the diagnosis of CD, and a weighted kappa of 87.4% (95% CI: 75.4%-99.4%) for Marsh classification of the tissue biopsy.

The true positive cases were linked via OHIP number to the Registered Persons Database (RPDB) at ICES to check for their OHIP eligibility, and 10 patients were further excluded as they did not have continuous OHIP eligibility during the study period. The final true positive reference cohort for algorithm derivation contained 115 patients.

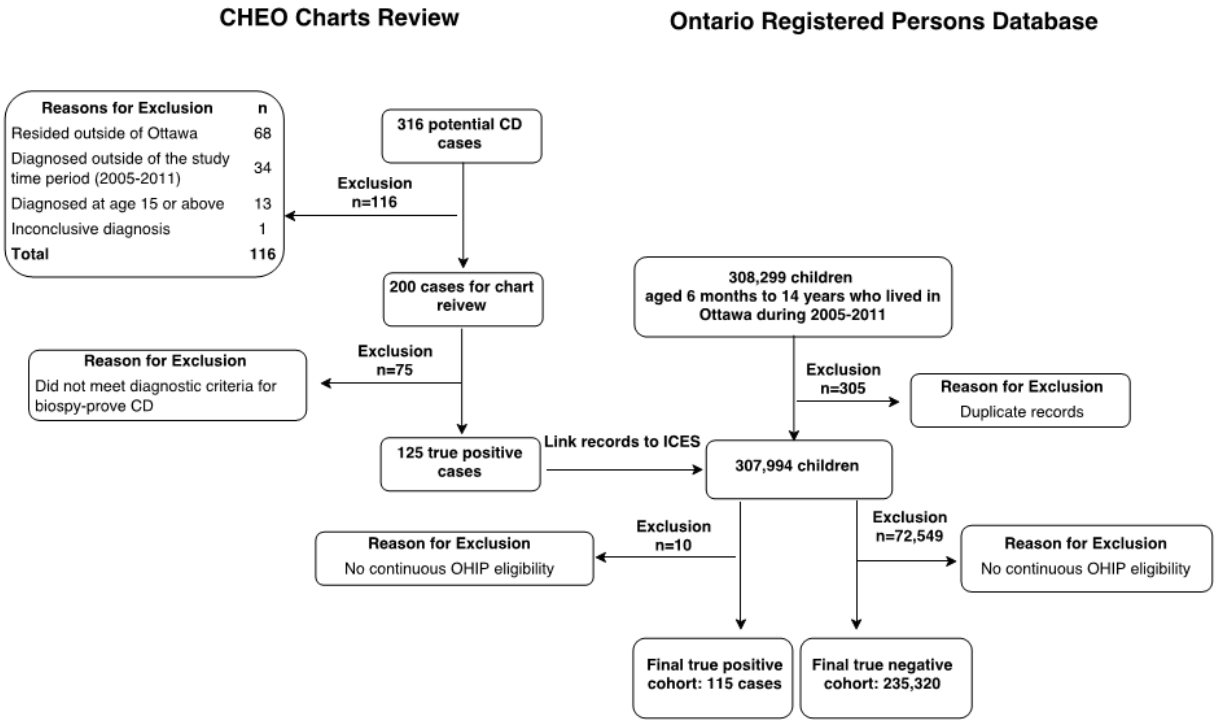

Figure3.1 Flow diagram for the creation of true positive and true negative reference cohort

- **3.1.2 True Negative Reference Cohort Creation**

The true negative reference cohort consisted of children who were determined not to have CD from the retrospective chart review and all other children aged 6 months to 14 years old who were residing in the Census Metropolitan Area of Ottawa 2005-2011. Based on the inclusion and exclusion criteria, 235,320 children were identified from the RPDB (see Figure 3.1). These children made up the true negative reference cohort.

- **3.1.3 Patient Demographics and CD Incidence Estimates in Ottawa**

From 2005-2011, there were 115 incident cases of CD in Ottawa, resulting in a crude incidence of 8.7 cases (95% CI: 7.26 to 10.46) per 100,000 PY. 62 cases were female (54%), and the median age of diagnosis was 7.4 years (IQR: 4.6-11.1years). The majority of the CD patients' duodenal biopsy were classified as Marsh 3b (70.4%), followed by Marsh 3a (15.7%) and Marsh 3c (13.9%). The annual standardized incidence rates are presented in Table 3.1.

Table 3-1 Yearly standardized incidence rate of biopsy-prove celiac disease among children in Ottawa (2005-2011)

Year	Cases (total=115)	Standardized rate (per 100,000 persons)[1]
2005	17	10.12 (5.83-16.34)
2006	11	6.54 (3.20-11.85)
2007	10	6.58 (3.15-12.12)
2008	12	7.59 (3.86-13.42)
2009	25	15.46 (9.89-23.02)
2010	16	9.06 (5.03-15.04)
2011	24	15.29 (9.67-22.97)

[1]*The annual incidence is standardized to the age and gender structure of the Canadian population in 2006.*

3.2 Results from Algorithm Derivation

- **3.2.1 Algorithms Development**

To illustrate the problems with using only a single diagnostic code in ascertaining disease cases from health administrative data, all children with a single diagnostic code for CD (579.0) were identified using the OHIP billing data from the Ottawa cohort, and the results were compared to the reference cohort created. The CD diagnostic code had a sensitivity of 88.7% (95% CI: 81.1% to 93.6%), specificity of 99.6% (95% CI: 99.5 to 99.6%), PPV of 8.9% (95% CI: 7.3 to 10.7%) and NPV >99.9% (95% CI: >99.9 to >99.9%) (see Table 3.2). The CD diagnostic code captured 102 of the 115 true positive cases. However, the case definition also classified 1045 false positive cases, meaning that 1045 children in Ottawa had at least 1 OHIP claim related to CD but were not shown to have CD based on true positive reference standard cohort. Therefore, the use of a single CD diagnostic code as the case definition could overestimate the number of CD cases by a factor of ten.

Table 3-2 2x2 cross tabulation demonstrating the accuracy of a single CD diagnostic code to identify biopsy-proven celiac disease

Presence of CD diagnostic code (579)	Biopsy-Proven Celiac Disease		Total
	Yes	No	
Yes	102	1045	1147
No	13	234,275	234,288
Total	115	235,320	235,435

Based on the pre-determined algorithm structure explained in Section 2.4, approximately 200 algorithms were tested in the current study. The candidate algorithms can be roughly divided into 3 groups based on how the health claims for the endoscopic procedures and the clinical encounters (outpatient visits and/or hospitalizations) were defined. The major characteristics of the 3 groups of algorithms are described in Table 3.3.

Table 3-3 Characteristics of the 3 groups of biopsy-proven CD algorithms tested

Components	Characteristics	Algorithm Groups		
		Group 1	Group 2	Group 3
Endoscopy	Source of endoscopy code	OHIP	OHIP	CIHI
	Procedure billing by a gastroenterologist only	No	Yes	Yes
	Contain diagnostic code for CD	No	No	Yes
Clinical Encounters	Outpatient visits to a gastroenterologist only	No	Yes	Yes

- ***3.2.2.1 Group 1 Algorithms***

The first group of algorithms included an OHIP endoscopy fee code followed by various numbers of clinical encounters and/or hospitalizations. The requirements to meet each component are the least stringent among all 3 groups of algorithms for two reasons: 1) the specialty of the physicians who billed the endoscopy or clinical encounters was not specified; 2)

the endoscopy billing claims did not necessarily have the CD ICD-9 diagnostic code (579) associated with them. The reasoning behind this is that when the endoscopy procedure was performed, the clinician performing the procedure may not have been certain of the diagnosis. This is because histological diagnosis is far more reliable than endoscopic diagnosis, and in milder CD cases the microscopic manifestation of the disease may not be as observable through endoscope. With the less restrictive nature of these algorithms, more cases were captured by this group of algorithms and this group had the highest sensitivity of all 3 groups of algorithm. However, a less restrictive case definition also resulted in more false positive cases identified as CD, thus decreasing the PPVs of the algorithms (Table 3.4).

Table 3-4 Operating Characteristics of the Group 1 Algorithms

Algorithms	Sensitivity	Specificity	PPV	NPV
Scope + 1 or more outpatient contacts	88.70 (81.11-93.60)	99.93 (99.92-99.94)	39.53 (33.58-45.81)	99.99 (99.99->99.9)
Scope + 2 or more outpatient contacts in 1 year	60.87 (51.30-69.70)	99.97 (99.96-99.98)	48.61 (40.25-57.04)	99.98 (99.97-99.99)
Scope + 2 or more outpatient contacts in 2 years	66.09 (57.59-74.49)	99.97 (99.96-99.97)	50.33 (42.12-58.52)	99.98 (99.98-99.99)
Scope+ 2 or more outpatient contacts in 3 years	66.09 (57.59-74.49)	99.97 (99.96-99.97)	50.33 (42.12-58.52)	99.98 (99.98-99.99)
Scope + 2 or more outpatient contacts in 4 years	66.09 (56.59-74.49)	99.97 (99.96-99.97)	49.03 (40.97-57.15)	99.98 (99.98-99.99)
Scope + 2 or more outpatient contacts in 5 years	66.09 (56.59-74.49)	99.97 (99.96-99.97)	49.03 (40.97-57.15)	99.98 (99.98-99.99)
Scope + 1 or more outpatient contacts or 1 hospitalization	91.30 (84.20-95.52)	99.93 (99.91-99.94)	37.77 (32.10-43.78)	>99.9 (99.9->99.9)
Scope + 2 or more outpatient contacts or hospitalizations in 1 year	63.48 (53.93-72.11)	99.97 (99.96-99.97)	48.99 (40.77-57.27)	99.98 (99.98-99.99)
Scope + 2 or more outpatient contacts or hospitalizations in 2 years	67.83 (58.38-76.06)	99.97 (99.96-99.97)	50.32 (42.24-58.40)	99.98 (99.98-99.99)

Scope + 2 or more outpatient contacts or hospitalizations in 3 years	68.70 (59.28-76.84)	99.97 (99.96-99.97)	50.32 (42.27-58.35)	99.98 (99.98-99.99)
Scope + 2 or more outpatient contacts or hospitalizations in 4 years	68.70 (59.28-76.84)	99.97 (99.96-99.97)	49.07 (41.16-57.03)	99.98 (99.98-99.99)
Scope + 2 or more outpatient contacts or hospitalizations in 5 years	68.70 (59.28-76.84)	99.97 (99.96-99.97)	49.07 (41.16-57.03)	99.98 (99.98-99.99)
Scope + 3 or more outpatient contacts in 1 year	29.57 (21.60-38.91)	99.98 (99.97-99.99)	42.50 (31.68-54.05)	99.97 (99.96-99.97)
Scope + 3 or more outpatient contacts in 2 years	33.91 (25.51-43.41)	99.98 (99.97-99.98)	43.82 (33.46-54.72)	99.97 (99.96-99.97)
Scope + 3 or more outpatient contacts in 3 years	37.39 (28.69-46.95)	99.98 (99.97-99.98)	44.79 (34.74-55.26)	99.97 (99.96-99.98)
Scope + 3 or more outpatient contacts in 4 years	38.26 (29.49-47.82)	99.98 (99.97-99.98)	45.36 (35.33-55.76)	99.97 (99.96-99.98)
Scope + 3 or more outpatient contacts in 5 years	38.26 (29.49-47.82)	99.98 (99.97-99.98)	45.36 (35.33-55.76)	99.97 (99.96-99.98)

As seen in Table 3.4, the presence of an endoscopy fee code plus 1 or more CD related visits corresponds to a sensitivity of 88.7%; 88.7% of the true positive cases met this case definition, while the remaining 11.3% of the true positive reference cases that were not captured by the algorithm would be considered false negative cases. An inverse relationship can be observed between the sensitivity and the number of visits specified in the algorithms; this is because as more requirements were added to an algorithm, the case definition became more stringent and sensitivity decreased. Less than 70% of the true positive reference cohort had 2 or more visits, and the sensitivity decreased to <40% with algorithms requiring 3 or more outpatient visits. The small number of CD-related visits generated by the true positive cohort was a limiting factor for the creation of algorithms with higher visit numbers. There were very few CD-related

hospitalizations and therefore algorithms that used DAD hospitalization codes exclusively for the visit component of the algorithms did not achieve adequate sensitivity.

All algorithms achieved specificity of >99.9%. Although all of the algorithms were shown to be highly specific, for a low-prevalence condition such as pediatric biopsy-proven CD, even a small number of false positive cases significantly reduced the PPV of the algorithms. Therefore, in the context of CD, it is more appropriate to evaluate the discriminatory power of the algorithms based on PPV. Unlike sensitivity, the PPV of the algorithms increased as the number of visits required to qualify increased, because a more stringent case definition helps to reduce the number of false positive cases captured. All algorithms in Group 1 had PPVs of 50% or less, meaning that half or more of the cases captured by these algorithms were non-CD patients identified as having CD. The algorithm with the highest PPV, while maintaining an optimal sensitivity, was: 1 endoscopy fee code plus ≥2 CD-related outpatient contacts or hospitalizations within 3 years (sensitivity 68.7% (95% CI: 59.3 to 76.8%), specificity >99.9% (95% CI: >99.9- >99.9%), PPV 50.3% (95% CI: 42.3-58.4%), NPV >99.9% (95% CI: >99.9- >99.9%)).

- ***3.2.2.2 Group 2 Algorithms***

Given that the diagnosis of pediatric biopsy-proven CD would often involve clinical encounters with gastroenterologists, the clinician experts on the thesis advisory committee suggested that the PPV of the algorithms could potentially be increased by limiting the endoscopy fee code and the outpatient contacts to those recorded only by gastroenterologists, and assessing outpatient contacts only after the endoscopy procedure. These added requirements made up the Group 2 algorithms and their performances are displayed in Table 3.5 below:

Table 3-5 Operating Characteristics of the Group 2 Algorithms

Algorithms	Sensitivity	Specificity	PPV	NPV
Scope + 1 or more GI outpatient contacts	83.48 (75.14-89.51)	99.95 (99.94-99.96)	44.04(39.38-52.90)	99.99 (99.99->99.9)
Scope + 2 or more GI outpatient contacts in 1 year	46.96 (37.67-56.45)	99.98 (99.97-99.98)	50.00 (40.29-59.71)	99.97 (99.97-99.98)
Scope + 2 or more GI outpatient contacts in 2 years	51.30 (41.85-60.67)	99.98 (99.97-99.98)	51.75(42.25-61.1)	99.98(99.97-99.98)
Scope+ 2 or more GI outpatient contacts in 3 years	52.17 (42.70-61.50)	99.98 (99.97-99.98)	52.17 (43.70-61.50)	99.98 (99.97-99.98)
Scope + 2 or more GI outpatient contacts in 4 years	52.17 (42.70-61.50)	99.98 (99.97-99.98)	51.28 (41.91-60.56)	99.98 (99.97-99.98)
Scope + 2 or more GI outpatient contacts in 5 years	52.17 (42.70-61.50)	99.98 (99.97-99.98)	51.28 (41.91-60.56)	99.98 (99.97-99.98)
Scope + 1 or more GI outpatient contact or 1 hospitalization	86.96 (79.09-92.27)	99.94 (99.93-99.95)	41.67 (35.41-48.20)	99.99 (99.99->99.9)
Scope + 2 or more GI outpatient contacts or hospitalization in 1 year	50.43 (41.01-59.83)	99.98 (99.97-99.98)	50.88 (41.40-60.30)	99.98 (99.97-99.98)
Scope + 2 or more GI outpatient contacts or hospitalization in 2 years	54.78 (45.25-63.99)	99.98 (99.97-99.98)	52.50 (43.22-61.62)	99.98 (99.97-99.98)
Scope + 2 or more GI outpatient contacts or hospitalization in 3 years	57.39 (48.82-66.46)	99.98 (99.97-99.98)	53.66 (44.47-62.62)	99.98 (99.97-99.98)
Scope + 2 or more GI outpatient contacts or hospitalization in 4 years	57.39 (47.83-66.46)	99.97 (99.97-99.98)	52.80 (43.70-61.72)	99.98 (99.97-99.98)
Scope + 2 or more GI outpatient contacts or hospitalization in 5 years	57.39 (47.83-66.46)	99.97 (99.97-99.98)	52.80 (43.70-61.72)	99.98 (99.97-99.98)

Scope + 1 GI or more outpatient contacts after scope[1]	70.43 (61.09-78.39)	99.97 (99.96-99.98)	53.29 (45.05-61.36)	99.99 (99.98-99.99)

[1]*Algorithm proposed by gastroenterologists on the research team*

Compared to the Group 1 algorithms, the algorithms in Group 2 had lower sensitivities. The PPV of the Group 2 algorithms increased marginally in comparison to the algorithms in Group 1, with several algorithms achieving PPVs of just over 50%. Of Group 2 algorithms, the best performing algorithm was: one endoscopy fee code plus ≥1 outpatient contacts after the endoscopy procedure with associated CD diagnostic code (sensitivity 70.4% (95% CI 61.1-78.4%), specificity >99.9% (95% CI: >99.9- >99.9%), PPV 53.3% (95% CI: 45.1-61.4), NPV >99.9% (95% CI: >99.9- >99.9%).

- ***3.2.2.3 Group 3 Algorithms***

The last group of algorithms involved changing the source of the endoscopy fee code from OHIP to SDS. In the SDS database, up to 20 diagnostic codes are associated with a single procedural record, while the OHIP record is only associated with a single diagnostic code. This increased the likelihood for the documentation of an endoscopy procedure to be associated with a CD diagnostic code, thus making the endoscopy component of the algorithms more specific to CD when compared to Group 1 and Group 2 algorithms. Table 3.6 demonstrates the operating characteristics of the Group 3 algorithms.

Table 3-6 Operating Characteristics of Group 3 Algorithms

Algorithms	Sensitivity	Specificity	PPV	NPV
Scope + 1 or more GI outpatient contacts	71.30 (62.00-79.16)	>99.9 (99.99->99.9)	92.13 (83.94-96.51)	99.99 (99.98-99.99)
Scope + 2 or more GI outpatient contacts in 1 year	38.26 (29.49-47.83)	>99.9 (>99.9->99.9)	93.62 (81.44-98.33)	99.97 (99.96-99.98)
Scope + 2 or more GI outpatient contacts in 2 years	41.74 (32.73-51.31)	>99.9 >(99.9->99.9)	94.12 (82.77->98.47)	99.97 (99.96-99.98)
Scope+ 2 or more GI outpatient contacts in 3 years	42.61 (33.54-52.17)	>99.9 (>99.9->99.9)	94.23 (83.08->98.50)	99.97 (99.96-99.98)
Scope + 2 or more GI outpatient contacts in 4 years	42.61 (33.54-52.17)	>99.9 (>99.9->99.9)	94.23 (83.08->98.50)	99.97 (99.96-99.98)
Scope + 2 or more GI outpatient contacts in 5 years	42.61 (33.54-52.17)	>99.9 (>99.9->99.9)	94.23 (83.08->98.50)	99.97 (99.96-99.98)
Scope + 1 or more GI outpatient contact or 1 hospitalization	76.52 (67.53-83.71)	99.99 (99.98->99.9)	76.52 (67.53-83.71)	99.99 (99.98-99.99)
Scope + 2 or more GI outpatient contacts or hospitalization in 1 year	41.74 (32.73-51.31)	>99.9 (>99.9->99.9)	92.31 (80.60-97.51)	99.97 (99.96-99.98)
Scope + 2 or more GI outpatient contacts or hospitalization in 2 years	45.22 (36.01-54.75)	>99.9 (>99.9->99.9)	92.86 (81.87-97.69)	99.97 (99.97-99.98)
Scope + 2 or more GI outpatient contacts or hospitalization in 3 years	47.83 (38.50-57.30)	>99.9 (>99.9->99.9)	93.22 (82.73-97.81)	99.97 (99.97-99.98)
Scope + 2 or more GI outpatient contacts or hospitalization in 4 years	47.83 (38.50-57.30)	>99.9 (>99.9->99.9)	93.22 (82.73-97.81)	99.97 (99.97-99.98)

Scope + 2 or more GI outpatient contacts or hospitalization in 5 years	47.83 (38.50-57.30)	>99.9 (>99.9->99.9)	93.22 (82.73-97.81)	99.97 (99.97-99.98)
Scope + 1 or more GI outpatient contacts after scope	60.87 (51.30-69.70)	>99.9 (99.99->99.9)	92.11 (83.00-96.75)	99.98 (99.97-99.99)

Overall, the algorithms in Group 3 had the lowest sensitivities when compared to the algorithms in Group 1 and Group 2; the endoscopy fee code plus one or more visits captured only 71.3% of the true positive cases in contrast to the >80% in Group 1 and Group 2. For the algorithms with two or more visits, the sensitivities decreased below 50%. This confirmed the hypothesis that during the disease investigation process, the coder responsible for the SDS record often did not include the CD diagnostic code with the procedural record, potentially due to diagnostic uncertainty.

The algorithms in Group 3 have the highest PPV of all 3 groups of algorithms. Almost all of the algorithms had a PPV of >90%. The algorithm with the best balance of sensitivity and PPV was: one SDS endoscopy code plus ≥1 outpatient contacts with a gastroenterologist (sensitivity 71.3% (95% CI: 62.0-79.2%), specificity >99.9% (95% CI: >99.9- >99.9%), PPV 92.1% (95% CI: 83.9-96.5%), NPV 99.99% (95% CI: 99.98-99.99%)).

- ***3.2.2.4 Testing the Performance of the Algorithms in Different Age Groups***

In the above analyses, the pediatric population used to test the performance of the algorithms were aged 6 months to 14 years. However, the algorithms may perform differently in different age groups due to different health service utilization patterns.(64) For example, children of younger age may require more clinical visits for the disease management process, while children of older age may be more independent and thus require fewer follow-ups. We tested the algorithms based on the age at diagnosis with CD: <12 years, <10 years, and <8 years (Table 3.7). The accuracies of the algorithms did not differ substantially in these age groups,

and the overall trends in sensitivities and PPVs were very similar to the results above. This may suggest that the health services utilization pattern for CD is similar across children in different age groups in Ottawa. Therefore, in the remainder of the thesis, the results will be discussed based on the children aged 6 months to 14 years population.

Table 3-7 Operating characteristics of the Group 1-3 algorithms in different age subgroups (<12, <10 and <8 years old)

	Age at Diagnosis					
	<12 years (113 true CD cases)		<10 years (108 true CD cases)		<8 years (97 true CD cases)	
Algorithms	**Sen**	**PPV**	**Sen**	**PPV**	**Sen**	**PPV**
Group 1 Algorithms						
Scope + 1 GI outpatient contact	88.5	39.06	87.96	38.62	86.60	37.00
Scope + 2 GI outpatient contacts in 1 year	60.18	47.89	58.33	47.01	57.73	44.44
Scope + 2 GI outpatient contacts in 2 years	65.49	49.66	63.89	48.94	62.89	46.21
Scope+ 2 GI outpatient contacts in 3 years	65.49	49.66	63.89	48.94	62.89	46.21
Scope + 2 GI outpatient contacts in 4 years	65.49	48.37	63.89	47.59	62.89	44.85
Scope + 2 GI outpatient contacts in 5 years	65.49	48.37	63.89	47.59	62.89	44.85
Scope + 1 GI outpatient contact or 1 hosp	91.15	37.45	90.74	37.55	89.69	36.40
Scope + 2 GI outpatient contacts or hosp in 1 year	62.83	48.30	61.11	47.48	60.82	45.38
Scope + 2 GI outpatient contacts or hosp in 2 years	67.26	49.67	65.74	48.97	64.95	46.67
Scope + 2 GI outpatient	68.14	49.68	66.67	48.98	65.98	46.72

contacts or hosp in 3 years						
Scope + 2 GI outpatient contacts or hosp in 4 years	68.14	48.43	66.67	47.68	65.98	45.39
Scope + 2 GI outpatient contacts or hosp in 5 years	68.14	48.43	66.67	47.68	65.98	45.39
Scope + 1 GI outpatient contacts 7 days after scope	76.11	49.43	75.00	49.09	73.20	47.02
Group 2 Algorithms						
Scope + 1 GI outpatient contact	83.19	43.52	82.41	43.00	81.44	41.58
Scope + 2 GI outpatient contacts in 1 year	46.02	49.06	45.37	48.51	45.36	46.32
Scope + 2 GI outpatient contacts in 2 years	50.44	50.89	50.00	50.47	50.52	48.51
Scope+ 2 GI outpatient contacts in 3 years	51.33	51.33	50.93	50.93	51.55	49.02
Scope + 2 GI outpatient contacts in 4 years	51.33	50.43	50.93	50.00	51.55	48.08
Scope + 2 GI outpatient contacts in 5 years	51.33	50.43	50.93	50.93	51.55	48.08
Scope + 1 GI outpatient contact or 1 hosp	86.73	41.35	86.11	41.52	85.57	40.69
Scope + 2 GI outpatient contacts or hosp in 1 year	49.56	50.0	49.07	49.53	48.45	47.47
Scope + 2 GI outpatient contacts or hosp in 2 years	53.98	51.69	53.70	51.33	53.61	49.52
Scope + 2 GI outpatient contacts or hosp in 3 years	56.64	52.89	56.48	52.59	56.70	50.93
Scope + 2 GI outpatient contacts or hosp in 4 years	56.64	52.03	56.48	51.69	56.70	50.00
Scope + 2 GI outpatient contacts or hosp in 5 years	56.64	52.03	56.48	51.69	56.70	50.00

Scope + 1 GI outpatient contacts 7 days after scope	69.91	52.67	69.44	52.45	68.04	50.77
Group 3 Algorithms						
Scope + 1 GI outpatient contact	70.80	91.95	69.44	91.46	68.04	91.67
Scope + 2 GI outpatient contacts in 1 year	37.17	93.33	36.11	92.86	36.08	92.11
Scope + 2 GI outpatient contacts in 2 years	40.71	93.88	39.81	93.48	39.18	92.68
Scope+ 2 GI outpatient contacts in 3 years	41.59	94.00	40.74	93.62	40.21	92.86
Scope + 2 GI outpatient contacts in 4 years	41.59	94.00	40.74	93.62	40.21	92.86
Scope + 2 GI outpatient contacts in 5 years	41.59	94.00	40.74	93.62	40.21	92.86
Scope + 1 GI outpatient contact or 1 hosp	76.11	76.79	75.00	78.64	74.23	80.90
Scope + 2 GI outpatient contacts or hosp in 1 year	40.71	92.00	39.81	91.49	39.18	90.48
Scope + 2 GI outpatient contacts or hosp in 2 years	44.25	92.59	43.52	92.16	42.27	91.11
Scope + 2 GI outpatient contacts or hosp in 3 years	46.90	92.98	46.30	92.59	45.36	91.67
Scope + 2 GI outpatient contacts or hosp in 4 years	46.90	92.98	46.30	92.59	45.36	91.67
Scope + 2 GI outpatient contacts or hosp in 5 years	46.90	92.98	46.30	92.59	45.36	91.67
Scope + 1 GI outpatient contacts 7 days after scope	60.15	91.89	59.26	91.43	57.73	90.32

- ***3.2.2.5 Testing the Assumption of CHEO as the Sole Pediatric Endoscopy Centre in Ottawa***

An incomplete reference cohort could falsely reduce the PPV of the tested algorithms.(66) For example, if some children underwent the endoscopic procedure outside of CHEO but within Ottawa, it would be missed from the CHEO chart review process, thus not being included in the true positive reference cohort. However, they could still meet the algorithm criteria proposed, and would be mistakenly classified as false positive cases when compared to the reference cohort.

To test the validity of the assumption that CHEO was the sole pediatric endoscopy centre for CD patients, we applied an algorithm from Group 2: one endoscopy fee code plus one or more CD related visits from gastroenterologists, and examine the source of the endoscopy health claims from the false positive patients. The algorithm had a PPV of 44.0%, and it captured 122 false positive cases. We then ascertained all of the endoscopy fee codes from these false positive cases; these 122 false positive patients had a total number of 150 health claims for endoscopy. 145 of the 150 claims (96.7%) were associated with a facility code billed for CHEO, with 1 claim from The Ottawa Hospital (Genera Campus), a local adult hospital, and the other 4 claims from hospitals outside of Ottawa. This revealed that CHEO remained the major endoscopy centre in the city. Therefore, the poor PPVs of the validated algorithms could not be attributed to the quality of the reference cohort.

3.3 Ontario-Wide Incidence of Biopsy-Proven CD Using the Selected Algorithms

In the study design phase, we pre-specified that to accurately capture CD cases from Ontario health administrative data, the best algorithm should have a PPV of >90% while maintaining sensitivity of >80%. None of the algorithms from the three groups met these criteria. The algorithms from Group 1 and Group 2 suffered from poor PPV, while the algorithms in Group 3 had poor sensitivity. Thus, depending on the algorithm in use, the incidence of CD in Ontario could be under- or over-estimated. Based on our results, we concluded that none of the

algorithms were suitable to accurately identify biopsy-proven CD from Ontario health administrative data. Nevertheless, to demonstrate the effect of using poorly performing algorithms, two “best” algorithms were chosen from the three groups of algorithms applied to Ontario health administrative data to demonstrate the range of incidence estimates. Due to the similarity in algorithm structure and performance, only one algorithm was selected from the Group 1 and Group 2 algorithms. The other algorithm would be chosen from the Group 3 algorithms. Upon consulting with the pediatric gastroenterologists in the research team, the two selected algorithms were: 1) OHIP endoscopy fee code plus ≥1 outpatient contacts with a gastroenterologist after the endoscopy from Group 2 and 2) SDS endoscopy fee code plus ≥1 gastroenterologist visits from Group 3. They will be referred to as Group 2 algorithm and Group 3 algorithm in the remainder of the thesis. To estimate the number of pediatric biopsy-proven CD from 1995-2011 in Ontario, the health claims related to CD were ascertained from the OHIP and SDS databases for children aged 6 months to 14 years who met either or both of the endoscopy components from the two selected algorithms.

The selected Group 2 algorithm identified a total number of 1289 cases of CD, while the Group 3 algorithm captured 948 cases. Using the Registered Persons Database (RPDB), basic demographic information from all children aged 6 months to 14 years who resided in Ontario during 1995-2011 (n=5,736,813) were obtained and used to calculate the incidence of biopsy-proven CD. The RPDB was used to determine OHIP eligibility status and the 2006 Census (with inter-censal annual estimates) was used to calculate annual standardized incidence. Using the Group 2 and Group 3 algorithms, the crude incidence were 1.98 (95% CI: 1.87 to 2.09) and 1.46 (95% CI: 1.37 to 1.55) cases per 100,000 PY respectively. The annual standardized incidence estimates are presented in Table 3.8 below. Figure 2 demonstrates the trends in biopsy-proven CD diagnosis based on the two selected algorithms.

Table 3-8 Incidence rate of pediatric biopsy-proven CD per 100,000 patients in Ontario, standardized to the 2006 Canadian census population based on sex and age structure

Year	Group 2 algorithm		Group 3 algorithm	
	Number of cases (total=1289)	Standardized rate(95% CI)	Number of cases (total=948)	Standardized rate(95% CI)
1995	26	1.11 (0.72-1.64)	13	0.53 (0.28-0.90)
1996	23	0.85 (0.53-1.29)	24	0.95 (0.60-1.41)
1997	31	1.21 (0.81-1.74)	32	1.28 (0.87-1.81)
1998	39	1.55 (1.08-2.14)	31	1.23 (0.83-1.76)
1999	32	1.17 (0.79-1.67)	26	1.06 (0.69-1.55)
2000	55	2.16 (1.61-2.83)	43	1.72 (1.24-2.33)
2001	50	2.00 (1.46-2.68)	30	1.06 (0.70-1.54)
2002	63	2.25 (1.71-2.91)	54	2.00 (1.49-2.63)
2003	58	2.31 (1.74-3.01)	33	1.33 (0.91-1.87)
2004	69	2.90 (2.25-3.69)	52	2.12 (1.58-2.79)
2005	88	3.42 (2.73-4.24)	67	2.68 (2.07-3.42)
2006	123	5.30 (4.39-6.35)	72	3.00 (2.34-3.79)
2007	121	5.08 (4.20-6.09)	92	3.81 (3.06-4.70)
2008	110	4.53 (3.71-5.47)	94	3.92 (3.16-4.81)
2009	147	6.09 (5.12-7.19)	112	4.70 (3.85-5.67)
2010	115	4.73 (3.89-5.70)	87	3.55 (2.82-4.40)
2011	139	5.62 (4.70-6.66)	86	3.51 (2.80-4.36)

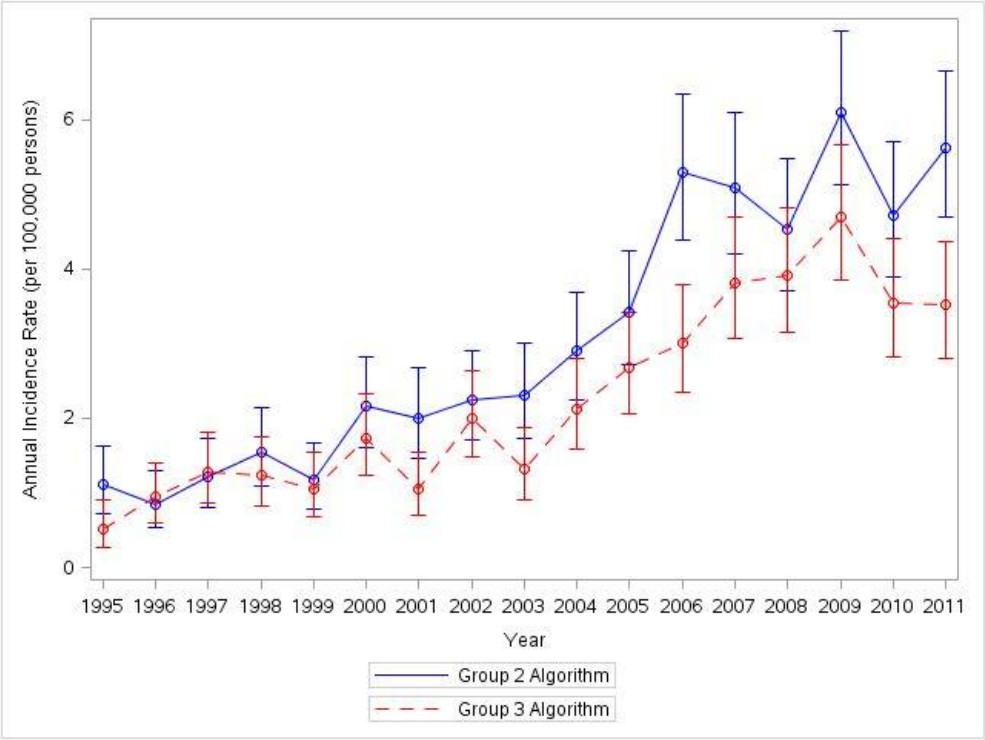

Figure 3.2 Annual standardized incidence rate of pediatric biopsy-prove CD as estimated by selected algorithms

Standardized incidence of pediatric CD increased significantly in 2002 when compared to 1995; the incidence rate doubled from 1.11 (95% CI 0.72-1.64) to 2.25 (95% CI 1.71-2.91) per 100,000 PY when the Group 2 algorithm was used, while the incidence rate obtained from Group 3 algorithm increased by four fold from 0.53 (95% CI 0.28-0.90) to 2.00 (95% CI 1.49-2.63) per 100,000 PY. For both algorithms, the incidence continued to increase after 2002 and reached its highest point within the study period in 2009 (Group 2 algorithm: 6.09 (95% CI 5.12 to 7.19) cases per 100,000 PY; Group 3 algorithm: 4.70 (95% CI 3.85-5.67) cases per 100,000 person years).

For both algorithms, Poisson regression was used to explore the trends in CD incidence from 1995 to 2011, adjusting for age and gender. Based on the deviance and Pearson Chi-square statistics of the models, both models fitted well and over-dispersion of data was not detected. We examined the presence of interaction among the three covariates, and an interaction between age and year of diagnosis was detected in the model based on Group 3 algorithm, but

not Group 2 algorithm. To test the linearity assumption between year of diagnosis and incidence rate, models with square and cubic terms of diagnosis year were tested and no significant quadratic effects were detected. Table 3.9 demonstrates the results from Poisson regression for the two algorithms.

Table 3-9 Trends in incidence over time by Poisson regression, adjusting for gender and age of diagnosis

Variables	Estimates (95% CI)	Rate Ratio(95% CI)[1]	P-value
Group 2 Algorithm			
Intercept	1.07 (0.90 to 1.25)	2.91 (2.45 to 3.50)	<0.001
Year of diagnosis	0.08 (0.07 to 0.10)	1.09 (1.07 to 1.10)	<0.001
Gender (Male)	-0.33 (-0.45 to -00.22)	0.72 (0.64 to 0.80)	<0.001
Age of diagnosis	-0.06 (-0.07 to -0.04)	0.94 (0.93 to 0.96)	<0.001
Group 3 Algorithm			
Intercept	1.23 (0.95 to 1.50)	3.41 (2.60 to 4.48)	<0.001
Year of diagnosis	0.04 (0.02 to 0.07)	1.04 (1.02 to 1.07)	<0.001
Gender (Male)	-0.36 (-0.48 to -0.24)	0.70 (0.62 to 0.79)	<0.001
Age of diagnosis	-0.09 (-0.12 to -0.05)	0.91 (0.88 to 0.95)	<0.001
Year of diagnosis x Age of diagnosis	0.004 (0.0004 to 0.007)	1.004 (1.0004 to 1.007)	0.028

[1]*Calculated by exponentiating the estimates*

For the model constructed based on Group 2 algorithm, a 9% yearly increase in incidence rate was observed. The exponentiated intercept represents the incidence of CD among female aged 6 months to 1 year who were diagnosed in 1995, which was 2.91 cases per 100,000 persons. The incidence of CD was lower among male and older children; the incidence for male was 28%

lower than the female's, and a unit increase in age of diagnosis represented a 6% decrease in the incidence.

While the model using the Group 3 algorithm included an interaction term between the year of diagnosis and the age of diagnosis, it yielded similar results as the model from Group 2 algorithm. The baseline incidence rate is 3.41 cases per 100,000 PY among female patients aged 6 months to 1 year who were diagnosed in 1995. The CD incidence rate among male is 30% lower than females. To interpret the trends in CD incidence, the effect of the year of diagnosis was considered along with the age of diagnosis. Table 3.10 demonstrates the incidence of CD for female and male children over time.

Table 3-10 Effect of the Year x Age interaction on the incidence rate estimates for Group 3 algorithm model

Year[1]	Age	Incidence rate for female[2]	Incidence rate for male[3]
1996	1	3.26 (2.33 to 4.57)	2.28 (1.44 to 3.59)
2000	1	3.92 (2.51 to 6.14)	2.73 (1.55 to 4.83)
2005	1	4.93 (2.74 to 8.89)	3.44 (1.69 to 7.00)
2010	1	6.21 (3.00 to 12.86)	4.33 (1.85 to 10.13)
1996	4	2.52 (1.61 to 3.96)	1.76 (0.99 to 3.12)
2000	4	3.16 (1.74 to 5.76)	2.20 (1.07 to 4.54)
2005	4	4.19 (1.91 to 9.21)	2.92 (1.18 to 7.25)
2010	4	5.56 (2.10 to 14.72)	3.88 (1.30 to 11.59)
1996	8	1.79 (0.98 to 3.27)	1.25 (0.61 to 2.58)
2000	8	2.37 (1.06 to 5.29)	1.65 (0.66 to 4.17)
2005	8	3.37 (1.18 to 9.65)	2.35 (0.73 to 7.60)
2010	8	4.80 (1.31 to 17.60)	3.35 (0.81 to 13.86)
1996	12	1.27 (0.60 to 2.70)	0.89 (0.37 to 2.13)
2000	12	1.78 (0.65 to 4.86)	1.24 (0.40 to 3.83)
2005	12	2.71 (0.73 to 10.12)	1.89 (0.45 to 7.97)
2010	12	4.14 (0.82 to 21.06)	2.89 (0.50 to 16.58)

[1]*Year is rescaled relative to 1995 such that 1995=0, 1996=1, 1997=2 etc.*

[2]*Rate ratio for female:* $e^{Intercept+\beta_{year}(year)+\beta_{age}(age)+\beta_{year\ x\ age}(year\ x\ age)}$

[3]*Rate ratio for male:* $e^{Intercept+\beta_{year}(year)+\beta_{age}(age)+\beta_{sex}(sex)+\beta_{year\ x\ age}(year\ x\ age)}$

As observed in Table 3.10, when age is held constant, the rate ratio for year generally increases as the year of diagnosis increases; among 8 year old female the incidence of CD increased from 1.79 (95% CI: 0.98 to 3.27) in 1996 to 4.80 (95% CI: 1.31 to 17.60) cases per 100,000 persons in 2010. Similar to the trends observed from the model based on Group 2 algorithm,

females had a higher CD incidence than males at any time point and age range. In addition, the incidence of CD decreased as age increased; the incidence among males decreased from 2.92 (95%CI: 1.18 to 7.25) cases per 100,000 PY among the 4 year olds to 1.89 (95%CI: 0.45 to 7.97) cases per 100,000 PY among the 12 year olds in 2005.

Chapter 4 : Discussion

The goals of this chapter are to:

- **Review the findings from the Ottawa reference standard cohort**
- **Interpret the results from the algorithms derivation**
- **Explore the reasons behind the poor performances of the algorithms**
- **Discuss the feasibility of algorithm adjustment**
- **Examine the findings from Ontario CD incidence estimation**
- **Study limitation and conclusion**

The work described in this thesis evaluated the feasibility of using health administrative data to capture children with biopsy-proven CD, and explored the trends in incidence over the past two decades in Ontario. Our findings demonstrate that health administrative data is not an appropriate source for capturing biopsy-proven CD among children, at least not with currently available Ontario data comprising mostly outpatient physician billing and hospitalization data. All of the algorithms derived in the study suffered from low sensitivity and/or low PPV, and we were not able to identify an algorithm that can maximize the ascertainment of true positive cases while keeping the number of false positive cases captured at a minimum. We applied two of the test algorithms to the Ontario health administrative data to illustrate the effect of using algorithms for incidence rate estimation. The Group 2 and Group 3 algorithms, differing by the source of their endoscopy procedural code, both detected an increased trend in CD-related health services during the study time period (1995 to 2011). However, due to the suboptimal performances of the two algorithms, we are hesitant to generate any conclusion regarding the trends in pediatric CD incidence over the past 2 decades.

4.1 The Ottawa reference standard cohort

To ensure maximum case ascertainment with high accuracy, we created a reference standard cohort consisting of the entire Ottawa pediatric population. We used this cohort to test the performance of different identification algorithms prior to applying them to Ontario health administrative data. We were only able to identify two others studies that evaluated the accuracy of the health administrative data codes for case ascertainment of celiac disease patients. One study from UK only used a positive reference cohort of 38 patients (48), while another study from Denmark used patients from a single hospital as reference.(20) Both reference cohorts were not accurate representations of the overall population, and as a result, they would not be able to evaluate how the codes would perform at greater population level. Specifically, since they either lacked a true negative reference cohort or did not use a true negative reference cohort that is representative of the population, the prevalence of CD in their reference cohort would be higher than the general population's, and as a result, the PPVs of their algorithms would be inflated. Therefore, their incidence estimates would likely be overestimated when compared to the true CD incidence estimate in the general population, and their algorithms would capture many false positive cases when applied at a population level. In the current study, the Ottawa population-based reference cohorts would allow us to better assess how the health administrative codes were being used in both the disease and non-disease population, thus producing more accurate estimates for the operating characteristics of the algorithms.

The Ottawa cohort represented an incidence of 8.7 cases per 100,000 PY. This is comparable to the only population-based pediatric CD incidence study in North America from Alberta, which reported an incidence of 7.3 cases per 100,000PY from 2000 to 2006.(5) Our incidence is lower than estimates from some of the European studies. In Scotland, the incidence was reported to be 11.7 cases per 100,000 PY from 2005 to 2009(6), while in Sweden the incidence was

reported to be 42 cases per 100,000 PYs from 2004 to 2009.(46) One major factor that could explain the difference in incidence estimates between North America and Europe is the diagnostic criteria for CD. In some European centres, positive results from serologic testing (tTG antibody titre) would suffice for the diagnosis of CD. The clinical practice guidelines recently published by ESPGHAN suggested that duodenal biopsy could be forgone if the patient's anti-TTG antibody titre is 10 times above the upper limit of normal.(36)

In our Ottawa reference cohort, we identified 115 true positive CD cases from 1995 to 2011. 54% of the cases diagnosed were female. Other studies have determined that female were at increased risk for CD, with an estimated female to male ratio of 2 to 1.(5, 46, 67) The median age of diagnosis from the current study is 7.4 years, which is comparable to the median age reported in other pediatric incidence studies.(5, 46) The age of CD diagnosed has been observed to be increasing steady over the past 2 decades; what was once a disease that was diagnosed mostly under the age of 2 years has now been commonly diagnosed among children of older age.(6) The Ottawa cohort only spanned 6 years (2005 to 2011); therefore, it may not be of sufficient size to draw accurate conclusion about the trends in age in diagnosis. Regardless, it is clear that the median age reported is greater than the age of diagnosis a decade ago as reported in the literature.

4.2 Algorithm Derivation

As mentioned in the previous chapters, it usually takes more than 1 clinical encounter for a chronic condition to be diagnosed. The use of a single diagnostic code, as demonstrated in our study, resulted in a large amount of false positive cases. In the Ottawa derivation cohort, the use of a single diagnostic code for CD (OHIP-based outpatient contact using ICD code 579) yielded a PPV of 8.9%, suggesting that over 90% of the cases captured by the algorithm would be false positive cases. West et al. reported that the application of a single disease code for CD within the Clinical Practice Research Datalink (CRPD) in UK yielded a PPV 81%.(48) However,

the code was evaluated in a sample of 38 patients; without a true negative cohort reflective of the lower prevalence of celiac in the population. Therefore, the PPV was likely inflated. The importance of a reference cohort with disease prevalence reflective of the database cohort has been addressed in multiple reports.(51, 57, 68)

To ensure that the CD cases captured were biopsy-proven, all of the algorithms proposed in the study required the presence of an endoscopy fee code. We established different criteria for the endoscopy component of the algorithm based on the source of the health claims. For the endoscopy fee codes that were obtained from the OHIP outpatient physician billing database (Group 1 and Group 2 algorithms), the claims for the endoscopy were not limited to those that were associated with the diagnostic code for CD, while the claims obtained from SDS (Group 3 algorithms) were associated with the diagnostic code. As discussed above, clinicians performing the endoscopy or coders may not associate the endoscopy claim with the diagnostic code for CD. One major difference between the OHIP and SDS billing database is that SDS allows for up to 20 diagnostic codes to be attached with a single health contact while OHIP only allows for one diagnostic code per claim. This may increase the likelihood for a CD diagnostic code to be billed with the endoscopy procedure in SDS. Therefore, we tested the Group 3 algorithms' performance by both including and excluding the CD diagnostic code from the endoscopy component of the algorithms. When the CD diagnostic code was removed from the endoscopy component of the Group 3 algorithms, their performances were similar to that of the Group 1 and Group 2 algorithms (not shown in result session). When the CD was included in the endoscopy component of the algorithms, almost every algorithms in the group achieved a PPV of >90% (see Table 3.6), suggesting that the presence of a CD diagnostic code in a SDS endoscopy procedural record is highly predictive of being a true positive biopsy-prove CD case.

These observations suggest that the difference in coding practice between these two databases may have contributed to the results observed in Group 3 algorithms. Unlike the OHIP billing

database, in SDS the procedural records are not recorded directly by clinicians but rather by professional hospital coders. It is possible that the SDS records were coded and submitted after the diagnostic outcome had become clear. As a result, the true positive cases' endoscopy procedural code in SDS would be more likely to be billed along with the CD diagnostic code, thus raising the PPVs of the Group 3 algorithms. Should such difference in coding practice exist at CHEO only, it would jeopardize the generalizability of the Group 3 algorithms, as this practice may not be in place in all Ontario hospitals. In addition, since SDS only documents procedures performed in hospitals, if an endoscopy for CD investigation is performed outside of a hospital, it would not be captured by the Group 3 algorithms. This will likely be an issue if we were to apply the Group 3 algorithms to the Ontario cohort, as some endoscopy procedures maybe performed outside of a hospital in areas where there is no pediatric care centres nearby.

In general, the more stringent a case identifying algorithm, the less likely it is for the algorithm to misclassify a non-case as positive case. However, as a tradeoff, the sensitivity of the algorithm would decrease as it becomes more difficult to meet the case definition. In the current study, the second component of all of the algorithms evaluated in the study specified the number of CD related visits before or after the endoscopy procedure. Interestingly, biopsy-proven CD patients in Ottawa did not seem to have many healthcare contacts with associated diagnostic code following the initial diagnosis period. This is demonstrated by the low sensitivity of all algorithms requiring more than one healthcare contact. Out of all of the patients diagnosed with CD, fewer than 70% had two or more visits or hospitalizations related to CD (Group 1 algorithms). When only clinical encounters from gastroenterologists were considered, fewer than 60% had two or more CD related visits (Group 2 algorithms). For Group 3 algorithms, fewer than 50% of patients had two or more CD related outpatient contacts or hospitalizations. This indicates that once the diagnosis is established and therapeutic diet initiated, the diagnosis of CD likely does not result in cause-specific health services utilization. This is supported by a recent study which

reported that pediatric CD patients had significantly lower health service usage after diagnosis.(69)

4.3 Adjusting Algorithms Based on Diagnostic Accuracy Measures

To remedy the poor performances of some health administrative data algorithms, Couris et al.(70) proposed a method to adjust for the poor sensitivity and/or specificity of the algorithms. They proposed the following formula:

$$Kc_{real} = \frac{Kc_{rec} - (1 - Specificity)N}{Sensitivity + Specificity - 1}$$

(where N is the population at risk of the condition, Kc_{real} is the corrected number of incident cases, and Kc_{rec} is the recorded number of cases as obtained by the algorithm)

The equation essentially subtracts the number of false positive cases estimated from the specificity while adding the estimated number of false negative cases based on the sensitivity of the algorithm. Using our Group 2 algorithm as an example, (Sensitivity: 70.4%, Specificity: 99.9%), we identified 152 cases when applied to the Ottawa health administrative data from 2005-2011. Based on the retrospective chart review, there were 115 cases of CD cases diagnosed at CHEO in Ottawa. With a sensitivity of 70.43% and a specificity of 99.97%, the algorithm only included 81 out of the 115 true cases with the remaining being false positive cases.

Applying the adjusting formula from Couris et al.:

$$Kc_{real} = \frac{152 - (1 - 0.9997)235{,}435}{0.7043 + 0.9997 - 1} = 116\ cases$$

Based on the adjustment, the Group 2 algorithm would estimate 116 CD cases, which is very close to the number of true cases from chart review (115 cases). However, when we apply the adjustment method to the Group 3 algorithm, which has a sensitivity of 71.30% and a specificity of 99.99%, the adjusted number of cases becomes 92 cases versus 89 cases obtained by the unadjusted Group 3 algorithm. When the algorithm requiring only a single diagnostic code (sensitivity: 88.70%, specificity: 99.95%) is adjusted using the equation, the number of adjusted

cases is 1160 compared to the 1146 cases from the unadjusted algorithm. Table 4.1 below summarizes the application of the adjustment method to the 3 algorithms mentioned above:

Table 4-1 Application of the algorithm adjustment method to selected algorithm

(Number of true positive cases=115)

Algorithms	Cases captured by algorithm	Algorithm application	Number of cases after adjustment
Single Diagnostic code for CD	1146	$\frac{1146 - (1 - 0.9995)235{,}435}{0.8870 + 0.9995 - 1}$	1160
Group 2 Algorithm	152	$\frac{152 - (1 - 0.9997)235{,}435}{0.7043 + 0.9997 - 1}$	116
Group 3 Algorithm	89	$\frac{89 - (1 - 0.9999)235{,}435}{0.71.30 + 0.9999 - 1}$	92

The performance of the adjustment formula is very inconsistent. The formula does not perform well when the disease is rare and the specificity of the algorithm is extremely high. Therefore, we feel that even with the existence of the adjustment method it would still not justify estimating incidence using algorithms with poor performance.

4.4 The Incidence of CD in Ontario

To illustrate the consequences of using algorithms with poor performance to estimate the incidence of CD, we applied the Group 2 and Group 3 algorithms to the Ontario health administrative data and compared the results. From 1995 to 2011, the Group 2 algorithms and Group 3 algorithms captured 1289 and 948 “CD cases” respectively. The Group 2 algorithm ascertained more cases in almost every year. Both algorithms were similarly sensitive (70.4% for Group 2 versus 71.3% for Group 3), but the PPV of the Group 3 algorithm was almost 40% higher than that of the Group 2 algorithm. This may explain the higher number of cases obtained from the Group 2 algorithm. Therefore, the Group 2 algorithm likely misclassified many more non-CD patients as having CD.

Overall, with a moderate sensitivity and high PPV, the incidence rates obtained from the Group 3 algorithm may underestimate the true incidence rate. The Group 2 algorithm had a moderate sensitivity but poor PPV, and therefore may either over- or under-estimate the true incidence. It is likely that the true incidence of biopsy-proven CD in Ontario children falls between the estimates provided by the two algorithms.

There are several specific time points in our study period (1995-2011) that warrant our special attention. Firstly, in 1997, the IgA-based EMA test was introduced in Canada (5). McGowan et al. examined the impact of serology screening for CD, and noted that the incidence of CD in the post-test era (2000-2006) was over 3 times that of the pre-test era (1990-1996).(5) We noticed a similar trend in the number of CD cases ascertained by the two algorithms. Assuming 1997-1999 to be transition years, more cases were captured by the algorithms in later years, compared to the pre-test era.

In 2002, the ICD-10 and new procedural variables were adopted by the SDS database. These changes would only have impacted the Group 3 algorithm, as its endoscopy component is based on the SDS. It is interesting to note that for the Group 3 algorithm, the number of cases obtained from 2002 onwards were all higher than the cases from 1995-2001. The Ottawa reference cohort only contained cases diagnosed from 2005-2011. As a result, we were only able to evaluate the performance of the algorithms in the years following the change to ICD-10. We were unable to determine whether the ICD-9 codes used before 2002 by SDS had similar diagnostic accuracy statistics. However, given that the trends in incidence rates were very similar between the two algorithms, there is no evidence that the change in coding system resulted in a change in the performance of the algorithms.

The standardized annual incidence in the Ottawa cohort was consistently higher than the Ontario estimates obtained from the two algorithms applied to Ontario-wide data. Since the

derivation was done in Ottawa, it is possible that the health services patterns from Ottawa could not be applied to other regions. Children with biopsy-proven CD may have fewer healthcare contacts outside of Ottawa, and the algorithms could have failed to capture CD cases outside of Ottawa. For example, in smaller communities without specialized pediatric care facilities, the endoscopies may be performed in community clinics by adult gastroenterologists. If true, this would particularly harm the validity of the Group 3 algorithm, since the endoscopy component of the algorithm was based in SDS, which only documents medical procedures performed in hospitals. Therefore, algorithms validated using reference standard cohorts from one region should also be validated in other regions and in other practice types.

With the suboptimal performances of the algorithms, we cannot rely on the accuracy of the incidence estimates. However, since all of the algorithms were composed of the health claims related to CD, one can still infer that the health system contacts related to CD has increased substantially over the past two decades regardless of whether these heath services were provided to true cases or non-cases. The increase in CD-related health services may be caused by increased patient or physician awareness of the condition and better disease detection. A pediatric longitudinal study from UK claimed that close to 90% of CD cases were left undiagnosed in the population.(71) However, the increased coding for CD may have resulted from the increased awareness of non-celiac gluten sensitivity driven by the popular press and reported in new scientific literature.(72) Patients may describe gluten sensitivity to their physician at health care visits. Whether because of misinformation or because no ICD code exists for non-celiac gluten sensitivity, these patients may be incorrectly assigned a diagnosis of CD, even in the presence of a normal duodenal biopsy.

Growth in the field of pediatric gastroenterology as a specialty may also have contributed to the increased CD-related health services observed. Using Ottawa as an example, in 1995 there was only one part-time pediatric gastroenterologist in the city; by 2005 that number increased to

five. The number of endoscopy procedures performed each year also increased between 2000 and 2014 from approximately 115 to 1,200 per year. The growth in the number of pediatric gastroenterologists improved the capacity for disease detection, thus increased the volume of health services provided for patients suspected of CD.

4.5 Study Limitations

In the current study, the algorithms were derived and tested within Ottawa but not in any other regions in the province of Ontario. Therefore, the algorithms may not perform well in the Ontario cohort if there is a difference in health service utilization pattern from Ottawa patients. In addition, the derivation time frame of the Ottawa cohort was 2005 to 2011, while the target time frame for the CD incidence estimation was from 1995 to 2011. Therefore, we were not able to test the performance of the algorithms in the decade prior to the derivation time period (1995 to 2004). Furthermore, the CIHI changed their coding practice from ICD-9 to ICD-10 in 2002, prior to our derivation time frame. In figure 3.2, the increase in incidence rates as estimated by Group 2 and Group 3 algorithms appeared to be sharper from 2005 onwards. This could potentially be caused by the "over-fitting" of the algorithms to the derivation time period (2005 to 2011).

In an editorial, Kaplan, G. warned of the danger of applying the same algorithm over a long period of time, and that the performance of the algorithm may change over time.(73) For reasons of feasibility, our derivation time period only spanned 7 years (2005 to 2011); therefore, we did not evaluate the performance of the algorithms over time. However, during the study time period (1995 to 2011), there was several notable changes to the diagnostic and care process of CD that could potentially influence the algorithms` performance over time. They included: 1) the availability of the new screening tests (EMA and tTG) (5); 2) increased supply of pediatric gastroenterologists; 3) the increased screening of at-risk individuals (e.g. first relatives of CD patients, patients with diabetes, Down and Turner syndromes, etc.) (74); 4) the increased awareness of CD and the popularization of gluten sensitivity and the gluten-free diet (75, 76). Therefore, even if we were able to obtain an algorithm with optimal performance, we would still

need to validate the algorithm at various time points during the study period and various geographic regions to ensure that it works well to identify patients with biopsy-proven CD.

This purpose of this study is not to close the door to future derivation and validation work. Other databases may contain better or more reliable information facilitating the ability to identify CD patients. For example, we suspect that we may have been better able to identify patients using the administrative data with the addition of certain data elements such as supplemental clinical information (i.e. linked from electronic health record data), a pathology database with searchable records, or laboratory testing results (i.e. from tTG or EMA screening tests). Perhaps in the future these elements will allow for the surveillance of CD patients using Ontario health administrative data. However, at the present time the algorithms were not of sufficient accuracy to produce reliable research.

4.6 Conclusions

In summary, we were unable to derive an algorithm that could accurately identify children with biopsy-proven CD from within Ontario health administrative database using the data available to us at present. The Group 1 and Group 2 algorithms suffered mainly from low PPVs. As a result, we were not able to discriminate true cases of CD from patients who have undergone endoscopy but did not have the disease. While the Group 3 algorithms demonstrated excellent PPVs with some of them achieving moderate sensitivity, we are skeptical of their generalizability outside of the Ottawa cohort. This group of algorithms could only capture endoscopy procedures that were performed in the hospital, which may not reflect the practice of all Ontario gastroenterologists.

Given that the algorithms are composed of health services related to CD, we observed a significant increase in CD-related health services in Ontario during the study period. Specifically, we saw an average annual increase of 9% in the number of children who received an endoscopy procedure followed by one or more CD related outpatient encounters with

gastroenterologists after the endoscopy. We are uncertain of whether this is the result of better disease awareness and/or detection. At any rate, as more children are seeking health services related to CD or gluten sensitivity, it is important for clinicians to adhere to the diagnostic guidelines for CD, such as the one published by the NASGHAN(1), such that suspected patients could be correctly diagnosed and that a gluten-free diet could be initiated in a timely manner.

This study demonstrates the limitations of using health administrative data (as it currently exists in Ontario) to study chronic diseases. Not all diseases can be accurately identified using health administrative data codes. Our study emphasizes the importance of evaluating the codes/algorithms involved with identifying populations derived from health administrative data, thus improving the reliability of the results generated from research using such data.

As health administrative data become more accessible for research purposes, one could anticipate that the number of studies that utilize health administrative data will continue to grow. As a relatively new source of research data, there is still much to learn about appropriate utilization of these data for research. This study could serve as a caution for researchers who are interested in studying chronic diseases using these data, in that the validity of their study findings are limited by the quality of the identification algorithm, as well as the codes and algorithms used to ascertain exposures/outcomes. Health administrative data, much like any type of data used for epidemiological research, has limitations. However, with proper data quality checks such as algorithms validation, these data can be extremely valuable resources for the advancement of our medical knowledge.

References

1. Hill ID, Dirks MH, Liptak GS, Colletti RB, Fasano A, Guandalini S, et al. Guideline for the diagnosis and treatment of celiac disease in children: Recommendations of the North American Society for Pediatric Gastroenterology, Hepatology and Nutrition. J Pediatr Gastroenterol Nutr. 2005;40(1):1-19.

2. Rubio-Tapia A, Hill ID, Kelly CP, Calderwood AH, Murray JA. ACG clinical guidelines: Diagnosis and management of celiac disease. Am J Gastroenterol. 2013;108(5):656-76.

3. Cilleruelo ML, Roman-Riechmann E, Sanchez-Valverde F, Donat E, Manuel-Ramos J, Martín-Orte E, et al. Spanish National Registry of Celiac Disease: Incidence and clinical presentation. J Pediatr Gastroenterol Nutr. 2014;59(4):522-6.

4. Kwon JH, Farrell RJ. Recent advances in the understanding of celiac disease: Therapeutic implications for the management of pediatric patients. Pediatric Drugs. 2006;8(6):375-88.

5. McGowan KE, Castiglione DA, Butzner JD. The changing face of childhood celiac disease in north america: impact of serological testing. Pediatrics. 2009 Dec;124(6):1572-8.

6. White LE, Bannerman E, McGrogan P, Kastner-Cole D, Carnegie E, Gillett PM. Childhood coeliac disease diagnoses in Scotland 2009-2010: the SPSU project. Arch Dis Child. 2013 Jan;98(1):52-6.

7. Roma E, Panayiotou J, Karantana H, Constantinidou C, Siakavellas SI, Krini M, et al. Changing pattern in the clinical presentation of pediatric celiac disease: A 30-year study. Digestion. 2009;80(3):185-91.

8. Rubio-Tapia A, Kyle RA, Kaplan EL, Johnson DR, Page W, Erdtmann F, et al. Increased Prevalence and Mortality in Undiagnosed Celiac Disease. Gastroenterology. 2009;137(1):88-93.

9. Rashid M, Cranney A, Zarkadas M, Graham ID, Switzer C, Case S, et al. Celiac disease: Evaluation of the diagnosis and dietary compliance in Canadian children. Pediatrics. 2005;116(6):e754-9.

10. Boersma B, Houwen RHJ, Blum WF, Van Doorn J, Wit JM. Catch-up growth and endocrine changes in childhood celiac disease: Endocrine changes during catch-up growth. Horm Res. 2002;58(SUPPL. 1):57-65.

11. Mora S, Barera G, Ricotti A, Weber G, Bianchi C, Chiumello G. Reversal of low bone density with a gluten-free diet in children and adolescents with celiac disease. Am J Clin Nutr. 1998;67(3):477-81.

12. Grainge MJ, West J, Solaymani-Dodaran M, Card TR, Logan RF. The long-term risk of malignancy following a diagnosis of coeliac disease or dermatitis herpetiformis: a cohort study. Aliment Pharmacol Ther. 2012 Mar;35(6):730-9.

13. Silano M, Volta U, Vincenzi AD, Dessi M, Vincenzi MD, Collaborating Centers of the Italian Registry of the Complications of Coeliac Disease. Effect of a gluten-free diet on the risk of enteropathy-associated T-cell lymphoma in celiac disease. Dig Dis Sci. 2008 Apr;53(4):972-6.

14. Askling J, Linet M, Gridley G, Halstensen TS, Ekström K, Ekbom A. Cancer incidence in a population-based cohort of individuals hospitalized with celiac disease or dermatitis herpetiformis. Gastroenterology. 2002;123(5):1428-35.

15. Logan RFA, Rifkind EA, Turner ID, Ferguson A. Mortality in celiac disease. Gastroenterology. 1989;97(2):265-71.

16. Van De Wal Y, Kooy Y, Van Veelen P, Peña S, Mearin L, Papadopoulos G, et al. Cutting edge: Selective deamidation by tissue transglutaminase strongly enhances gliadin-specific T cell reactivity. J Immunol. 1998;161(4):1585-8.

17. Molberg Ø, Mcadam SN, Körner R, Quarsten H, Kristiansen C, Madsen L, et al. Tissue transglutaminase selectively modifies gliadin peptides that are recognized by gut-derived T cells in celiac disease. Nat Med. 1998;4(6):713-7.

18. Stankovic B, Radlovic N, Lekovic Z, Ristic D, Radlovic V, Nikcevic G, et al. HLA genotyping in pediatric celiac disease patients. Bosn j basic med sci. 2014;14(3):171-6.

19. Cerf-Bensussan N, Cellier C, Heyman M, Brousse N, Schmitz J. Coeliac disease: An update on facts and questions based on the 10th International Symposium on Coeliac Disease. J Pediatr Gastroenterol Nutr. 2003;37(4):412-21.

20. Dydensborg S, Toftedal P, Biaggi M, Lillevang ST, Hansen DG, Husby S. Increasing prevalence of coeliac disease in Denmark: a linkage study combining national registries. Acta Paediatr. 2012 Feb;101(2):179-84.

21. Greco L, Romino R, Coto I, Di Cosmo N, Percopo S, Maglio M, et al. The first large population based twin study of coeliac disease. Gut. 2002;50(5):624-8.

22. Dubé C, Rostom A, Sy R, Cranney A, Saloojee N, Garritty C, et al. The prevalence of celiac disease in average-risk and at-risk Western European populations: A systematic review. Gastroenterology. 2005;128(4 SUPPL. 1):S57-67.

23. Karell K, Louka AS, Moodie SJ, Ascher H, Clot F, Greco L, et al. HLA types in celiac disease patients not carrying the DQA1 *05-DQB1 *02 (DQ2) heterodimer: Results from the European genetics cluster on celiac disease. Hum Immunol. 2003;64(4):469-77.

24. Green PHR, Cellier C. Medical progress: Celiac disease. New Engl J Med. 2007;357(17):1731-43.

25. Størdal K, Bakken IJ, Surén P, Stene LC. Epidemiology of coeliac disease and comorbidity in Norwegian children. J Pediatr Gastroenterol Nutr. 2013;57(4):467-71.

26. Ivarsson A, Persson LÅ, Nyström L, Ascher H, Cavell B, Danielsson L, et al. Epidemic of coeliac disease in Swedish children. Acta Paediatr Int J Paediatr. 2000;89(2):165-71.

27. Namatovu F, Stromgren M, Ivarsson A, Lindgren U, Olsson C, Lindkvist M, et al. Neighborhood conditions and celiac disease risk among children in Sweden. Scand J Public Health. 2014 Nov;42(7):572-80.

28. Szajewska H, Chmielewska A, Piescik-Lech M, Ivarsson A, Kolacek S, Koletzko S, et al. Systematic review: Early infant feeding and the prevention of coeliac disease. Aliment Pharmacol Ther. 2012;36(7):607-18.

29. Vriezinga SL, Auricchio R, Bravi E, Castillejo G, Chmielewska A, Crespo Escobar P, et al. Randomized feeding intervention in infants at high risk for celiac disease. N Engl J Med. 2014 Oct 2;371(14):1304-15.

30. Health Quality O. Clinical utility of serologic testing for celiac disease in ontario: an evidence-based analysis. Ont Health Technol Assess Ser. 2010;10(21):1-111.

31. Rostom A, Murray JA, Kagnoff MF. American Gastroenterological Association (AGA) Institute Technical Review on the Diagnosis and Management of Celiac Disease. Gastroenterology. 2006;131(6):1981-2002.

32. Hill ID. What are the sensitivity and specificity of serologic tests for celiac disease? Do sensitivity and specificity vary in different populations? Gastroenterology. 2005;128(4 SUPPL. 1):S25-32.

33. Telega G, Bennet TR, Werlin S. Emerging new clinical patterns in the presentation of celiac disease. Archives of Pediatrics and Adolescent Medicine. 2008;162(2):164-8.

34. Cataldo F, Marino V, Bottaro G, Greco P, Ventura A. Celiac disease and selective immunoglobulin A deficiency. J Pediatr. 1997;131(2):306-8.

35. Cataldo F, Marino V, Ventura A, Bottaro G, Corazza GR. Prevalence and clinical features of selective immunoglobulin A deficiency in coeliac disease: An Italian multicentre study. Gut. 1998;42(3):362-5.

36. Husby S, Koletzko S, Korponay-Szabó IR, Mearin ML, Phillips A, Shamir R, et al. European society for pediatric gastroenterology, hepatology, and nutrition guidelines for the diagnosis of coeliac disease. J Pediatr Gastroenterol Nutr. 2012;54(1):136-60.

37. Marsh MN. Gluten, major histocompatibility complex, and the small intestine - A molecular and immunobiologic approach to the spectrum of gluten sensitivity ('celiac sprue'). Gastroenterology. 1992;102(1):330-54.

38. Harewood GC, Holub JL, Lieberman DA. Variation in small bowel biopsy performance among diverse endoscopy settings: Results from a national endoscopic database. Am J Gastroenterol. 2004;99(9):1790-4.

39. Wakim-Fleming J, Pagadala MR, Lemyre MS, Lopez R, Kumaravel A, Carey WD, et al. Diagnosis of Celiac Disease in Adults Based on Serology Test Results, Without Small-Bowel Biopsy. Clin Gastroenterol Hepatol. 2013;11(5):511-6.

40. Marsh MN. Gluten, major histocompatibility complex, and the small intestine - A molecular and immunobiologic approach to the spectrum of gluten sensitivity ('celiac sprue'). Gastroenterology. 1992;102(1):330-54.

41. Rostom A, Dubé C, Cranney A, Saloojee N, Sy R, Garritty C, et al. Celiac disease. Evid Rep Technol Assess (Summ). 2004(104):1-6.

42. Makovicky P, Rimarova K, Boor A, Makovicky P, Vodicka P, Samasca G, et al. Correlation between antibodies and histology in celiac disease: Incidence of celiac disease is higher than expected in the pediatric population. Molecular Medicine Reports. 2013;8(4):1079-83.

43. Ress K, Luts K, Rägo T, Pisarev H, Uibo O. Nationwide study of childhood celiac disease incidence over a 35-year period in Estonia. Eur J Pediatr. 2012;171(12):1823-8.

44. Rewers M. Epidemiology of celiac disease: What are the prevalence, incidence, and progression of celiac disease? Gastroenterology. 2005;128(4 SUPPL. 1):S47-51.

45. Hurley JJ, Lee B, Turner JK, Beale A, Jenkins HR, Swift GL. Incidence and presentation of reported coeliac disease in Cardiff and the Vale of Glamorgan: the next 10 years. Eur J Gastroenterol Hepatol. 2012 May;24(5):482-6.

46. Namatovu F, Sandstrom O, Olsson C, Lindkvist M, Ivarsson A. Celiac disease risk varies between birth cohorts, generating hypotheses about causality: evidence from 36 years of population-based follow-up. BMC Gastroenterol. 2014;14:59.

47. Myléus A, Ivarsson A, Webb C, Danielsson L, Hernell O, Högberg L, et al. Celiac disease revealed in 3% of Swedish 12-year-olds born during an epidemic. J Pediatr Gastroenterol Nutr. 2009;49(2):170-6.

48. West J, Fleming KM, Tata LJ, Card TR, Crooks CJ. Incidence and prevalence of celiac disease and dermatitis herpetiformis in the UK over two decades: population-based study. Am J Gastroenterol. 2014 May;109(5):757-68.

49. Viljamaa M, Collin P, Huhtala H, Sievänen H, Mäki M, Kaukinen K. Is coeliac disease screening in risk groups justified? A fourteen-year follow-up with special focus on compliance and quality of life. Alimentary Pharmacology and Therapeutics. 2005;22(4):317-24.

50. Vilppula A, Kaukinen K, Luostarinen L, Krekelä I, Patrikainen H, Valve R, et al. Clinical benefit of gluten-free diet in screen-detected older celiac disease patients. BMC Gastroenterology. 2011;11.

51. Benchimol EI, Manuel DG, To T, Griffiths AM, Rabeneck L, Guttmann A. Development and use of reporting guidelines for assessing the quality of validation studies of health administrative data. J Clin Epidemiol. 2011;64(8):821-9.

52. Benchimol EI, Guttmann A, Mack DR, Nguyen GC, Marshall JK, Gregor JC, et al. Validation of international algorithms to identify adults with inflammatory bowel disease in health administrative data from Ontario, Canada. J Clin Epidemiol. 2014;67(8):887-96.

53. Ducharme R, Benchimol EI, Deeks SL, Hawken S, Fergusson DA, Wilson K. Validation of diagnostic codes for intussusception and quantification of childhood intussusception incidence in Ontario, Canada: A population-based study. J Pediatr. 2013;163(4):1073,1079.e3.

54. Hux JE, Ivis F, Flintoft V, Bica A. Diabetes in Ontario: Determination of prevalence and incidence using a validated administrative data algorithm. Diabetes Care. 2002;25(3):512-6.

55. Cohen D, Manuel DG, Tugwell P, Sanmartin C, Ramsay T. Direct healthcare costs of acute myocardial infarction in Canada's elderly across the continuum of care. Journal of the Economics of Ageing. 2014;3:44-9.

56. Mullen K-, Coyle D, Manuel D, Nguyen HV, Pham B, Pipe AL, et al. Economic evaluation of a hospital-initiated intervention for smokers with chronic disease, in Ontario, Canada. Tob Control. 2014.

57. Manuel DG, Rosella LC, Stukel TA. Importance of accurately identifying disease in studies using electronic health records. BMJ. 2010;341.

58. Benchimol EI, Manuel DG, To T, Griffiths AM, Rabeneck L, Guttmann A. Development and use of reporting guidelines for assessing the quality of validation studies of health administrative data. J Clin Epidemiol. 2011;64(8):821-9.

59. Shiff NJ, Jama S, Boden C, Lix LM. Validation of administrative health data for the pediatric population: A scoping review. BMC Health Services Research. 2014;14(1).

60. Dolan D, Grainger J, MacCallum N, Creatura D, Forrester J, Shiller S. The Institute for Clinical Evaluative Sciences: 20 years and counting. Healthcare quarterly (Toronto, Ont.). 2012;15(4):19-21.

61. ICES Data Dictionary [Internet].; 2015. Available from: https://datadictionary.ices.on.ca/Applications/DataDictionary/Default.aspx.

62. Discharge Abstract Database (DAD) Metadata [Internet].; 2015. Available from: https://www.cihi.ca/en/types-of-care/hospital-care/acute-care/dad-metadata.

63. Newcombe RG. Two-sided confidence intervals for the single proportion: Comparison of seven methods. Stat Med. 1998;17(8):857-72.

64. Benchimol EI, Guttmann A, Griffiths AM, Rabeneck L, Mack DR, Brill H, et al. Increasing incidence of paediatric inflammatory bowel disease in Ontario, Canada: Evidence from health administrative data. Gut. 2009;58(11):1490-7.

65. Fay MP, Feuer EJ. Confidence intervals for directly standardized rates: A method based on the gamma distribution. Stat Med. 1997;16(7):791-801.

66. Benchimol EI. Epidemiology and health administrative data: Focus on methodology and transparency. Inflamm Bowel Dis. 2014;20(10):1780-1.

67. White LE, Merrick VM, Bannerman E, Russell RK, Basude D, Henderson P, et al. The rising incidence of celiac disease in Scotland. Pediatrics. 2013;132(4):e924-31.

68. Langan SM, Benchimol EI, Guttmann A, Moher D, Petersen I, Smeeth L, et al. Setting the RECORD straight: Developing a guideline for the reporting of studies conducted using observational routinely collected data. Clin Epidemiol. 2013;5(1):29-31.

69. Mattila E, Kurppa K, Ukkola A, Collin P, Huhtala H, Forma L, et al. Burden of illness and use of health care services before and after celiac disease diagnosis in children. J Pediatr Gastroenterol Nutr. 2013;57(1):53-6.

70. Couris CM, Colin C, Rabilloud M, Schott AM, Ecochard R. Method of correction to assess the number of hospitalized incident breast cancer cases based on claims databases. J Clin Epidemiol. 2002;55(4):386-91.

71. Ravikumara M, Nootigattu VKT, Sandhu BK. Ninety percent of celiac disease is being missed. J Pediatr Gastroenterol Nutr. 2007;45(4):497-9.

72. Mansueto P, Seidita A, D'Alcamo A, Carroccio A. Non-celiac gluten sensitivity: literature review. J Am Coll Nutr. 2014;33(1):39-54.

73. Kaplan GG. Pitfalls and perils of using administrative databases to evaluate the incidence of inflammatory bowel disease overtime. Inflamm Bowel Dis. 2014;20(10):1777-9.

74. Burger JP, Roovers EA, Drenth JP, Meijer JW, Wahab PJ. Rising incidence of celiac disease in the Netherlands; an analysis of temporal trends from 1995 to 2010. Scand J Gastroenterol. 2014 Aug;49(8):933-41.

75. Carroccio A, Mansueto P, Iacono G, Soresi M, D'Alcamo A, Cavataio F, et al. Non-celiac wheat sensitivity diagnosed by double-blind placebo-controlled challenge: Exploring a new clinical entity. Am J Gastroenterol. 2012;107(12):1898-906.

76. Digiacomo DV, Tennyson CA, Green PH, Demmer RT. Prevalence of gluten-free diet adherence among individuals without celiac disease in the USA: Results from the continuous national health and nutrition examination survey 2009-2010. Scand J Gastroenterol. 2013;48(8):921-5.

Appendix I: CHEO chart review data collection form

Chart Review Data Collection Form

MRN: _____
OHIP: _____
Date of Birth (DOB) (YYYY/MM/DD):_____

1) Has the patient been excluded?

☐ Yes

☐ No

If Yes to Q.1- Please provide reason for exclusion:

_____
_____
_____

2) Diagnosis of Celiac Disease:

☐ Yes

☐ No

3) Scope Status:

☐ Yes

☐ No

4) Date of diagnosis/scope (YYYY/MM/DD): _____

5) Marsh Classification:

☐ 1

☐ 2

☐ 3a

☐ 3b

☐ 3c

6) IgA TTG Level (Ref: <20 units): _____

7) Serum IgA Level (If IgA TTG<20 units): _____