

**Dyslipidemia in Newfoundland: Findings from Canadian Primary Care
Sentinel Surveillance Network in Newfoundland and Labrador**

By

Justin D. Oake

A thesis submitted to the School of Graduate Studies in partial fulfillment of the
requirements for the degree of Master of Science in Medicine

Clinical Epidemiology Program, Faculty of Medicine, Memorial University of
Newfoundland

St. John's, NL

May 2019

Abstract

Newfoundland and Labrador (NL) has a higher level of cardiovascular disease (CVD) mortality than any other Canadian province. One factor which may explain this trend is the lipid profile pattern in this province. Given the limited lipid profile data which has been reported from NL, we organized three studies in this thesis to describe the lipid profile of Newfoundlanders. The first study was a secondary analysis of Canadian Primary Care Sentinel Surveillance Network (CPCSSN) data to document single and mixed dyslipidemia in NL. The second study compared lipid profiles and the prevalence of dyslipidemia between NL CPCSSN data and the Canadian Health Measures Survey (CHMS). The third study used electronic medical record (EMR) data in assessing the validity of ICD codes for identifying patients with dyslipidemia. This was a secondary analysis of EMR data in NL. Most recent lipid profile scores, co-morbidities, and demographic information were extracted from the CPCSSN database. We demonstrated that single and mixed dyslipidemia are quite prevalent in the NL population. Unhealthy levels of HDL were also more prevalent in NL men, compared to the Canadian sample. Of importance, the use of the ICD coding, either alone or in combination with laboratory data or lipid-lowering medication records, was an inaccurate indicator in identifying dyslipidemia.

Acknowledgements

This study was funded by the Newfoundland and Labrador Centre for Applied Health Research. This work would not be completed without the assistance of many individuals. I would like to take this opportunity to thank all those who helped me along the way.

First and foremost, I would like to thank my supervisor Dr. Shabnam Asghari. Shabnam has stuck with me all these years and I know I wouldn't have completed this without her extraordinary help. I would also like to thank my supervisory committee members Dr. Marshall Godwin, Dr. Kris Aubrey, Dr. Kayla Collins, and Dr. Pauline Duke for their invaluable input, support, and assistance. I would also be remiss if I didn't thank all the members of the PHRU.

I thank my family and friends who have supported me throughout this process. Lastly, I thank my wife, Megan for her undivided support and encouragement during this time. Without her motivation, this work would not have been possible.

Table of Contents

Abstract.....	ii
Acknowledgements	iii
List of Tables	viii
List of Figures.....	x
List of Abbreviations	xi
Chapter 1: Introduction and Overview	1
1.1 OVERVIEW OF THESIS	1
1.2 CO-AUTHORSHIP STATEMENT	2
1.2.1 Design and Identification of the Research Proposal:	2
1.2.2 Practical Aspects of the Research:	3
1.2.3 Data Analysis:	3
1.2.4 Manuscript Preparation:	3
1.2.5 References:	3
Chapter 2: Literature Review	4
2.1 OVERVIEW OF THE LITERATURE REVIEW	4
2.2 LITERATURE REVIEW STRATEGY	4
2.3 DYSLIPIDEMIA AND LIPID PROFILES.....	5
2.3.1 Epidemiology of Dyslipidemia (Internationally, Nationally, Newfoundland and Labrador)	7
2.4 CARDIOVASCULAR DISEASE	10
2.5 ELECTRONIC HEALTH DATABASES	11
2.5.1 Advantages vs. Disadvantages	12

2.5.2 Electronic Health Databases for Epidemiologic Investigation and Prevalence	
Estimation.....	14
2.6 LITERATURE REVIEW SUMMARY AND THESIS RATIONALE	15
2.7 TESTABLE QUESTIONS	16
2.8 REFERENCES	16
Chapter 3: Prevalence of Single and Mixed Dyslipidemia in Adults in Primary Care	
Settings: Findings from Canadian Primary Care Sentinel Surveillance Network	
Newfoundland and Labrador Region	27
3.1 ABSTRACT	27
3.2 INTRODUCTION	28
3.3 METHODS	30
3.3.1 Study design	30
3.3.2 Data source	30
3.3.3 Variables.....	30
3.3.4 Study population.....	32
3.3.5 Statistical analysis	32
3.3.6 Ethics	32
3.4 RESULTS	32
3.4.1 Population characteristics	32
3.4.2 Prevalence of single and mixed dyslipidemia in patients of EMR primary care setting in NL.....	35
3.4.3 Factors associated with single and mixed dyslipidemia.....	36
3.4.4 Comparison of single and mixed dyslipidemia between lipid-lowering medication users and non-medication users.....	40
3.5 DISCUSSION.....	42

3.6 CONCLUSION.....	45
3.7 REFERENCES	45
Chapter 4: Comparison of Canadian Primary Care Sentinel Surveillance Network Database and Canadian Health Measures Survey on the Lipid Profile and Prevalence of Dyslipidemia in Newfoundland.....	
4.1 ABSTRACT	51
4.2 INTRODUCTION	52
4.3 METHODS	53
4.3.1 Study Design	53
4.3.2 Data Sources	53
4.3.3 Study Population	55
4.3.4 Lipid Profiles	55
4.3.5 Variables.....	55
4.3.6 Statistical Analysis	56
4.3.7 Ethics	57
4.4 RESULTS	57
4.4.1 Lipid Profile of Newfoundlanders.....	57
4.4.2 NL Mean Cholesterol Levels Compared to the Rest of Canada	63
4.4.3 NL Prevalence of Dyslipidemia Compared to the Rest of Canada	66
4.5 DISCUSSION.....	71
4.6 CONCLUSION.....	73
4.7 REFERENCES:	74
Chapter 5: Using Electronic Medical Records to Identify Patients with Dyslipidemia in Primary Care Settings: International Classification of Disease Code Matters from One Region to a National Database	
	78

5.1 ABSTRACT	78
5.2 INTRODUCTION	79
5.3 METHODS	81
5.3.1 Study Design	81
5.3.2 Study Database	82
5.3.3 Study Population	82
5.3.4 Algorithm Development and Evaluation.....	83
5.3.5 Statistical Analysis	88
5.3.6 Ethics	89
5.4 RESULTS	89
5.5 DISCUSSION	93
5.6 CONCLUSION.....	96
5.7 REFERENCES	96
Chapter 6: Summary	102
6.1 REFERENCES	104
Chapter 7: Overall Thesis References (Alphabetical)	107

List of Tables

Table 3.1. Characteristics of the study population of primary care patients in NL (n=4,382)	34
Table 3.2. The prevalence of cardiovascular risk factors with various combinations of single and mixed dyslipidemia in patients of EMR primary care settings in NL (n=4382)	38
Table 3.3. Factors which influence single and mixed dyslipidemia in patients of EMR primary care settings in NL (n=4382)	39
Table 4.1. Lipid Profile of Patients of Canadian Primary Care Sentinel Surveillance Network in NL (n=3893).....	59
Table 4.2. Lipid Profile of Patients of Canadian Primary Care Sentinel Surveillance Network in NL by Medication Use (n=3893)	61
Table 4.3. Lipid Profile and Some Associated Factors in NL Canadian Primary Care Sentinel Surveillance Patients (n=3893).....	63
Table 4.4. Lipid profile by Sex and Age, Comparison between NL sample (n=3893) and Canadian Sample (n=3508)	65
Table 5.1. Healthy Levels of Serum Lipids for Canadian Adults [18].....	84
Table 5.2. Number of Patients with Dyslipidemia and Associated Prevalence Categorized by Algorithm.....	87
Table 5.3. Definitions for sensitivity, specificity, NPV, and PPV	89

Table 5.4. Sensitivity, Specificity, and Predictive Values of all Combinations of Situations A, B, and C as Compared with Comprehensive Criterion Standard	91
Table 5.5. Factors Associated with ICD Coding for Dyslipidemia.....	93

List of Figures

Figure 3-1. Prevalence of single and mixed dyslipidemia of LDL, HDL, and TG in patients of EMR primary care settings in NL (n=4382).	36
Figure 3-2. Single and mixed dyslipidemia in lipid-lowering medication users and non-medication users in patients of EMR primary care settings in NL who had lipid test during the study period (n=4382).	41
Figure 4-1. Percentage of adults with hypertriglyceridemia by age group. Comparison between NL study (n=3983) and CHMS (n=3508).	67
Figure 4-2. Percentage of adults with LDL dyslipidemia by age group. Comparison between NL study (n=3983) and CHMS (n=3508).	68
Figure 4-3. Percentage of adults with HDL dyslipidemia by age group. Comparison between NL study (n=3983) and CHMS (n=3508).	69
Figure 4-4. Percentage of adults with hypercholesterolemia by age group. Comparison between NL study (n=3983) and CHMS (n=3508).	70
Figure 5-1. Venn diagram of the three components of the criterion standard algorithm (n=4382).	90

List of Abbreviations

Area under the receiver operating characteristic curve (AUC)

Body mass index (BMI)

Canadian Health Measures Survey (CHMS)

Canadian Primary Care Sentinel Surveillance Network (CPCSSN)

Cardiovascular disease (CVD)

Centers for Disease Control (CDC)

Electronic Medical Records (EMR)

High density lipoprotein (HDL)

International Classification of Disease (ICD)

Low density lipoprotein (LDL)

Medical Subject Headings (MeSH)

Myocardial infarction (MI)

National Health and Nutrition Examination Surveys (NHANES)

Negative predictive value (NPV)

Newfoundland and Labrador (NL)

Positive predictive value (PPV)

Receiver operating characteristic (ROC)

Retinol-binding protein 4 (RBP4)

Triglycerides (TG)

Chapter 1: Introduction and Overview

1.1 OVERVIEW OF THESIS

This chapter provides background information on the purpose and objectives of the study, the research problem, and the rationale for this study. It also outlines the organization of the dissertation.

In Chapter 2 a review of the literature is given. The existing literature on dyslipidemia and CVD, including its epidemiology, etiology, clinical features and risk factors, as well as the use of EMRs in medical research was reviewed. We first review the definition of dyslipidemia and a lipid profile. We also discuss the epidemiology of dyslipidemia locally, nationally, and internationally. We then review the mortality and morbidity, as well as the risk factors associated with CVD. The review then shifts to a focus of how EMRs are increasingly being utilized in medical research. The literature review concludes with posing the hypotheses and testable questions addressed in the three studies in this thesis.”

Chapter 3 describes the prevalence of various single and mixed dyslipidemias in a primary care setting in NL using data collected from the NL subset of the CPCSSN database. In Chapter 4 the most recent lipid profile scores extracted from the NL subset of the CPCSSN database were compared to Canadian lipid profile data from the CHMS. Chapter 5 reports the investigation of the validity of EMR data using ICD code data for identifying patients with dyslipidemia. Multiple testing algorithms were also developed to best identify patients with dyslipidemia from EMR data.

In Chapter 6 the responses to the research hypothesis and testable questions to the three studies were answered. In addition, a brief comparison and integration is made for the results from the three studies. Chapters 3, 4, and 5 in the thesis each has its own unique list of references and in Chapter 7 an alphabetical list of all references in the thesis is given.

1.2 CO-AUTHORSHIP STATEMENT

This thesis resulted in one published manuscript. Chapter 4: Comparison of Canadian Primary Care Sentinel Surveillance Network Database and Canadian Health Measures Survey on the Lipid Profile and Prevalence of Dyslipidemia in Newfoundland was published in the journal of Biomedical Informatics Insights in 2017 [1]. Justin Oake conceptualized the study, performed data analysis and wrote the early draft of the manuscript. Shabnam Asghari contributed to data analysis plan, interpretation of the results, reviewed and edited the manuscript.

1.2.1 Design and Identification of the Research Proposal:

The design and identification of my three-part research proposal was a collaborative effort between Dr. Shabnam Asghari and myself. Between October 2010 and November 2010, Dr. Asghari and I discussed ideas of possible areas of examination for my research proposal.

1.2.2 Practical Aspects of the Research:

For the first part of my research, the application to use existing health data, CPCSSN data, was completed by Dr. Asghari and myself. Data was stored in the Primary Healthcare Research Unit (PHRU) databases and all staff at the PHRU signed oaths of confidentiality. All data analyses were conducted in the PHRU.

1.2.3 Data Analysis:

The data analysis for each of the three manuscripts was a collaborative effort between Dr. Asghari and myself. For the published paper all coauthors reviewed and provided feedback.

1.2.4 Manuscript Preparation:

For the three studies in my thesis, I prepared first drafts of each of the manuscripts and I also prepared the first draft of the literature review. Members of my supervisory committee provided the necessary feedback and corrections, as they were needed for all sections of my thesis writing.

1.2.5 References:

1. Oake J, Aref-Eshghi E, Godwin M, et al. Using electronic medical record to identify patients with dyslipidemia in primary care settings: International classification of disease code matters from one region to a national database. *Biomed Inform Insights*. 2017;9:1178222616685880.

Chapter 2: Literature Review

2.1 OVERVIEW OF THE LITERATURE REVIEW

This literature review is divided into four sections. In *Section 2.1*, the outline of the chapter is given. *Section 2.2* outlines the literature review strategy. Next, *Section 2.3* reviews the definition of dyslipidemia and a lipid profile. We also discuss the epidemiology of dyslipidemia locally, nationally, and internationally. This is followed by *Section 2.4* where we review the mortality and morbidity, as well as the risk factors associated with CVD. In *Section 2.5*, the review shifts to a focus of how EMRs are increasingly being utilized in medical research. The literature review concludes with the *2.6 Literature Review Summary and Thesis Rationale*, and *2.7 Testable Questions* addressed in the three studies in this thesis.

2.2 LITERATURE REVIEW STRATEGY

To identify relevant articles about dyslipidemia and electronic health records, a literature search was performed for relevant articles published between 1970-2018. This search used electronic databases including PUBMED, EMBASE, COCHRANE and CINAHL. The initial search comprised the following MeSH terms: Dyslipidemias, Cardiovascular Diseases, Lipoproteins, Newfoundland and Labrador, Electronic Health Records, International Classification of Diseases, and Clinical Laboratory Information Systems. The title and abstract of each paper retrieved from the initial literature search was analyzed to identify potentially eligible studies. Reference lists of relevant papers were also surveyed to identify additional literature on the topic.

2.3 DYSLIPIDEMIA AND LIPID PROFILES

Atherogenic dyslipidemia (abnormally elevated cholesterol), high levels of LDL (or “bad cholesterol”) and TG, and low levels of HDL (or “good cholesterol”) are known as important risk factors for the development of CVD [1-3]. Abnormal levels of blood cholesterol are a key component of arterial plaques which can give rise to atherosclerosis, and sometimes myocardial ischemia [1, 3].

A lipid profile is typically an estimation of the levels of total cholesterol, HDL cholesterol, LDL cholesterol, and TG. Together, these values are helpful in understanding different lipid disorders and the risk of cardiovascular complications.

Cholesterol has important functions in cell membranes, and also functions as a precursor for both bile acids and steroid hormones. Cholesterol is transported through the blood in distinct particles composed of lipids and proteins (lipoproteins) [1, 4]. Although abnormal levels of certain lipoproteins are associated with serious health consequences [2], they are essential in transportation of free fatty acids and cholesterol throughout the body [5].

LDL particles typically make up 60-70 percent of the total serum cholesterol [1]. LDL particles carry a core of cholesterol esters and small amounts of TG [6]. LDL can be internalized by the liver and other tissues for useful biologic functions, including bile acid formation, which is important in the digestion of cholesterol and fat [5]. Nonhepatic LDL cholesterol is used for steroid hormone production and cell membrane synthesis. Macrophage recruitment appears to play a significant role in atherosclerosis. These cells

accumulate large amounts of lipids through the uptake of modified lipoproteins resulting in foam cell formation. Foam cell deposition into the arterial wall under conditions of inflammation or oxidant stress, can lead to atherosclerosis and subsequent cardiovascular conditions, such as coronary artery disease, stroke, peripheral vascular disease, aneurysm formation, and sudden death [5].

LDL, the major atherogenic lipoprotein, has long been a primary target of cholesterol-lowering therapy [1, 6, 7]. Statins, or HMG CoA reductase inhibitors, are noted as being the most effective and practical class of drugs for reducing LDL levels. Evidence from trials of statin therapy consistently shows that benefits in CVD event reduction are proportional to the magnitude of LDL lowering [8, 9]. Furthermore, statins are known to modestly increase levels of HDL (up to 10%), although the precise mechanism is not known [10]. The reduction in LDL levels and the overall improvement in lipid profile associated with statin therapy reduces the risk of essentially every clinical manifestation of the atherosclerotic processes [1, 7, 10]. Numerous studies and reports have consistently concluded that a reduction in the levels of plasma cholesterol would enhance the cardiovascular health of populations such as Canada's [11-13].

HDL cholesterol normally makes up 20-30 percent of the total serum cholesterol [1]. HDL are anti-atherogenic lipoproteins [5] and are inversely correlated with risk for coronary heart disease [1, 6]. Some evidence indicates that HDL protects against the development of atherosclerotic vascular disease [1, 6], through the net effect of reverse cholesterol transport which removes excess cholesterol from cells [5]; however, a low HDL level often reflects the presence of other atherogenic factors [1]. In addition, HDL particles appear to have antioxidant, antiadhesive, anti-aggregatory, anti-inflammatory,

and profibrinolytic effects as well. All of these effects may play a role in reducing CVD [6].

2.3.1 Epidemiology of Dyslipidemia (Internationally, Nationally, Newfoundland and Labrador)

The prevalence of dyslipidemia remains an important issue worldwide [14-18], especially in developing countries [19]. Lopez-Jimenez et al. in 2009 reviewed trends in the 10-year predicted risk of CVD in the US from 1976 to 2004 [20]. This review incorporated data from three NHANES. Overall prevalence of dyslipidemia was 52.1%. The prevalence of dyslipidemia also decreased over time [20]. CDC data from 2007 to 2014 demonstrated that the percentage of adults with high total cholesterol and low HDL declined. During 2011-2014, 12.1% of adults had high total cholesterol and 18.5% had low HDL [21]. The American Heart Association also showed that the age-adjusted prevalence of elevated LDL decreased from 42.9% in 1999 to 2000 to 28.5% in 2013 to 2014 [22]. While the prevalence of some forms of dyslipidemia has decreased over recent years, during 2003 to 2012, data from NHANES revealed that the percentage of adults aged 40 and over using a cholesterol-lowering medication increased from 20% to 28%. During this same time, about 71% of adults with CVD and 54% of adults with hypercholesterolemia used a cholesterol-lowering medication [23].

The Mexican National Health and Nutrition Survey 2006 reported the 1-year prevalence of mixed hyperlipidemia, a form of dyslipidemia characterized by high TG and cholesterol levels, at 18.2%, while hypercholesterolemia was reported at 25.3% [15]. According to Janus et al., a population survey conducted in rural Australia, 2004-2006,

reported the prevalence of hypercholesterolemia to be 48%. It should be noted that this study only sampled rural participants [17]. A review by Fuentes et al. reported that the 1-year prevalence of hypercholesterolemia ranged from 39.1% in Rajasthan (an urban center in India) to 30.5% in rural Malaysia [19]. It should be noted that the rural and urban values from the Indian and Malaysian reports may not be representative of the countries as a whole.

The 1-year prevalence of dyslipidemia in Beijing, China according to the population-based Beijing Eye Study 2006, was estimated at 45.1% [18]. Li et al. also reported on the 1-year prevalence of dyslipidemia in Beijing, China. The 1-year prevalence observed in males was 51.9% and 40.8% in females [24]. Additional cross-sectional studies performed in China reported the prevalence of dyslipidemia to be approximately 36.5%, 30.3%, and 35.5% respectively [25-27]. Further, a cross-sectional study of 5,375 adult patients in Chongqing, China revealed that 44.2% had elevated TG, 14.7% had elevated total cholesterol, 13.2% had mixed dyslipidemia, and 28% had low HDL [27]. These studies reflect how dyslipidemia has become one of the most important health risk factors in the Chinese population. A recent series of studies performed in Iran also reported a high prevalence of dyslipidemia among Iranian adults [14, 28-30]. A recent report from Iran assessing metabolic syndrome also noted a high prevalence of low HDL (34.3%) and elevated TG (25%) [31]. A Spanish study demonstrated that in adults at high risk for CVD, and over 30 years of age, low HDL and high total cholesterol/HDL ratio were associated with a higher risk of all-cause mortality and hospitalization due to cardiovascular events [32].

According to Statistics Canada and the Public Health Agency of Canada, it is estimated that as many as 10 million Canadian adults have a cholesterol level higher than the recommended target [3]. The CHMS presented findings on blood cholesterol levels of adult Canadians from 2007 to 2009. This survey reported that 41% of Canadian adults had a high total cholesterol level; about 36% had unhealthy levels of LDL cholesterol, while 30% had unhealthy levels of HDL cholesterol. Additionally, about 25% of adults had unhealthy levels of TG [33]. There was also a noticeable increase of total cholesterol, LDL cholesterol, and TG level with age. About 29% of adults aged 20 to 39 had high total cholesterol, compared with 47% among those aged 40 to 59, and 54% of those aged 60 to 79. LDL cholesterol levels peaked at 43% among adults aged 40 to 59. TG levels were 17% among adults aged 20 to 39 and 34% among those aged 60 to 79 [33]. Petrella et al. conducted a retrospective cohort analysis in a Southwestern Ontario database, which comprised data from more than 150,000 adult patients in rural and urban primary care practices. Of the 49,667 patients (mean age 21 years old) who were included in the study, dyslipidemia was identified in 12% of patients [34]. The results of this study are limited, however, in that the data is not representative of the entire country.

NL has a higher level of CVD mortality than any other province in Canada [35,36]. There are several factors that may explain these trends; one is the lipid profile pattern and high prevalence of dyslipidemia in this province. Several studies have found a high prevalence of hypercholesterolemia in NL [37-39]. These studies also suggest that NL's population may exhibit a unique pattern of dyslipidemia characterized by low HDL cholesterol as compared to other Canadian provinces. For example, Chockalingam and Fodor, in a study involving several rural areas of NL, found that 61% of study volunteers

had hypercholesterolemia. However, this study involved a limited sample that is not representative of the province as a whole [38]. The unique genetic pool in NL also makes it prone to various genetic diseases. One recently discovered gene is RBP4. RBP4 is considered to be a novel adipokine. Several studies in mice have suggested that RBP4 is involved in lipid metabolism [40]. Shea et al. discussed the association of the gene variants with serum HDL cholesterol levels within the NL population. This study used a large sample of third generation Newfoundlanders and showed a significant association between the two noncoding regions of RBP4 and HDL level [40].

2.4 CARDIOVASCULAR DISEASE

CVDs are chronic, lifelong diseases caused by interactions among genetic predisposition, health behaviours, and the environment [2]. This class of diseases is caused by atherosclerosis, a hardening of the arteries, which over time leads to blood flow becoming restricted [41]. This narrowing has the potential to lead to stroke, heart attacks, or other deadly complications and are among the leading causes of death throughout the world. According to the CDC, CVD is the leading cause of death for both men and women. More than half of the deaths due to heart disease in 2015 were in men. CVD also costs the US approximately \$200 billion each year. The costs include health care services, medications, and lost productivity [42]. CVD accounts for the death of more Canadians than any other disease [36]. US data from 2014 demonstrate that approximately 630,000 Americans die from heart disease each year [42]. In the US, someone has a heart attack every 40 seconds and each minute, more than one person dies from a CVD-related event [43]. Similarly, every seven minutes in Canada, someone dies

from heart disease or stroke [36]. In 2004, 29% of the total deaths in the world and 32% in the Americas were caused by CVD [44]. CVD affects a broad range of people. In Canada, people of all ages and backgrounds are affected. Statistics from 2007 show that 1.3 million Canadians (4.8% of Canadians – 5.3% of boys and men and 4.2% of girls and women 12 years of age and older) reported having heart disease diagnosed by a health professional [2]. CVD sufferers are also more commonly of lower socioeconomic status, less educated, and of diverse races, in particular Aboriginal/Indigenous or South Asian Canadians [41].

CVD is a well-researched topic and often a primary focus in Canadian medical research. This is certainly attributable to the fact that CVD is the number one killer of Canadians [36], but is also because of the preventable nature of its risk factors [41]. Risk factors for CVD have been widely documented in multiple epidemiological studies, and include cholesterol, age, sex, overweight or obesity, cigarette smoking, diabetes mellitus, and hypertension [45,46]. Reports show that nine out of ten Canadians over the age of 20 years have at least one of those listed risk factors. Addressing these risk factors should help to decrease the incidence of not only CVD mortality, but other chronic diseases as well [2].

2.5 ELECTRONIC HEALTH DATABASES

Secondary data are data collected or generated for purposes other than research activities. The interest in using these data for health research is increasing because of their population coverage, feasibility, availability, and relatively low costs [47-49]. Examples of these data in our province are EMR data and laboratory data.

An EMR can be defined as a partial health record under the custodianship of a health care provider that holds a portion of the relevant health information about a person over their lifetime. This is often described as a provider-centric or health organization-centric health record of a person [50]. EMR data for this study is identified using an electronic chart abstraction of family physicians who are part of the NL component of the CPCSSN. CPCSSN is the first pan-Canadian multi-disease EMR surveillance system. Longitudinal data are collected from primary care practices across Canada. Data collected can be utilized to facilitate and encourage research by health care providers, contributing to a stronger national knowledge base in the area of primary health care and chronic disease management. Personal information is removed from the data prior to being collected, and further de-identified after collection to ensure it is completely anonymous. The EMRs at these clinics regularly capture the lipid values of their patients.

Laboratory data is collected through the provincial Laboratory Information System and was able to be accessed from the CPCSSN records. It includes each laboratory service used, the patient's identification, date of service, and laboratory test result. Test results are linked or integrated with individuals' EMRs, providing additional resources for diagnosing and treating patients [51].

2.5.1 Advantages vs. Disadvantages

EMR data present researchers the opportunity to obtain large quantities of clinical information. Secondary data, such as that captured in EMRs, can be useful tools for planning, monitoring, and programming across the healthcare system, and ultimately providing evidence that can be used to improve the health of a population [52,53]. EMRs

provide an ideal means to track variables over time and, thus, allow longitudinal analyses of relationships between risk factors and disease prevalence and progression [48]. EMRs can include risk factors for important health outcomes such as smoking status, clinical and laboratory measurements (cholesterol levels and BMI) [54], patient medication lists [55], and the presence of chronic conditions such as dyslipidemia, diabetes mellitus, and hypertension. The utility of EMRs has also been proven for chronic disease surveillance management and prevention [56-58].

Analyzing EMR data is a cost-effective way to test hypotheses that have not been examined before. Furthermore, these datasets have been used to cross-validate exploratory analyses and to test advanced statistical models. The accuracy of these data depends on several factors including the reliability of the codes, the condition being identified, the use of unstructured or uncoded data, high-quality data entry, data entry consistency, and completeness of the data. As the reliability of the findings from such studies depends on the accuracy of the medico-administrative billing data, studies have attempted to assess the reliability of the way in which ICD codes are used in billing data to capture diagnoses and procedures. ICD codes have been particularly shown to have restricted potentials for patients with dyslipidemia. This has been shown by the results of few previous studies available using secondary data for lipid research. Li et al. found that in a large US medical insurance claims database, only 23% of laboratory-defined patients had a dyslipidemia diagnosis [59]. Additionally, an algorithm developed by an American study reported that 62.3% of patients with dyslipidemia were not recorded by the ICD codes [60].

2.5.2 Electronic Health Databases for Epidemiologic Investigation and Prevalence Estimation

Several studies have used EMR data alone, or with other resources such as laboratory data, to help understand disease pattern, therapeutic response, and healthcare utilization [58, 61-66]. To our knowledge, at the time of this study, there was no study in our province or Canada to use these databases to assess the prevalence of dyslipidemia and describe lipid profiles.

An American study analyzed EMR data to examine the effects of BMI and obesity on the prevalence of three chronic diseases: type II diabetes, hyperlipidemia, and hypertension [67]. Overall prevalence of hyperlipidemia among patients was 21.6%. The prevalence of hyperlipidemia increased steadily with age, and was highest with patients who were diagnosed with a BMI ranging from 35 to 40. Hyperlipidemia was also markedly lower among females, 19.2%, compared to males, 25.1% [67]. According to healthcare data included in the Madrid Regional Public Health System, an EMR, the prevalence of dyslipidemia in men was 8.1%. Additional chronic health problems measured included hypertension, diabetes, allergies, and chronic obstructive pulmonary disease [68].

In addition to utilizing EMRs to understand disease patterns for dyslipidemia, these electronic health databases have also been used extensively to measure and investigate other chronic conditions. The prevalence of high-risk cardiovascular conditions among hypertensive patients receiving at least one antihypertensive medication was examined using a primary care EMR [58]. Patients in the dataset were separated into two cohorts. The first cohort were the nonelderly group and they consisted

of patients aged 18 to 64, while the second, elderly, cohort were patients aged 65 years or older. The presence of comorbid conditions were more prevalent in the elderly cohort, with the exceptions of dyslipidemia (elderly 47.32% vs nonelderly 51.81%) and diabetes (elderly 11.66% vs nonelderly 12.80%) [58]. A retrospective study by McAdam-Marx et al. identified the prevalence of resistant hypertension in patients using an EMR [61]. This EMR based study supported findings from prospective trials that resistant hypertension is an important clinical problem and that more effective management is needed to enable patients with, or at risk for, resistant hypertension to achieve blood pressure goals [61]. Tu et al. used data from primary care physician EMRs and identified patients who had a MI. Of the 969 patients, 51 had solid evidence of a documented MI [69].

2.6 LITERATURE REVIEW SUMMARY AND THESIS RATIONALE

As previously mentioned, NL has the highest level of CVD mortality of all the Canadian provinces. Dyslipidemia is one of the most modifiable risk factors for CVD; however, the prevalence and patterns of dyslipidemia in NL and the overall lipid profile of Newfoundlanders has not been well documented. Further, comparisons of our provincial lipid profile to a national cohort have also not been investigated. Given that this study was a secondary analysis of existing medico-administrative data, that these databases have been increasingly used in large-scale studies in recent years, and that the inaccuracy of the ICD codes for the purpose of diagnosis of medical conditions is frequently reported [60, 70, 71], we sought to investigate the validity of ICD coded data for identifying patients with dyslipidemia in EMR data.

Through investigating the prevalence and pattern of dyslipidemia in NL and the overall lipid profile of Newfoundlanders, this project should help policymakers and healthcare providers to better understand the magnitude of dyslipidemia as an important risk factor for CVD and other chronic diseases. Additionally, by assessing the validity of ICD codes for identifying patients with dyslipidemia and then developing testing algorithms to best identify patients with dyslipidemia, we hope to provide information which may help develop and improve upon existing government health policies to decrease the high prevalence of cardiovascular risk factors.

2.7 TESTABLE QUESTIONS

1. What is the prevalence of single and mixed dyslipidemia in primary care settings in NL?
2. How does the lipid profile and prevalence of dyslipidemia in primary care settings in NL compare to the rest of Canada?
3. Is ICD coding alone an accurate method to identify patients with dyslipidemia in an EMR in a primary care setting?

2.8 REFERENCES

1. National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). Third report of the national cholesterol education program (NCEP) expert

- panel on detection, evaluation, and treatment of high blood cholesterol in adults (adult treatment panel III) final report. *Circulation*. 2002;106(25):3143-3421.
2. Public Health Agency of Canada. Tracking heart disease and stroke in Canada. www.phac-aspc.gc.ca/publicat/2009/cvd-avc/pdf/cvd-avs-2009-eng.pdf. Updated 2009. Accessed 09/25, 2017.
 3. Heart and Stroke Foundation of Canada. <http://www.heartandstroke.com/site/c.ikIQLcMWJtE/b.3483991/k.34A8/Statistics.htm#bloodcholesterol>. Updated 2011. Accessed 09/29, 2017.
 4. Franssen R, Monajemi H, Stroes ESG, Kastelein JJP. Obesity and dyslipidemia. *Endocrinology & Metabolism Clinics of North America*. 2008;37(3):623-633.
 5. Eaton CB. Hyperlipidemia. *Prim Care*. 2005;32(4):1027-55.
 6. Tulenko TN, Sumner AE. The physiology of lipoproteins. *J Nucl Cardiol*. 2002;9(6):638-649.
 7. Grundy SM, Cleeman JJ, Merz CN, et al. Implications of recent clinical trials for the national cholesterol education program adult treatment panel III guidelines. *Circulation*. 2004;110(2):227-239.
 8. Arca M, Gasparone A. Atorvastatin efficacy in the primary and secondary prevention of cardiovascular events. *Drugs*. 2007;67 Suppl 1:29-42.
 9. Soran H, Dent R, Durrington P. Evidence-based goals in LDL-C reduction. *Clin Res Cardiol*. 2017;106(4):237-248.
 10. Lardizabal JA, Deedwania P. Lipid-lowering therapy with statins for the primary and secondary prevention of cardiovascular disease. *Cardiol Clin*. 2011;29(1):87-103.

11. Carleton RA, Dwyer J, Finberg L, et al. Report of the expert panel on population strategies for blood cholesterol reduction. A statement from the national cholesterol education program, national heart, lung, and blood institute, national institutes of health. *Circulation*. 1991;83(6):2154-2232.
12. Grover SA, Coupal L, Kaouache M, Lowensteyn I. Preventing cardiovascular disease among Canadians: What are the potential benefits of treating hypertension or dyslipidemia? *Can J Cardiol*. 2007;23(6):467-473.
13. Genest J, McPherson R, Frohlich J, et al. 2009 Canadian cardiovascular Society/Canadian guidelines for the diagnosis and treatment of dyslipidemia and prevention of cardiovascular disease in the adult - 2009 recommendations. *Can J Cardiol*. 2009;25(10):567-579.
14. Esteghamati A, Meysamie A, Khalilzadeh O, et al. Third national surveillance of risk factors of non-communicable diseases (SuRFNCD-2007) in Iran: Methods and results on prevalence of diabetes, hypertension, obesity, central obesity, and dyslipidemia. *BMC Public Health*. 2009;9:167.
15. Aguilar-Salinas CA, Gomez-Perez FJ, Rull J, Villalpando S, Barquera S, Rojas R. Prevalence of dyslipidemias in the Mexican national health and nutrition survey 2006. *Salud Publica Mex*. 2010;52 Suppl 1:S44-53.
16. Goodman SG, Langer A, Bastien NR, et al. Prevalence of dyslipidemia in statin-treated patients in Canada: Results of the DYSlipidemia international study (DYSIS). *Can J Cardiol*. 2010;26(9):e330-5.

17. Janus ED, Tideman PA, Dunbar JA, et al. Dyslipidaemia in rural Australia: Prevalence, awareness, and adherence to treatment guidelines in the greater green triangle risk factor study. *Med J Aust*. 2010;192(3):127-132.
18. Wang S, Xu L, Jonas JB, You QS, Wang YX, Yang H. Prevalence and associated factors of dyslipidemia in the adult Chinese population. *PLoS One*. 2011;6(3):e17326.
19. Fuentes R, Uusitalo T, Puska P, Tuomilehto J, Nissinen A. Blood cholesterol level and prevalence of hypercholesterolaemia in developing countries: A review of population-based studies carried out from 1979 to 2002. *Eur J Cardiovasc Prev Rehabil*. 2003;10(6):411-419.
20. Lopez-Jimenez F, Batsis JA, Roger VL, Brekke L, Ting HH, Somers VK. Trends in 10-year predicted risk of cardiovascular disease in the United States, 1976 to 2004. *Circ Cardiovasc Qual Outcomes*. 2009;2(5):443-450.
21. Total and high-density lipoprotein cholesterol in adults: United states, 2011-2014. <https://www.cdc.gov/nchs/data/databriefs/db226.htm>. Updated 2015. Accessed 10/01, 2017.
22. Benjamin EJ, Blaha MJ, Chiuve SE, et al. Heart disease and stroke statistics-2017 update: A report from the american heart association. *Circulation*. 2017;135(10):e146-e603.
23. Gu Q, Paulose-Ram R, Burt V, Kit B. Prescription cholesterol-lowering medication use in adults aged 40 and over: United states, 2003–2012. 2014.
24. Li ZY, Xu GB, Xia TA. Prevalence rate of metabolic syndrome and dyslipidemia in a large professional population in Beijing. *Atherosclerosis*. 2006;184(1):188-192.

25. Wu JY, Duan XY, Li L, et al. Dyslipidemia in Shanghai, China. *Prev Med.* 2010;51(5):412-415.
26. Zhang L, Qin LQ, Liu AP, Wang PY. Prevalence of risk factors for cardiovascular disease and their associations with diet and physical activity in suburban Beijing, China. *J Epidemiol.* 2010;20(3):237-243.
27. Qi L, Ding X, Tang W, Li Q, Mao D, Wang Y. Prevalence and risk factors associated with dyslipidemia in Chongqing, China. *Int J Environ Res Public Health.* 2015;12(10):13455-13465.
28. Fesharakinia A, Zarban A, Sharifzadeh GR. Lipid profiles and prevalence of dyslipidemia in schoolchildren in south Khorasan province, eastern Iran. *Arch Iran Med.* 2008;11(6):598-601.
29. Ghoddusi K, Ameli J, Kachuee H, Pourfarziani V, Saadat A, Karami Q. Association of diabetes mellitus and dyslipidaemias in the Tehran population. *East Mediterr Health J.* 2008;14(3):647-653.
30. Hadaegh F, Harati H, Ghanbarian A, Azizi F. Prevalence of coronary heart disease among Tehran adults: Tehran lipid and glucose study. *East Mediterr Health J.* 2009;15(1):157-166.
31. Mohammadbeigi A, Farahani H, Moshiri E, et al. Prevalence of metabolic syndrome and associations with lipid profiles in Iranian men: A population-based screening program. *World J Mens Health.* 2018;36(1):50-56.
32. Orozco-Beltran D, Gil-Guillen VF, Redon J, et al. Lipid profile, cardiovascular disease and mortality in a Mediterranean high-risk population: The ESCARVAL-RISK study. *PLoS One.* 2017;12(10):e0186196.

33. Statistics Canada. Canadian health measures survey.
http://www.statcan.gc.ca/access_acces/archive.action?loc=/pub/82-623-x/82-623-x2010002-eng.pdf&archive=1. Updated 2011. Accessed 09/28, 2017.
34. Petrella RJ, Merikle E, Jones J. Prevalence and treatment of dyslipidemia in Canadian primary care: A retrospective cohort analysis. *Clin Ther*. 2007;29(4):742-750.
35. Filate WA, Helen LJ, Courtney CK, Jack VT. Regional variations in cardiovascular mortality in Canada. *Can J Cardiol*. 2003;19(11):1241-8.
36. Statistics Canada. Mortality, summary list of causes.
<http://www.statcan.gc.ca/pub/84f0209x/84f0209x2007000-eng.htm>. Updated 2007.
Accessed 09/25, 2017.
37. Fodor JG, Balram BC, Heyden S. Lipid and lipoprotein levels of Newfoundland school children. *Ann Nutr Metab*. 1987;31(4):231-236.
38. Chockalingam A, J GF. Cardiovascular risk factors in Newfoundland population and modification of their level through nonpharmacological intervention. *J Cardiovasc Pharmacol*. 1990;16 Suppl 8:S54-6.
39. Fodor JG, Chockalingam A. Modification of cardiovascular risk factors in rural population of Newfoundland through nutritional counselling. *Cor Vasa*. 1991;33(2):107-13.
40. Shea JL, Loredó-Ostí JC, Sun G. Association of RBP4 gene variants and serum HDL cholesterol levels in the Newfoundland population. *Obesity (Silver Spring)*. 2010;18(7):1393-1397.
41. Smith ER. The Canadian heart health strategy and action plan. *Can J Cardiol*. 2009;25(8):451-452.

42. Centers for disease control and prevention, national center for health statistics.
Multiple cause of death 1999-2015 on CDC WONDER online.
<https://wonder.cdc.gov/mcd-icd10.html>. Updated 2016. Accessed 10/01, 2017.
43. Heron M. Deaths: Leading causes for 2014. 2016;65(5).
44. World Health Organization. Estimated deaths per 100,000 population by cause, and member state. 2008.
45. Yusuf S, Hawken S, Ounpuu S, et al. Obesity and the risk of myocardial infarction in 27,000 participants from 52 countries: A case-control study. *Lancet*. 2005;366(9497):1640-1649.
46. Smith DG. Epidemiology of dyslipidemia and economic burden on the healthcare system. *Am J Manag Care*. 2007;13 Suppl 3:S68-71.
47. Birtwhistle R, Keshavjee K, Lambert-Lanning A, et al. Building a pan-Canadian primary care sentinel surveillance network: Initial development and moving forward. *J Am Board Fam Med*. 2009;22(4):412-422.
48. Crawford AG, Cote C, Couto J, et al. Prevalence of obesity, type II diabetes mellitus, hyperlipidemia, and hypertension in the United States: Findings from the GE centricity electronic medical record database. *Popul Health Manag*. 2010;13(3):151-161.
49. Young WB, Ryu H. Secondary data for policy studies: Benefits and challenges. *Policy, Politics, & Nursing Practice*. 2000;1(4):302-307.
50. Canada Health Infoway. EMR, EHR, and PHR – why all the confusion?
<http://infowayconnects.infoway-inforoute.ca/blog/electronic-health-records/374-emr->

[ehr-and-phr---why-all-the-confusion/#axzz1LCQ7XyAm](#). Updated 2011. Accessed 09/28, 2017.

51. Canada Health Infoway. Laboratory information systems. <https://www.infoway-inforoute.ca/about-infoway/approach/investment-programs/laboratory-information-systems>. Updated 2011. Accessed 09/28, 2017.
52. Alaghehbandan R. *Assessing administrative databases for surveillance of depressive disorders in Newfoundland and Labrador*. Memorial University of Newfoundland, Newfoundland; 2010.
53. Henkin RE, Harolds JA. Health information technology and the electronic medical record. *Clin Nucl Med*. 2010;35(10):788-789.
54. Mitiku TF, Tu K. Using data from electronic medical records: Theory versus practice. *Healthc Q*. 2008;11(4):23-25.
55. Protti D. Comparison of information technology in general practice in 10 countries. *Healthc Q*. 2007;10(2):107-116.
56. Holt TA, Stables D, Hippisley-Cox J, O'Hanlon S, Majeed A. Identifying undiagnosed diabetes: Cross-sectional survey of 3.6 million patients' electronic records. *Br J Gen Pract*. 2008;58(548):192-196.
57. Lazarus R, Klompas M, Campion FX, et al. Electronic support for public health: Validated case finding and reporting for notifiable diseases using electronic medical data. *J Am Med Inform Assoc*. 2009;16(1):18-24.
58. Biskupiak JE, Kim J, Phatak H, Wu D. Prevalence of high-risk cardiovascular conditions and the status of hypertension management among hypertensive adults 65

- years and older in the United States: Analysis of a primary care electronic medical records database. *J Clin Hypertens*. 2010;12(12):935-944.
59. Li J, Motsko SP, Goehring EL, Jr, Tave A, Pezzullo JC, Jones JK. Prevalence of pediatric dyslipidemia: Comparison of a population-based claims database to national surveys. *Pharmacoepidemiol Drug Saf*. 2010;19(10):1031-1040.
 60. Kandula S, Zeng-Treitler Q, Chen L, Salomon WL, Bray BE. A bootstrapping algorithm to improve cohort identification using structured data. *J Biomed Inform*. 2011;44 Suppl 1:S63-8.
 61. McAdam-Marx C, Ye X, Sung JC, Brixner DI, Kahler KH. Results of a retrospective, observational pilot study using electronic medical records to assess the prevalence and characteristics of patients with resistant hypertension in an ambulatory care setting. *Clin Ther*. 2009;31(5):1116-1123.
 62. Minard JP, Turcotte SE, Loughheed MD. Asthma electronic medical records in primary care: An integrative review. *J Asthma*. 2010;47(8):895-912.
 63. Nishida Y, Takahashi Y, Nakayama T, Soma M, Kitamura N, Asai S. Effect of candesartan monotherapy on lipid metabolism in patients with hypertension: A retrospective longitudinal survey using data from electronic medical records. *Cardiovasc Diabetol*. 2010;9:38.
 64. Pettersson B, Ambegaonkar B, Sazonov V, Martinell M, Stalhammar J, Wandell P. Prevalence of lipid abnormalities before and after introduction of lipid modifying therapy among Swedish patients with dyslipidemia (PRIMULA). *BMC Public Health*. 2010;10:737.

65. Tu K, Mitiku T, Lee DS, Guo H, Tu JV. Validation of physician billing and hospitalization data to identify patients with ischemic heart disease using data from the electronic medical record administrative data linked database (EMRALD). *Can J Cardiol.* 2010;26(7):e225-8.
66. Zhang Q, Rajagopalan S, Mavros P, et al. Differences in baseline characteristics between patients prescribed sitagliptin versus exenatide based on a US electronic medical record database. *Adv Ther.* 2010;27(4):223-232.
67. Crawford AG, Cote C, Couto J, et al. Comparison of GE centricity electronic medical record database and national ambulatory medical care survey findings on the prevalence of major conditions in the united states. *Popul Health Manag.* 2010;13(3):139-150.
68. Esteban-Vasallo MD, Dominguez-Berjon MF, Astray-Mochales J, et al. Epidemiological usefulness of population-based electronic clinical records in primary care: Estimation of the prevalence of chronic diseases. *Fam Pract.* 2009;26(6):445-454.
69. Tu K, Mitiku T, Guo H, Lee DS, Tu JV. Myocardial infarction and the validation of physician billing and hospitalization data using electronic medical records. *Chronic Dis Can.* 2010;30(4):141-146.
70. Birman-Deych E, Waterman AD, Yan Y, Nilasena DS, Radford MJ, Gage BF. Accuracy of ICD-9-CM codes for identifying cardiovascular and stroke risk factors. *Med Care.* 2005;43(5):480-485.
71. O'Malley KJ, Cook KF, Price MD, Wildes KR, Hurdle JF, Ashton CM. Measuring diagnoses: ICD code accuracy. *Health Serv Res.* 2005;40(5 Pt 2):1620-1639.

Chapter 3: Prevalence of Single and Mixed Dyslipidemia in Adults in Primary Care Settings: Findings from Canadian Primary Care Sentinel Surveillance Network Newfoundland and Labrador Region

3.1 ABSTRACT

Dyslipidemia is a leading risk factor for CVD. People in NL have a higher risk of CVD mortality than any other province. Anecdotal evidence also suggests that Newfoundlanders exhibit a unique lipid profile, characterized by low levels of HDL. The primary objective of this study was to describe the prevalence of various types of single and mixed dyslipidemia in primary care settings in NL. This was a secondary analysis of EMR data in NL. The most recent lipid profiles of adults aged 20 years and older were identified from January 1, 2009 to December 31, 2010. This cross-sectional study included 4382 primary healthcare patients. 39.4% of this population had no lipid disorder, whereas 41.2%, 16.5% and 2.9% had abnormalities in one, two and three lipid components, respectively. The most common shared abnormality is between those with low HDL and elevated TG dyslipidemia, representing 7.3% of the population. Patients who were obese were more likely to have low HDL (OR, 1.98; 95% CI, 1.23-3.19), as well as abnormal levels of mixed dyslipidemia (HDL+LDL+TG) (OR, 2.34; 95% CI, 1.09-5.05). Nearly one of every five patients in primary care settings in NL has mixed dyslipidemia. These findings highlight the need for adopting strategies which will benefit early prevention and management of dyslipidemia, as well as further investigation into lipid profiles in NL.

3.2 INTRODUCTION

The components of atherogenic dyslipidemia - elevated total cholesterol, high levels of LDL, elevated TG, and low levels of HDL - are known as important risk factors for the development of vascular disease [1]. Environmental risk factors such as smoking, diets high in calories, saturated fats, carbohydrates and salt, with low fruit and vegetable intake, and sedentary lifestyles have exacerbated the increase in CVD [2]. Metabolic syndrome and type II diabetes mellitus [3], as well as hypertension [4] are also individually associated with increased cardiovascular risk. The prevalence of dyslipidemia remains an important issue worldwide [5-9]. According to the Heart and Stroke Foundation of Canada, it is estimated that as many as 10 million Canadian adults have a cholesterol level higher than the recommended target [10]. The 2007-2009 CHMS reported that 41% of Canadian adults had a high total cholesterol level, about 36% had unhealthy levels of LDL, 30% had unhealthy levels of HDL, and about 25% of adults had unhealthy levels of TG [11]. Although single dyslipidemia has been extensively reported, very little is known about mixed dyslipidemia. Recently, Asghari et al. studied patients from across Canada and reported that one of every five primary care patients had mixed dyslipidemia [12]. In France, Laforest et al. described a mixed dyslipidemia prevalence of 31% among those patients treated with a statin [13], while a similar German study noted a prevalence of 32% [14].

Lowering LDL has long been the primary focus in treating dyslipidemia [15, 16], given that LDL is a major lipid component associated with the increasing risk of CVD [16]. However, abnormal levels of HDL and TG can also contribute to CVD [17-19].

NL has a higher level of CVD mortality than any other province in Canada [1, 20]. There are several factors that may explain this trend; one is the lipid profile pattern and high prevalence of dyslipidemia in this province. Several studies have found a high prevalence of hypercholesterolemia in NL [21-23]. These studies also suggest that NL's population may exhibit a unique pattern of dyslipidemia characterized by low HDL as compared to other Canadian provinces. For example, in a study involving several rural areas of NL, it was found that 61% of study volunteers had hypercholesterolemia [23]. However, this study involved a limited sample that is not representative of the province as a whole. Due to the settling of NL by English, Irish, and French settlers in the 17th century onward, these "founder families" present a unique genetic pool which makes it prone to various genetic diseases [24]. One recently discovered gene, RBP4, is considered to be a novel adipokine. Shea and colleagues discussed the association of the gene variants with serum HDL levels within the NL population. This study used a large sample of third generation Newfoundlanders and showed a significant association between the two noncoding regions of RBP4 and HDL level [25]. These findings are supported by the relative homogeneity of both genetics and culture in NL that is not present in other provinces. The objective of this study was to describe the prevalence of single and mixed dyslipidemia among patients of primary care settings in NL using EMR data from family physician's practices.

3.3 METHODS

3.3.1 Study design

This is a cross-sectional study using secondary analysis of existing data. The variables influencing dyslipidemia were extracted from the CPCSSN database.

3.3.2 Data source

Data from this study was collected from the NL subset of the CPCSSN database. The CPCSSN is a multi-disease electronic record surveillance system, used specifically for examining chronic diseases in primary care as well as for primary care research. Participating sentinel primary care practices allow electronic extraction of de-identified data from their EMRs on a quarterly basis in order to build a Canadian database of primary care [26]. Data from four family physicians practices' in St. John's, NL, who are part of CPCSSN, were used in this study. All the available lipid profiles from January 1, 2009 to December 31, 2010 were extracted.

3.3.3 Variables

The most recent lipid profile (total cholesterol, HDL, LDL, TG) and the date it was performed was identified for each patient. The most recent Canadian guidelines for the diagnosis and management of dyslipidemia at the time of the study was used to define the abnormal lipid profile [15]. These dyslipidemia elements include total cholesterol >5.2 mmol/L, HDL <1.0 mmol/L, LDL >3.4 mmol/L, total cholesterol/HDL ratio >5.0 and TG >1.7 mmol/L [15,27]. The ratio of total cholesterol to HDL was calculated by dividing the total cholesterol by HDL. Single dyslipidemia was defined as the existence of only one abnormal lipid element in the individual; whereas mixed dyslipidemia was defined as the existence of more than one lipid disorder. In the initial descriptive statistics

of the study population, all five components were presented; however, in the regression analyses, the total cholesterol and ratio were not considered as they both contained elements of the other three components.

The following demographic variables were used in this study: sex, age, and place of residence. The rural/urban residence was determined by using the second character of each individual's postal code address. Obesity was defined as individuals having a BMI ≥ 30 . BMI was defined as the individual's body mass divided by the square of his or her height, using the most recent height and weight measurements. Individuals were classified as smokers or non-smokers according to the ICD code records in the EMRs. BMI and smoking status were included for a subsample of the study population. History of dyslipidemia was defined as any previous record of ICD-9 code 272 (disorders of lipid metabolism) or text record for dyslipidemia. CPCSSN case definitions were used to identify both diabetes mellitus and hypertension [28]. These definitions identify occurrences of the following indicators to index the disease. These include diagnoses within two years, selected from the first diagnosis code available by the physician claim, laboratory results, and medication used for treatment of the particular disease [26, 29]. Previous studies show a high sensitivity and specificity for these case definitions [26, 30, 31]. Individuals prescribed a lipid-lowering medication were stratified into two categories: Medication users (those with any record of lipid-lowering agents in the database during the 3 months before the date of a blood test); and non-medication users (those with no record of lipid-lowering medication use within 3 months before the data of a blood test). This classification of medication use has been used in previous studies utilizing lipid profile databases [32].

3.3.4 Study population

Adults aged 20 years or older from the NL CPCSSN database were included in this study. Moreover, the patients must have had a complete lipid profile taken during the study timeframe (January 1, 2009-December 31, 2010). Among the 15,865 patients 4,382 had a complete lipid profile conducted. Pregnant women were excluded from the analysis by identifying ICD code records, text code records within the EMR, or any event related to pregnancy [33].

3.3.5 Statistical analysis

Demographics, prevalence of single and mixed dyslipidemia, co-morbidities, and lipid-lowering medication use of the study population was summarized using descriptive statistics and compared using student's t test and the chi-square (X^2) test. Multinomial logistic regression analyses were conducted to identify factors which influenced single and mixed dyslipidemia. Significance of effects was evaluated at $\alpha=0.05$. SPSS version 22.0 [34] was used to perform all the statistical analyses.

3.3.6 Ethics

The Human Research Ethics Authority, Memorial University of Newfoundland, reviewed and approved the study protocol, reference number 11.090. All the data was de-identified prior to the analysis.

3.4 RESULTS

3.4.1 Population characteristics

Among 4,382 individuals who met the study criteria, the majority (59%, n=2,578) were female and the average age was 58 years (SD=16). The majority of the sample

(89%, n=3,794) resided in urban areas. At least one-third of patients were hypertensive (34%, n=1,469), a quarter had a history of dyslipidemia (26%, n=1,144), and 15% (n=666) were diabetic. The mean BMI was 31 kg/m² (SD=16) and 42% of the sample had a previous or current smoking status recorded (Table 3.1). The mean lipid profile levels in the study population are presented in Table 3.1.

Table 3.1. Characteristics of the study population of primary care patients in NL (n=4,382)

	Male (n=1,804)	Female (n=2,578)	Total (n=4,382)	p-value
Age*	58.2±14.6	58.0±14.9	58.1±14.8	NS
BMI**	31.3±14.7	31.0±16.7	31.1±15.9	NS
Residence (rural)	11.0% (n=190)	10.5% (n=263)	10.7% (n=453)	NS
Smoking (current)#	46.9% (n=307)	38.7% (n=326)	42.1% (n=633)	p<0.001
Hypertension	34.8% (n=628)	32.6% (n=841)	33.4% (n=1,469)	NS
Diabetes mellitus	18.7% (n=1,466)	12.7% (n=328)	15.2% (n=666)	p<0.0001
History of dyslipidemia	32.3% (n=583)	21.8% (n=561)	26.1% (n=1,144)	p<0.0001
Lipid-lowering medication use	44.4% (n=801)	29.2% (n=752)	35.4% (n=1,553)	p<0.0001
Lipid profile				
Total cholesterol*	4.8±1.1	5.2±1.1	5.0±1.1	p<0.0001
HDL*	1.1±0.3	1.4±0.4	1.3±0.4	p<0.0001
LDL*	3.0±1.0	3.1±0.9	3.1±0.9	p<0.0001
TG*	1.5±1.1	1.3±0.9	1.4±0.9	p<0.0001
Total cholesterol/HDL ratio*	4.5±1.4	3.9±1.1	4.1±1.3	p<0.0001

Figures are a percentage except for:

*Mean ± SD.

#Besides smoking and BMI which have missing rates 66% and 67% respectively, the missing rates in all other variables are below 5% of the total population (n=4,382).

Comparisons were obtained from student's t test and chi-square (X²) test.

Comparisons were made between male and female.

BMI, body mass index; HDL, high-density lipoprotein; LDL, low-density lipoprotein; TG, triglyceride; NS, not significantly different.

3.4.2 Prevalence of single and mixed dyslipidemia in patients of EMR primary care setting in NL

A total of 4,382 patients had lipid levels available for all three components of HDL, LDL and TG in their EMRs concomitantly; and hence, were considered for further exploration. The prevalence of dyslipidemia for each component (regardless of the existence of other forms of dyslipidemia) was 25% for low HDL, 43% for high total cholesterol, 36% for high LDL, 25% for high TG, and 22% for elevated total cholesterol/HDL ratio. Figure 3-1 illustrates the prevalence of single (mutually exclusive dyslipidemia) and mixed dyslipidemia among individuals with abnormal levels of LDL, HDL and TG dyslipidemia. In approximately 39.4% of this population, all three components were in normal ranges, whereas 41.2%, 16.5% and 2.9% had abnormalities in one, two and three lipid components, respectively. The most common shared abnormality is between those with low HDL and elevated TG dyslipidemia, representing 7.3% of the population. A lesser proportion of the population shared elevated LDL and TG (6.2%), and even fewer shared low HDL and elevated LDL (3.0%).

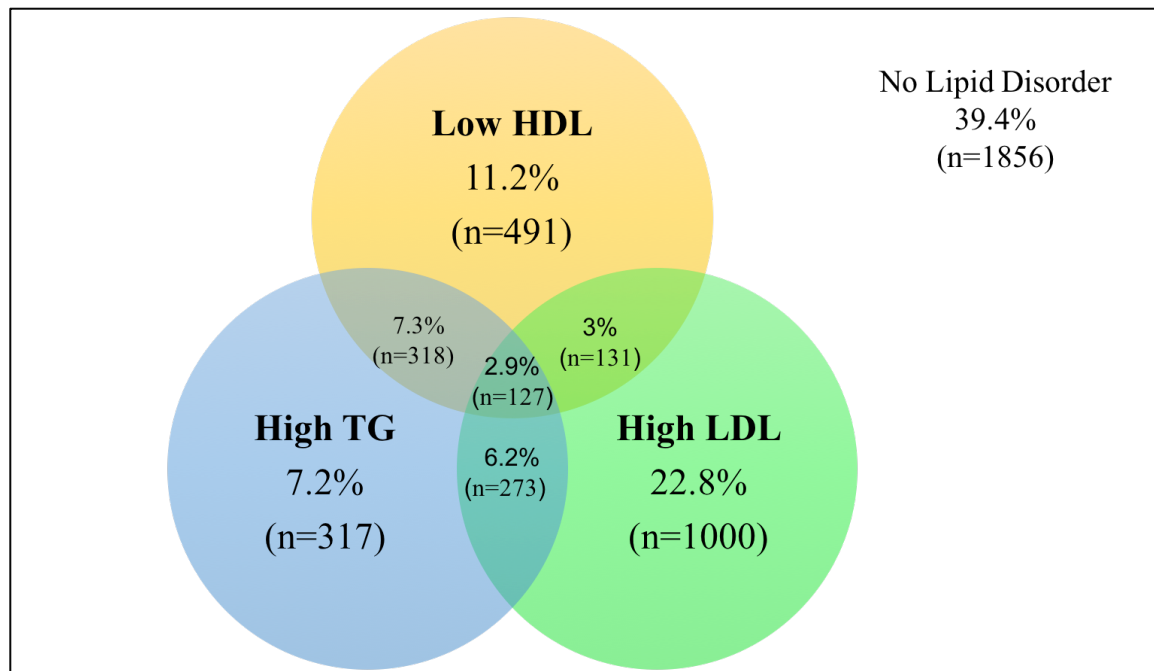


Figure 3-1. Prevalence of single and mixed dyslipidemia of LDL, HDL, and TG in patients of EMR primary care settings in NL (n=4382).

EMR, electronic medical records; HDL, high-density lipoprotein; LDL, low-density lipoprotein; TG, triglyceride.

The figure does not account for the possibility of having dyslipidemia of total cholesterol or total cholesterol to HDL ratio. The sizes of the circles are schematic and do not represent their true values.

3.4.3 Factors associated with single and mixed dyslipidemia

The prevalence of each kind of dyslipidemia has been stratified by the risk factors associated with CVD in Table 3.2. Among female patients, 6.5% had low HDL, 27.5% had elevated LDL, and 8.1% had elevated TG. Nearly 1.5% of female patients had mixed dyslipidemia including all three lipid profile components (HDL+LDL+TG). 34.5% of patients with hypertension had a normal lipid profile, while 13.9% had low HDL, 18.5% had elevated LDL, and 10.5% had high TG. Table 3.3 represents the results of the multinomial logistic regression modelling for factors associated with dyslipidemia.

Among demographic factors, females were less likely to have low HDL dyslipidemia (OR, 0.23; 95% CI, 0.14-0.38), whereas they were more likely to have high LDL dyslipidemia (OR, 1.86; 95% CI, 1.28-2.70), high TG (OR, 1.95; 95% CI, 1.09-3.48), and accordingly elevated LDL & TG (OR, 2.57; 95% CI, 1.32-4.99). Patients with hypertension (OR, 0.58; 95% CI, 0.35-0.96) were less likely to have low HDL dyslipidemia. In a subsample, (n=550) patients who were obese were more likely to have low HDL dyslipidemia (OR, 1.98; 95% CI, 1.23-3.19), as well as abnormal levels of mixed dyslipidemia (HDL+LDL+TG) (OR, 2.34; 95% CI, 1.09-5.05).

Table 3.2. The prevalence of cardiovascular risk factors with various combinations of single and mixed dyslipidemia in patients of EMR primary care settings in NL (n=4382)

	Males n=1804 (%)	Females n=2578 (%)	Non- smoker n=863 (%)	Current Smoker n=633 (%)	Normal/ underweight n=855 (%)	Obese n=570 (%)	Non- hypertensive n=2913 (%)	Hypertensive n=1469 (%)	Non- diabetic n=3716 (%)	Diabetic n=666 (%)	Non- medication user n=2829 (%)	Lipid- lowering medication user n=1553 (%)
Normal	30.2	43.6	34.2	34.5	38.9	27.5	39.9	34.5	38.6	35.1	38.4	37.6
HDL	18.5	6.5	10.8	13.9	10.6	16.5	10.2	13.9	9.7	21.2	7.8	18.2
LDL	17.2	27.5	28.7	18.0	25.6	16.5	25.7	18.5	26.4	5.5	31.2	8.5
TG	6.4	8.1	6.9	7.5	5.8	9.6	5.8	10.5	6.5	12.4	4.4	12.9
HDL & LDL	4.9	1.7	2.8	3.9	3.2	4.2	3.4	2.4	3.4	1.3	3.7	1.9
HDL & TG	12.6	3.8	6.8	9.8	6.7	15.1	6.0	10.3	5.5	18.2	4.5	12.9
LDL & TG	5.2	7.2	6.9	8.2	7.0	5.5	6.2	6.8	6.8	4.1	7	5.1
HDL & LDL & TG	5.0	1.5	3.0	4.2	2.3	5.1	2.9	3.1	3.1	2.4	3	2.9

EMR, electronic medical records; HDL, high-density lipoprotein; LDL, low-density lipoprotein; TG, triglycerides

Table 3.3. Factors which influence single and mixed dyslipidemia in patients of EMR primary care settings in NL (n=4382)

Variable	Odds Ratio (95% Confidence Interval)						
	HDL	LDL	TG	HDL & LDL	HDL & TG	LDL & TG	HDL & LDL & TG
Sex (Female)	0.23 (0.14-0.38)^	1.86 (1.28-2.70)*	1.95 (1.09-3.48)*	0.29 (0.13-0.63)*	0.26 (0.15-0.44)^	2.57 (1.32-4.99)*	0.22 (0.10-0.51)^
Age	0.98 (0.97-1.00)^	1.01 (0.99-1.02)	1.03 (1.00-1.05)*	0.98 (0.95-1.01)	1.00 (0.98-1.02)	1.02 (1.00-1.05)	0.99 (0.96-1.02)
Hypertension	0.58 (0.35-0.96)*	0.93 (0.63-1.38)	1.73 (0.96-3.12)	0.98 (0.41-2.37)	1.28 (0.76-2.15)	1.20 (0.65-2.21)	1.20 (0.55-2.64)
Diabetes	1.12 (0.65-1.93)	0.46 (0.24-0.86)*	1.25 (0.66-2.36)	0.36 (0.80-1.67)	1.74 (1.00-3.04)	0.67 (0.28-1.56)	0.39 (0.13-1.22)
Current smoker^a	1.58 (0.99-2.50)	0.73 (0.50-1.06)	0.89 (0.50-1.57)	1.89 (0.88-4.05)	0.89 (0.54-1.47)	0.97 (0.53-1.76)	1.47 (0.70-3.08)
Obese^b	1.98 (1.23-3.19)*	0.79 (0.54-1.15)	1.46 (0.83-2.58)	1.11 (0.50-2.48)	1.55 (0.92-2.59)	0.82 (0.44-1.51)	2.34 (1.09-5.05)*
Lipid lowering medication use	2.60 (1.52-4.43)^	0.28 (0.18-0.45)^	2.37 (1.25-4.53)*	0.42 (0.16-1.11)	2.26 (1.25-4.08)*	0.79 (0.41-1.51)	1.05 (0.46-2.38)
Rural residence	1.14 (0.67-1.96)	0.98 (0.59-1.62)	1.31 (0.69-2.51)	0.89 (0.32-2.48)	1.47 (0.83-2.58)	1.43 (0.70-2.91)	0.95 (0.38-2.35)

*p<0.05

^p<0.0001

^a Smoking status was compared with non-smokers.

^b Obesity status was compared with normal and underweight individuals.

EMR, electronic medical records; HDL, high-density lipoprotein; LDL, low-density lipoprotein; TG, triglycerides.

3.4.4 Comparison of single and mixed dyslipidemia between lipid-lowering medication users and non-medication users

The pattern of single and mixed dyslipidemia varied among lipid-lowering and non-medication users (Figure 3-2). Among patients with single dyslipidemia, the prevalence of low HDL was higher among lipid-lowering medication users (18.2% vs. 7.8%, $p<0.0001$), whereas elevated LDL was higher among non-medication users (31.2% vs. 8.5%, $p<0.0001$). For those with mixed dyslipidemia, the combination of low HDL and high TG was higher among lipid-lowering medication users (12.9% vs. 4.5%, $p<0.0001$). All forms of mixed dyslipidemia that contain LDL were less prevalent among lipid-lowering medication users than non-medication users.

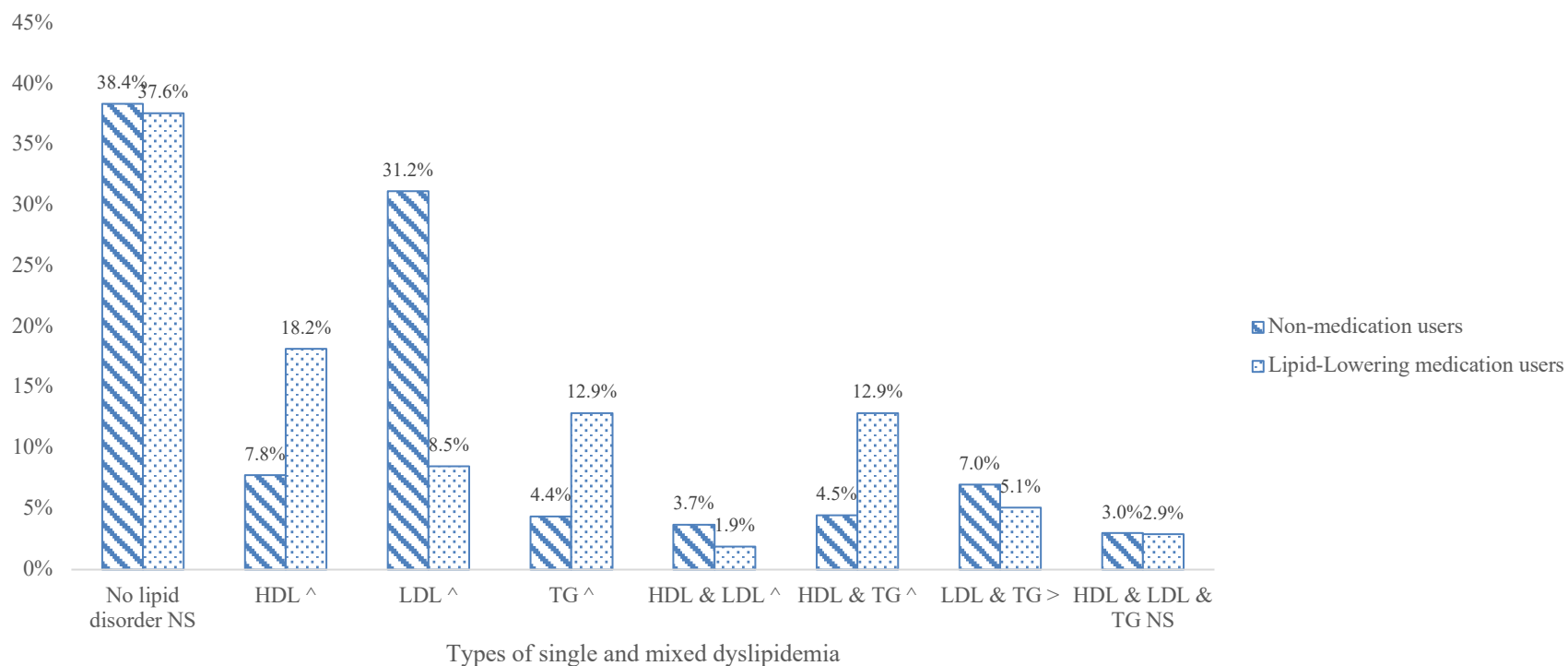


Figure 3-2. Single and mixed dyslipidemia in lipid-lowering medication users and non-medication users in patients of EMR primary care settings in NL who had lipid test during the study period (n=4382).

EMR, electronic medical record; HDL, high-density lipoprotein; LDL, low-density lipoprotein; TG, triglycerides.

>p-value for two independent sample proportion test <0.01.

^p-value for two independent sample proportion test <0.0001.

NS, not significantly different

3.5 DISCUSSION

By applying Canadian national guidelines, we observed that nearly one out of five (19.4%) of the NL-CPCSSN primary care population had mixed dyslipidemias, with low HDL and elevated TG occurring in approximately 7.3% of patients, compared to 2.9% with all three dyslipidemias in NL. Similarly, Asghari et al. studied the Canada-CPCSSN primary care population and reported that one of every five primary care patients had mixed dyslipidemia [12]. This group has a high risk for developing CVDs [17-19]. These results are in contrast with the findings by Halcox and Misra who described a mixed dyslipidemia prevalence in one of ten patients in the general population and in 15% of statin-treated patients [35].

Among primary care patients in NL, many of whom were treated with lipid-lowering medications, the prevalence of high LDL dyslipidemia, decreased significantly with treatment, from 31.2% to 8.5%. This trend was also appreciable for mixed dyslipidemia which included high LDL. This is in line with most guidelines, as targeting LDL cholesterol with statins is generally the primary goal when managing dyslipidemia [15]. However, in our review we noted that cholesterol lowering guidelines do not specifically address mixed dyslipidemia. This is surprising given the high prevalence of mixed dyslipidemia among primary care patients. Conversely, other forms of dyslipidemia (HDL and TG) had a higher prevalence of dyslipidemia among lipid-lowering medication users. An Australian study noted similar findings where following lipid-lowering therapy, a substantial number of patients reached their LDL goals,

however, those patients with low HDL, elevated TG, and approximately one third of those with mixed dyslipidemia did not attain their goal [36].

Multiple co-morbidities were prevalent in the patient population. Males had a higher prevalence of smoking, diabetes, history of dyslipidemia, and lipid-lowering medication use compared to females. Males also tended to have lower levels of HDL, higher levels of TG, and a higher total cholesterol to HDL ratio. Unexpectedly, in our study diabetes was not associated with increased TGs and low HDL as has been regularly reported [37-39]. Although the absolute differences in serum lipid levels reported between males and females were small, minor changes in serum cholesterol levels are often clinically meaningful. For example, in the 2009 Canadian cardiovascular Society/Canadian guidelines for the diagnosis and treatment of dyslipidemia and prevention of cardiovascular disease, lowering the LDL level to a mean of 2.0 mmol/L or less is associated with the lowest risk of recurrent cardiovascular events. Furthermore, every 1.0 mmol/L reduction in LDL is associated with a corresponding 20-25% reduction in cardiovascular mortality [15].

Results of this study illustrate the high prevalence of both single and mixed dyslipidemias in the NL population as well as a high prevalence of complicated patients who have different types of dyslipidemia in addition to other co-morbidities such as diabetes. Dyslipidemia in patients with diabetes is known to have considerable CVD-related morbidity and these patients are also at high risk for complications associated with atherosclerosis and should therefore receive targeted intervention [40]. A high prevalence of mixed dyslipidemia among NL patients with diabetes may also be a contributing factor to NL's high level of CVD mortality.

The findings of this study should be considered in light of the following limitations. Ethics approval and secondary use of data approval were only granted for the study timeframe (January 1, 2009-December 31, 2010), therefore having more recent data would have required a new request which was not feasible in the timeline of this two-year program. This cross-sectional study using secondary analysis of data was aimed to provide a snapshot of dyslipidemia in a primary healthcare setting in NL. The study sample includes patients who received healthcare services in primary healthcare settings with EMRs in St. John's, NL, therefore, the sample is not representative and generalizable to the entire population of NL. This has implications for selection bias and potentially impacts on our prevalence estimates. This study only included patients with a complete lipid profile. This did limit our patient population; however, a complete lipid profile was important for our analysis on mixed dyslipidemia. Similar to other studies using secondary data, there is also potential for incomplete and inaccurate data. Moreover, our reported prevalence of single and mixed dyslipidemia may be underreporting the true values as lipid-lowering medication therapy may be masking the actual figures. The study also does not provide information on patient adherence to lipid-lowering medication use. Further, we did not have any data on some of the lifestyle factors that are associated with dyslipidemia, such as exercise and diet. Abnormal lipid profile parameters were defined according to 2009 Canadian Cardiovascular Society guidelines for the diagnosis and treatment of dyslipidemia. The most recent 2016 guidelines do not change how abnormal lipid profiles are defined, which should not change the interpretation of our results [41].

3.6 CONCLUSION

Dyslipidemia is a leading risk factor of CVD. Results of this study illustrate the high prevalence of both single and mixed dyslipidemias in the NL population, as well as a high prevalence of complicated patients who have abnormal lipid levels in addition to other co-morbidities such as diabetes and hypertension. Although the use of lipid-lowering medications appears to have reduced the levels of elevated LDL, a significant number do not attain their goals for other lipid components. These findings highlight the need for adopting a strategy for the early prevention and management of dyslipidemia, as well as further investigation into lipid profiles in NL. We suspect that the high prevalence of dyslipidemia in this province may be a contributing factor to NL's high level of CVD mortality.

3.7 REFERENCES

1. Filate WA, Helen LJ, Courtney CK, Jack VT. Regional variations in cardiovascular mortality in Canada. *Can J Cardiol.* 2003;19(11):1241-8.
2. National Clinical Guideline Centre (UK). Lipid Modification: Cardiovascular Risk Assessment and the Modification of Blood Lipids for the Primary and Secondary Prevention of Cardiovascular Disease. 2014.
3. Cannon CP. Mixed dyslipidemia, metabolic syndrome, diabetes mellitus, and cardiovascular disease: Clinical implications. *Am J Cardiol.* 2008;102(12A):5L-9L.
4. Padwal R, Hackam D, Khan N, Tobe S. Primary prevention of CVD: Modification of diet in people with hypertension. *BMJ Clin Evid.* 2016;2016:0214.

5. Esteghamati A, Meysamie A, Khalilzadeh O, et al. Third national surveillance of risk factors of non-communicable diseases (SuRFNCD-2007) in Iran: Methods and results on prevalence of diabetes, hypertension, obesity, central obesity, and dyslipidemia. *BMC Public Health*. 2009;9:167.
6. Aguilar-Salinas CA, Gomez-Perez FJ, Rull J, Villalpando S, Barquera S, Rojas R. Prevalence of dyslipidemias in the Mexican national health and nutrition survey 2006. *Salud Publica Mex*. 2010;52 Suppl 1:S44-53.
7. Goodman SG, Langer A, Bastien NR, et al. Prevalence of dyslipidemia in statin-treated patients in Canada: Results of the DYSlipidemia international study (DYSIS). *Can J Cardiol*. 2010;26(9):e330-5.
8. Janus ED, Tideman PA, Dunbar JA, et al. Dyslipidaemia in rural australia: Prevalence, awareness, and adherence to treatment guidelines in the greater green triangle risk factor study. *Med J Aust*. 2010;192(3):127-132.
9. Wang S, Xu L, Jonas JB, You QS, Wang YX, Yang H. Prevalence and associated factors of dyslipidemia in the adult Chinese population. *PLoS One*. 2011;6(3):e17326.
10. Heart and Stroke Foundation of Canada.
<http://www.heartandstroke.com/site/c.ikIQLcMWJtE/b.3483991/k.34A8/Statistics.htm#bloodcholesterol>. Updated 2011. Accessed 09/29, 2017.
11. Statistics Canada. Canadian health measures survey. <http://www.statcan.gc.ca/daily-quotidien/100323/dq100323a-eng.htm>. Updated 2011. Accessed 09/28, 2017.
12. Asghari S, Aref-Eshghi E, Godwin M, Duke P, Williamson T, Mahdavian M. Single and mixed dyslipidaemia in Canadian primary care settings: Findings from the

- Canadian primary care sentinel surveillance network database. *BMJ Open*. 2015;5(12):e007954-2015-007954.
13. Laforest L, Ambegaonkar BM, Souchet T, Sazonov V, Van Ganse E. Mixed dyslipidemias in primary care patients in France. *Vasc Health Risk Manag*. 2012;8:247-254.
14. Gitt AK, Junger C, Smolka W, Bestehorn K. Prevalence and overlap of different lipid abnormalities in statin-treated patients at high cardiovascular risk in clinical practice in Germany. *Clin Res Cardiol*. 2010;99(11):723-733.
15. Genest J, McPherson R, Frohlich J, et al. 2009 Canadian cardiovascular Society/Canadian guidelines for the diagnosis and treatment of dyslipidemia and prevention of cardiovascular disease in the adult - 2009 recommendations. *Can J Cardiol*. 2009;25(10):567-579.
16. Soran H, Dent R, Durrington P. Evidence-based goals in LDL-C reduction. *Clin Res Cardiol*. 2017;106(4):237-248.
17. Gordon T, Castelli WP, Hjortland MC, Kannel WB, Dawber TR. High density lipoprotein as a protective factor against coronary heart disease. the framingham study. *Am J Med*. 1977;62(5):707-714.
18. Gordon DJ, Probstfield JL, Garrison RJ, et al. High-density lipoprotein cholesterol and cardiovascular disease. four prospective American studies. *Circulation*. 1989;79(1):8-15.
19. Sarwar N, Danesh J, Eiriksdottir G, et al. Triglycerides and the risk of coronary heart disease: 10,158 incident cases among 262,525 participants in 29 western prospective studies. *Circulation*. 2007;115(4):450-458.

20. Statistics Canada. Mortality, summary list of causes.
<http://www.statcan.gc.ca/pub/84f0209x/84f0209x2007000-eng.htm>. Updated 2007.
Accessed 09/25, 2017.
21. Fodor JG, Balram BC, Heyden S. Lipid and lipoprotein levels of Newfoundland school children. *Ann Nutr Metab*. 1987;31(4):231-236.
22. Chockalingam A, J GF. Cardiovascular risk factors in newfoundland population and modification of their level through nonpharmacological intervention. *J Cardiovasc Pharmacol*. 1990;16 Suppl 8:S54-6.
23. Fodor JG, Chockalingam A. Modification of cardiovascular risk factors in rural population of Newfoundland through nutritional counselling. *Cor Vasa*. 1991;33(2):107-13.
24. Pope AM, Carr SM, Smith KN, Marshall HD. Mitogenomic and microsatellite variation in descendants of the founder population of Newfoundland: High genetic diversity in an historically isolated population. *Genome*. 2011;54(2):110-119.
25. Shea JL, Loredó-Osti JC, Sun G. Association of RBP4 gene variants and serum HDL cholesterol levels in the Newfoundland population. *Obesity (Silver Spring)*. 2010;18(7):1393-1397.
26. Birtwhistle R, Keshavjee K, Lambert-Lanning A, et al. Building a pan-Canadian primary care sentinel surveillance network: Initial development and moving forward. *J Am Board Fam Med*. 2009;22(4):412-422.
27. Statistics Canada. Heart health and cholesterol levels of Canadians, 2007 to 2009.
<http://www.statcan.gc.ca/pub/82-625-x/2010001/article/11136-eng.pdf>. Updated 2010. Accessed 09/25, 2017.

28. CPCSSN disease definitions. <http://cpcssn.ca/research-resources/case-definitions/>. Updated 2016. Accessed 04/01, 2018.
29. Public Health Agency of Canada. Report from the Canadian chronic disease surveillance system: Hypertension in Canada, 2010. http://www.phac-aspc.gc.ca/cd-mc/cvd-mcv/ccdss-snsmc-2010/pdf/CCDSS_HTN_Report_FINAL_EN_20100513.pdf. Updated 2010. Accessed 02/07, 2017.
30. 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):512-516.
31. Robitaille C, Dai S, Waters C, et al. Diagnosed hypertension in Canada: Incidence, prevalence and associated mortality. *CMAJ*. 2012;184(1):E49-56.
32. Jones PH, Nair R, Thakker KM. Prevalence of dyslipidemia and lipid goal attainment in statin-treated subjects from 3 data sources: A retrospective analysis. *J Am Heart Assoc*. 2012;1(6):e001800.
33. Devine S, West S, Andrews E, et al. The identification of pregnancies within the general practice research database. *Pharmacoepidemiol Drug Saf*. 2010;19(1):45-50.
34. IBM corp. Released 2013. IBM SPSS statistics for windows, version 22.0. Armonk, NY: IBM corp. <https://www.ibm.com/analytics/data-science/predictive-analytics/spss-statistical-software>.
35. Halcox J, Misra A. Type 2 diabetes mellitus, metabolic syndrome, and mixed dyslipidemia: How similar, how different, and how to treat? *Metab Syndr Relat Disord*. 2015;13(1):1-21.

36. Colquhoun D, Chirovsky D, Sazonov V, Cui YA, Ambegaonkar B. Prevalence of mixed dyslipidemia among Australian patients undergoing lipid-modifying therapy. *Exp Clin Cardiol.* 2013;18(1):e32-6.
37. Verges B. New insight into the pathophysiology of lipid abnormalities in type 2 diabetes. *Diabetes Metab.* 2005;31(5):429-439.
38. Bruckert E, Baccara-Dinet M, McCoy F, Chapman J. High prevalence of low HDL-cholesterol in a pan-European survey of 8545 dyslipidaemic patients. *Curr Med Res Opin.* 2005;21(12):1927-1934.
39. Bruckert E, Baccara-Dinet M, Eschwege E. Low HDL-cholesterol is common in European type 2 diabetic patients receiving treatment for dyslipidaemia: Data from a pan-European survey. *Diabet Med.* 2007;24(4):388-391.
40. Krentz AJ. Lipoprotein abnormalities and their consequences for patients with type 2 diabetes. *Diabetes Obes Metab.* 2003;5 Suppl 1:S19-2.
41. Anderson TJ, Gregoire J, Pearson GJ, et al. 2016 Canadian Cardiovascular Society Guidelines for the Management of Dyslipidemia for the Prevention of Cardiovascular Disease in the Adult. *Can J Cardiol.* 2016;32(11):1263-1282.

**Chapter 4: Comparison of Canadian Primary Care Sentinel
Surveillance Network Database and Canadian Health Measures Survey
on the Lipid Profile and Prevalence of Dyslipidemia in Newfoundland**

4.1 ABSTRACT

NL has higher mortality from CVD among younger adults than any other Canadian province. Lipid profiles of NL residents might partially explain this difference, but no lipid profile data from NL has been recently reported. Most recent lipid profile scores, medical history, and demographic information, were extracted from 2 years of CPCSSN patient EMRs from 4 clinics in NL. This was compared to Canadian lipid profile data from the CHMS Cycle 1. The CPCSSN database held 3,983 NL lipid profiles that met the study criteria. There were no significant overall differences in mean lipid component measures between the samples, including when the sample was stratified by age and sex. However, significantly more men (40%) in NL had HDL dyslipidemia compared to Canada (25%; $p<0.001$) and, among adults aged 20 to 39, hypercholesterolemia (35% vs. 27%, $p<0.001$), hypertriglyceridemia (22% vs. 17%, $p<0.01$), and LDL dyslipidemia (32% vs. 27%, $p<0.001$) were significantly more prevalent in NL than Canada. In NL, hypercholesterolemia, hypertriglyceridemia, and LDL dyslipidemia are more prevalent in young adults, and HDL dyslipidemia levels are more prevalent in men in our study population, compared to the Canadian sample. This may partially explain NL's higher level of young adult mortality from CVD.

4.2 INTRODUCTION

CVD, including heart disease and stroke, are the leading causes of death for Canadians [1]. NL ranks in the bottom three Canadian provinces in terms of healthy behaviors [2] and has a higher level of CVD mortality in younger adults than any other province [3-5]. Poor blood lipid levels (including total cholesterol, HDL, LDL, and TG levels) are a major risk factor for CVD and may partially explain the higher level of CVD mortality in NL. There is also a higher degree of homogeneity of genetics and culture in NL than in the other provinces [6], both of which may contribute to abnormal lipid profiles and CVD.

Previous comparative studies document findings of relatively poor blood lipid levels in NL. In the 1980s, school-age children in NL were shown to have higher total cholesterol levels than those of American children matched for age, sex, and race [7]. In the 1990s, among a rural NL sample, 72% of the population had at least one risk factor (high blood pressure, high blood cholesterol, smoking, and high BMI) noted in their medical history and hypercholesterolemia was prevalent in 61% of the sample [8].

The most well-known surveys on cardiovascular risk in the general population of Canada are the Canadian Heart Health Study (CHHS) [9] and CHMS [10-12]. These studies suggest that nearly 40% of Canadian adults have hypercholesterolemia, 36% have unhealthy levels of LDL, 30% have unhealthy levels of HDL and 25% have hypertriglyceridemia [10-12]. However, these studies are either outdated or do not include a representative sample from NL.

Addressing cardiovascular risk factors can help prevent CVD, as well as many other chronic diseases that share the same risk factors. Many of these risk factors were examined in previous studies; however, a major limitation of these studies is the lack of biochemical measures of cardiovascular risk factors such as blood lipids. The purposes of the current study are to (1) describe the lipid profile pattern of a sample from NL and (2) compare this pattern to that of the rest of Canada.

4.3 METHODS

4.3.1 Study Design

This is a cross-sectional study using secondary analysis of existing data. The variables influencing dyslipidemia were extracted from the CPCSSN database. Data from the CHMS was used as a comparison for the sample of NL data.

4.3.2 Data Sources

Canadian Primary Care Sentinel Surveillance Network Database (CPCSSN)

CPCSSN is a pan-Canadian network of primary care sentinel practices that use EMRs. Data from these EMRs is abstracted quarterly and uploaded in a de-identified format to both regional and central (pan-Canadian) databases. The databases are used for chronic disease surveillance in primary care and are also used as a tool for conducting primary care research [13]. Data for this study was extracted from the NL regional component of the CPCSSN database. At the time of this study only four clinics were part of the CPCSSN and had laboratory

information system linked to their EMR. The abstracted de-identified data came from these four clinics.

Canadian Health Measures Survey (CHMS; Cycle 1)

Aggregated data from the CHMS Cycle 1 [10] was used as a comparison for the sample of NL data. Cycle 1 was the most recently released cycle at the time of the study. Data for CHMS Cycle 1 was collected from March 1, 2007 to March 31, 2009 and included persons residing in the 10 provinces and three territories [11].

This survey is a nationally representative sample of approximately 5,000 respondents aged 6-79, including 1,000 individuals per age-group strata (6 to 11, 12 to 19, 20 to 39, 40 to 59, and 60 to 79 years). Statistics Canada reports that responses are representative of 97% of Canadians. Cycle 1 involved data collection in 15 sites across the country, allocated by region in proportion to their populations as follows: Atlantic (1), Quebec (4), Ontario (6), Prairies (2), and British Columbia (2). The respondents were randomly selected using a systematic sampling method with probability proportional to the size of each site's population [12]. No data from NL were collected for this cycle of the CHMS.

Response to the survey was voluntary. Data were collected directly from survey respondents using a combination of computer-assisted personal interviews and a visit to a mobile clinic, where phlebotomists and laboratory technologists collected blood samples. Qualified health measures specialists evaluated the biological data [12].

4.3.3 Study Population

The study population included adults aged 20-79 from the NL CPCSSN database and the national CHMS Cycle 1. For the NL database, patients who had a lipid profile in their record between January 1, 2009 and December 31, 2010 were included. Pregnant women were excluded.

4.3.4 Lipid Profiles

Lipid profiles (including total cholesterol, HDL, LDL, and TG) of NL residents were identified from individuals' most recent lipid profiles in the CPCSSN database; the year of the most recent lipid test was recorded. Lipid profiles of Canadians were identified from the Statistics Canada report on the CHMS (Cycle 1) [10-12]. Dyslipidemia was defined using the Canadian guidelines for lipids [14, 15].

4.3.5 Variables

Demographic variables of interest from both the NL CPCSSN database and CHMS included sex and age. Place of residence (rural or urban), current diabetes or hypertension, history of dyslipidemia, and medication use were recorded because of potential influences on lipid profile results. Patients with postal codes corresponding to communities with populations of 1,000 or greater were deemed urban and populations of less than 1,000 were deemed rural [16].

CPCSSN case definitions for EMR data were used to ascertain both diabetes and hypertension [17]. Patients were classified as having dyslipidemia if there was any record related to dyslipidemia in their EMR.

Medication use of the drug classes (Lipid Lowering Medication [18], ACE Inhibitors [18, 19], Beta Blockers [20-22] and Diuretics [20, 21, 23]) that have influence on lipid levels was also recorded. Lipid lowering medications, including statins and fibrates, are the recommended treatment for patients with high lipid levels [20]. Medication usage was stratified as: current user (Started before January 1, 2009 and have not stopped); previous user (Stopped before January 1, 2009); short term user (Started since January 1, 2009 and stopped before December 31, 2010); new user (Started since January 1, 2009 and have not stopped); and non-user (no record for medication use from January 1, 2007 to December 31, 2010).

Smoking status and BMI were included for patients who had information recorded on these variables (n = 1,373, n = 1,304 respectively). Patients were classified as smokers or non-smokers by any record related to positive smoking status in the EMR.

4.3.6 Statistical Analysis

Characteristics of the study population, as well as the mean and confidence intervals of the individual lipid components (total cholesterol, HDL, LDL, TG) were summarized using descriptive statistics. Multivariate linear regression and ANOVA with Bonferroni correction was performed to show the association between each lipid component, sex, and age, controlling for potential influential factors. The analysis was limited to a subsample for comparisons that included smoking and BMI. Aggregated data from CHMS Cycle 1 represented the lipid profiles of the other Canadian provinces [10, 12, 16]. Statistics Canada reports the data by age group and sex of the respondent. For our comparisons, we used data from both sexes in the 20-79 year age groups to match the

NL data. Mean, confidence interval, and percentiles were noted. Aggregated measures were compared using Chi-Square tests, t-tests for independent samples with unequal variance, and proportion tests for independent samples. Significance of effects was evaluated at $\alpha=0.05$. SPSS version 22.0 [23] was used to perform all the statistical analyses.

4.3.7 Ethics

The Human Research Ethics Authority, Memorial University of Newfoundland, reviewed and approved the study protocol, reference number 11.090. All the data was de-identified prior to the analysis.

4.4 RESULTS

4.4.1 Lipid Profile of Newfoundlanders

In total, 15,865 patients visited the four sentinel clinics in St. John's, NL between January 1, 2009 and December 31, 2010 and lipid profiles were collected from 4,664 patients. Approximately 3,983 patients met the study criteria. There were more women ($n = 2,350$) than men ($n = 1,633$) in the sample but no significant differences in age between the sexes. The majority of the sample (89%) resided in urban areas. Forty percent of the study subsample was obese ($BMI \geq 30$) and 43% were smokers. Approximately 31% of the sample had a diagnosis of hypertension, 25% had a history of dyslipidemia, and 14% were diabetic.

The mean values and 95% confidence intervals for all four components of the lipid profile are presented in Table 4.1. Women had significantly higher total

cholesterol (5.19 vs. 4.80mmol/L) ($p<0.001$), HDL (1.43 vs. 1.12mmol/L) ($p<0.001$), and LDL (3.17 vs. 2.99mmol/L) ($p<0.001$), while men had significantly higher TG levels (1.55 vs. 1.32mmol/L) ($p<0.001$). The mean level of cholesterol in patients aged 40 to 59 years was significantly higher than patients both younger than 39 ($p<0.001$) and older than 60 ($p<0.001$). The mean level of HDL in patients younger than 39 was significantly lower than patients older than 60 years ($p<0.001$). The mean level of LDL in patients aged 40 to 59 was higher than both other age groups ($p<0.001$). There were no significant differences between the age groups in TG levels.

Table 4.1. Lipid Profile of Patients of Canadian Primary Care Sentinel Surveillance Network in NL (n=3893)

Characteristics N (%)		Lipid profile			
		Total Cholesterol Mean (95%CI)	LDL Mean (95%CI)	HDL Mean (95%CI)	TG Mean (95%CI)
Sex	Female 2326 (58%)	5.19 (5.14-5.23)	3.17 (3.13-3.20)	1.43 (1.41-1.45)	1.32 (1.28-1.35)
	Male 1657 (42%)	4.84 (4.79-4.90)	3.03 (2.98-3.08)	1.12 (1.11-1.14)	1.55 (1.49-1.60)
		P < 0.001	P < 0.001	P < 0.001	P < 0.001
Age	20-39 489 (13%)	4.8 (4.8-5.0)	3.1 (3.0-3.1)	1.2 (1.1-1.2)	1.3 (1.2-1.4)
	40-59 1768 (44%)	5.1 (5.1-5.2)	3.2 (3.2-3.3)	1.2 (1.2-1.3)	1.4 (1.4-1.5)
	60- 79 1726 (43%)	4.9 (4.9-5.0)	3.0 (3.0-3.1)	1.3 (1.3-1.4)	1.4 (1.4-1.5)
		# P _a < 0.001	# P _a < 0.001	# P _a < 0.001	NS
Place of residence	Rural 426 (11%)	4.88 (4.78-4.99)	3.01 (2.92-3.10)	1.22 (1.2-1.3)	1.56 (1.47-1.66)
	Urban 3437 (89%)	5.06 (5.01-5.09)	3.13 (3.09-3.16)	1.32 (1.30-1.33)	1.40 (1.36-1.42)
		P < 0.01	P < 0.01	P < 0.001	P < 0.008
Hypertension	Yes 1225 (31%)	4.87 (4.81-4.93)	2.93 (2.87-2.98)	1.24 (1.22-1.27)	1.58 (1.53-1.63)
	No 3441 (87%)	5.04 (5.08-5.16)	3.19 (3.16-3.23)	1.33 (1.32-1.35)	1.34 (1.30-1.38)
		P < 0.001	P < 0.001	P < 0.001	P < 0.001
Diabetes	Yes 542 (14%)	4.25 (4.16-4.34)	2.37 (2.30-2.45)	1.10 (1.07-1.12)	1.74 (1.65-1.84)
	No 3351 (86%)	5.17 (5.14-5.21)	3.23 (3.20-3.26)	1.33 (1.32-1.35)	1.36 (1.33-1.39)
		P < 0.001	P < 0.001	P < 0.001	P < 0.001
History of Dyslipidemia	Yes 971 (25 %)	4.89 (4.80-4.97)	2.94 (2.88-3.01)	1.19 (1.16-1.21)	1.67 (1.29-1.36)
	No 3012 (75%)	5.09 (5.06-5.13)	3.16 (3.14-3.20)	1.34 (1.33-1.36)	1.32 (1.29-1.36)
		P < 0.001	P < 0.001	P < 0.001	P < 0.001
BMI *	≥ 30 529 (40 %)*	4.84 (4.74-4.93)	2.95 (2.87-3.04)	1.14 (1.11-1.17)	1.67 (1.62-1.79)
	≤ 30 775 (60%)	5.13 (5.04-5.20)	3.19 (3.13-3.26)	1.34 (1.31-1.36)	1.35 (1.28-1.41)
		P < 0.001	P < 0.001	P < 0.001	P < 0.001
Smoking*	Current/Former 593 (43%)*	5.00 (4.90-5.09)	3.06 (2.98-3.15)	1.23 (1.19-1.26)	1.59 (1.51-1.68)
	Nonsmoker 780 (57%)	5.19 (5.12-5.28)	3.25 (3.18-3.23)	1.32 (1.29-1.35)	1.38 (1.42-1.53)
		P < 0.001	P < 0.001	P < 0.001	P < 0.001
Total Study Population		5.02 (4.99-5.05)	3.08 (3.06-3.11)	1.30 (1.29-1.31)	1.42 (1.39-1.45)

*Percentage calculated out of number of patients with condition, not total patients;

P: p-value for T- test

NS: not significant

Pa: p-value for ANOVA;

The results of Bonferroni tests show significant difference ($P < 0.001$) between Patients aged 40-59 years and patients both younger than 39 and older than 60 for Cholesterol; Patients younger than 39 and patients older than 60years for HDL; Patients between 40 and 59 and both other age groups for LDL.

HDL, high-density lipoprotein; LDL, low-density lipoprotein; TG, triglycerides

As shown in Table 4.1, the mean levels of cholesterol ($p < 0.001$), LDL ($p < 0.001$), and TG ($p < 0.001$) for patients with diabetes were significantly lower compared to non-diabetic patients. Mean levels of total cholesterol ($p < 0.001$), LDL ($p < 0.001$), and TG ($p < 0.001$) were also significantly lower for patients with hypertension compared to those without hypertension. Furthermore, mean levels of total cholesterol ($p < 0.001$), LDL ($p < 0.001$), and TG ($p < 0.001$) for patients with a history of dyslipidemia were significantly lower compared to patients without this history. The results of a linear regression showed that lipid lowering medication use, stratified among the subcategories of medication users (new user, short term user, current user, previous user), was statistically significant and negatively associated with total cholesterol ($p < 0.001$), LDL ($p < 0.001$), and HDL ($p < 0.001$) compared to non-users (Table 4.2).

Table 4.2. Lipid Profile of Patients of Canadian Primary Care Sentinel Surveillance Network in NL by Medication Use (n=3893)

Medication use		B coefficient (95% Confidence Interval)			
		Total Cholesterol	HDL	LDL	TG
Lipid Lowering	Current user	-1.149 (-1.263 - -1.036)^	-0.244 (-0.287 - -0.200)^	-1.044 (-1.139 - -0.949)^	0.326 (0.225 - 0.427)^
	Previous user	-0.557 (-0.715 - --0.399)^	-0.202 (-0.263 - -0.141)^	-0.565 (-0.698 - -0.432)^	0.454 (0.313 - 0.594)^
	New user	-0.701 (-0.794 - -0.609)^	-0.195 (-0.231 - -0.159)^	-0.670 (-0.748 - -0.593)^	0.398 (0.316 - 0.480)^
	Short term user	-0.598 (-0.708 - -0.487)^	-0.171 (-0.214 - -0.129)^	-0.623 (-0.716 - -0.530)^	0.433 (0.335 - 0.530)^
ACE Inhibitors	Current user	-0.793 (-0.919 - -0.666)^	-0.198 (-0.246 - -0.150)^	-0.669 (-0.775 - -0.562)^	0.152 (0.041 - 0.263)*
	Previous user	-0.667 (-0.836 - -0.499)^	-0.173 (-0.237 - -0.109)^	-0.613 (-0.756 - -0.471)^	0.273 (0.125 - 0.421)^
	New user	-0.646 (-0.746 - -0.546)^	-0.143 (-0.181 - -0.105)^	-0.635 (-0.720 - -0.550)^	0.294 (0.206 – 0.382)^
	Short term user	-0.596 (-0.733 - -0.459)^	-0.103 (-0.155 - -0.051)^	-0.563 (-0.678 - -0.447)^	0.158 (0.038 - 0.277)
Beta Blockers	Current user	-0.776 (-0.938 - -0.614)^	-0.225 (-0.274 - -0.083)^	-0.681 (-0.818 - -0.544)^	0.297 (0.157 - 0.438)^
	Previous user	-0.692 (-0.945 - -0.439)^	-0.179 (-0.274 - -0.083)^	-0.605 (-0.819 - -0.391)^	0.215 (-0.008 - 0.437)
	New user	-0.745 (-0.886 - -0.604)^	-0.167 (-0.220 - -0.114)^	-0.680 (-0.799 - -0.561)^	0.229 (0.107 - 0.351)^
	Short term user	-0.600 (-0.806 - -0.393)^	-0.101 (-0.179 - -0.024)^	-0.558 (-0.733 - -0.383)^	0.144 (-0.035 - 0.323)
Diuretics	Current user	-0.300 (-0.447 - -0.153)^	-0.089 (-0.143 - -0.034)*	-0.300 (-0.425 - -0.176)^	0.180 (0.056 - 0.305)*
	Previous user	-0.274 (-0.470 - -0.079)*	-0.089 (-0.162 - -0.017)*	-0.241 (-0.407 - -0.075)*	0.136 (-0.030 - 0.302)
	New user	-0.257 (-0.380 - -0.135)^	-0.110 (-0.155 - -0.064)^	-0.257 (-0.361 - -0.153)^	0.240 (0.137 - 0.344)^
	Short term user	-0.148 (-0.306 - 0.010)	-0.079 (-0.138 - -0.021)*	-0.232 (-0.366 - -0.097)*	0.395 (0.262 - 0.528)^

* $p < 0.05$

[^] $p < 0.0001$

HDL, high-density lipoprotein; LDL, low-density lipoprotein; TG, triglycerides
Medication use compared to non-users.

Results of the multivariate linear regression included adjustments for age, sex, medication use, place of residence, hypertension, diabetes and history of dyslipidemia (Table 4.3). The results were also adjusted for LDL, HDL, and TG levels wherever applicable. As shown, age was positively associated with total cholesterol ($\beta=0.014$, $p<0.001$), LDL ($\beta=0.006$, $p<0.001$), and HDL ($\beta=0.006$, $p<0.001$) and sex was negatively associated with HDL ($\beta= -0.25$, $p<0.001$) and total cholesterol ($\beta= -0.27$, $p<0.001$) indicating that men had lower levels of HDL and cholesterol. These models are shown in Table 4.3.

Table 4.3. Lipid Profile and Some Associated Factors in NL Canadian Primary Care Sentinel Surveillance Patients (n=3893)

Variable	B coefficient (95% Confidence Interval)			
	Total Cholesterol	HDL	LDL	TG
Sex (Female)	-0.27 (-0.33--0.21)^	-0.25 (-0.27- 0.23)^	0.012 (-0.045-0.070)	-0.023 (-0.073-0.027)
Age	0.014 (0.011-0.016)^	0.006 (0.005-0.007)^	0.006 (0.004-0.009)^	0.0006 (-0.001-0.002)
Hypertension	0.034 (-0.058-0.127)	-0.0041 (-0.0394-0.0311)	0.037 (-0.045-0.12)	0.11 (0.032-0.18)*
Diabetes	-0.61 (-0.71 - -0.51)^	-0.087 (-0.126- -0.049)^	-0.51 (-0.59 - -0.42)^	0.12 (0.048-0.21)^
History of Dyslipidemia	0.25 (0.17-0.33)^	-0.021 (-0.052-0.009)	0.24 (0.17-0.31)^	0.13 (0.066-0.19)^
TG	0.39 (0.36-0.42)^	-0.16 (-0.17- -0.14)^	0.18 (0.15-0.22)^	--
LDL	--	0.027 (0.013-0.040)^	---	0.14 (0.12-0.17)^
HDL	--	--	0.15 (0.077-0.23)^	-0.67 (-0.73 - -0.61)^

The results are adjusted for medication use and place of residence as well as sex, age, hypertension, diabetes, history of dyslipidemia. The results are also adjusted for LDL, HDL, and TG wherever applicable.

* $p<0.05$

^ $p<0.0001$

HDL, high-density lipoprotein; LDL, low-density lipoprotein; TG, triglycerides.

4.4.2 NL Mean Cholesterol Levels Compared to the Rest of Canada

Overall, the cholesterol levels from the NL and Canadian samples were very similar. Mean scores for lipid components were not significantly different, as

shown by the t-test for independent samples with unequal variance. A breakdown of mean lipid scores by gender and age also showed no statistically significant differences between the samples for any of the lipid profile components (Table 4.4).

Table 4.4. Lipid profile by Sex and Age, Comparison between NL sample (n=3893) and Canadian Sample (n=3508)

	Men		Women		Both Sexes	
Cholesterol						
Age Group	NL Study	CHMS	NL Study	CHMS	NL Study	CHMS
20-39	4.9 (4.7-5.0)	4.8 (4.7-5.0)	4.9 (4.8-5.0)	4.7 (4.6-4.8)	4.9 (4.8-5.0)	4.8 (4.7-4.8)
40-59	5.1 (5.0-5.2)	5.2 (5.1-5.3)	5.2 (5.1-5.2)	5.2 (5.1-5.4)	5.1 (5.1-5.2)	5.2 (5.1-5.3)
60-79	4.6 (4.5-4.7)	5.1 (4.9-5.2)	5.3 (5.3-5.4)	5.4 (5.3-5.5)	5.0 (5.0-5.1)	5.2 (5.1-5.3)
20-79	4.8 (4.8-4.9)	5.0 (4.9-5.2)	5.2 (5.1-5.2)	5.1 (5.0-5.1)	5.0 (5.0-5.1)	5.0 (5.0-5.1)
LDL						
Age Group	NL Study	CHMS	NL Study	CHMS	NL Study	CHMS
20-39	3.2 (3.0-3.3)	3.1 (3.0-3.2)	3.0 (2.9-3.1)	2.6 (2.4-2.7)	3.1 (3.0-3.1)	2.8 (2.7-2.9)
40-59	3.3 (3.2-3.3)	3.2 (3.1-3.4)	3.2 (3.1-3.2)	3.2 (3.0-3.3)	3.2 (3.2-3.3)	3.2 (3.1-3.3)
60-79	2.8 (2.7-2.9)	3.1 (2.9-3.3)	3.2 (3.2-3.3)	3.1 (3.1-3.2)	3.0 (3.0-3.1)	3.1 (3.0-3.2)
20-79	3.0 (2.9-3.0)	3.1 (3.0-3.3)	3.2 (3.1-3.2)	2.9 (2.9-3.0)	3.1 (3.0-3.1)	3.0 (3.0-3.1)
HDL						
Age Group	NL Study	CHMS	NL Study	CHMS	NL Study	CHMS
20-39	1.1 (1.0-1.1)	1.2 (1.1-1.2)	1.4 (1.3-1.4)	1.5 (1.4-1.5)	1.2 (1.2-1.3)	1.3 (1.3-1.4)
40-59	1.1 (1.0-1.1)	1.2 (1.1-1.2)	1.4 (1.3-1.4)	1.5 (1.4-1.5)	1.3 (1.2-1.3)	1.3 (1.3-1.4)
60-79	1.2 (1.1-1.2)	1.2 (1.2-1.3)	1.5 (1.4-1.5)	1.5 (1.4-1.6)	1.3 (1.3-1.4)	1.4 (1.3-1.4)
20-79	1.1 (1.1-1.4)	1.2 (1.1-1.2)	1.4 (1.4-1.5)	1.5 (1.4-1.5)	1.3 (1.2-1.3)	1.3 (1.3-1.4)
TG						
Age Group	NL Study	CHMS	NL Study	CHMS	NL Study	CHMS
20-39	1.5 (1.3-1.7)	1.3 (1.1-1.4)	1.2 (1.1-1.3)	1.1 (1.0-1.2)	1.3 (1.2-1.4)	1.2 (1.1-1.3)
40-59	1.6 (1.6-1.7)	1.5 (1.3-1.6)	1.3 (1.2-1.3)	1.3 (1.2-1.5)	1.4 (1.4-1.5)	1.4 (1.3-1.5)
60-79	1.5 (1.4-1.5)	1.7 (1.5-1.8)	1.4 (1.4-1.5)	1.5 (1.4-1.7)	1.4 (1.4-1.5)	1.6 (1.5-1.7)
20-79	1.5 (1.5-1.6)	1.4 (1.3-1.5)	1.3 (1.3-1.4)	1.3 (1.2-1.4)	1.4 (1.4-1.5)	1.4 (1.3-1.4)

*Values for LDL and TG from the CHMS are calculated using fasted sub-samples, not full samples. The result was not significantly different using t-test for independent samples with unequal variance.

HDL, high-density lipoprotein; LDL, low-density lipoprotein; TG, triglycerides.

4.4.3 NL Prevalence of Dyslipidemia Compared to the Rest of Canada

The prevalence of TG, LDL, HDL, and total cholesterol dyslipidemia is shown in Figures 4-1, 4-2, 4-3, 4-4. In the NL sample, 37.3% of patients had LDL dyslipidemia, 24.4% had HDL dyslipidemia, 25.8% had hypertriglyceridemia, and 43.6% had hypercholesterolemia. The NL rate was not significantly different than the overall Canadian rate for LDL dyslipidemia (36%), hypertriglyceridemia (25%), and hypercholesterolemia (41%), but it was significantly different than the Canadian rate for HDL dyslipidemia (30%; $p<0.001$). For patients aged 20 to 39 years, a significantly higher proportion had hypercholesterolemia (35% vs 27%, $p<0.001$), hypertriglyceridemia (22% vs 17%, $p<0.01$), LDL dyslipidemia (32% vs 27%, $p<0.001$) in NL than Canada. Among patients aged 40 to 59 years, hypertriglyceridemia (23.9% vs 27.7%, $p<0.01$) and HDL dyslipidemia (24.8% vs 30.9%, $p<0.001$) were less prevalent in NL than Canada. Among those aged 60 to 79 years, hypercholesterolemia (44.1% vs 54.4%, $p<0.001$), TG (33.3% vs 28.6%, $p<0.01$), LDL (34.8% vs 40.9%, $p<0.001$), and HDL dyslipidemia (21.9% vs 30.5%, $p<0.001$) were significantly less frequent in NL than Canada.

About 40% of men and 13% of women in NL had a HDL score lower than recommended by Canadian Lipid Guidelines ($\text{HDL}<1.0\text{mmol/L}$) compared to 25% of men and 10% of women in the Canadian sample; the difference is significant for men ($p<0.001$).

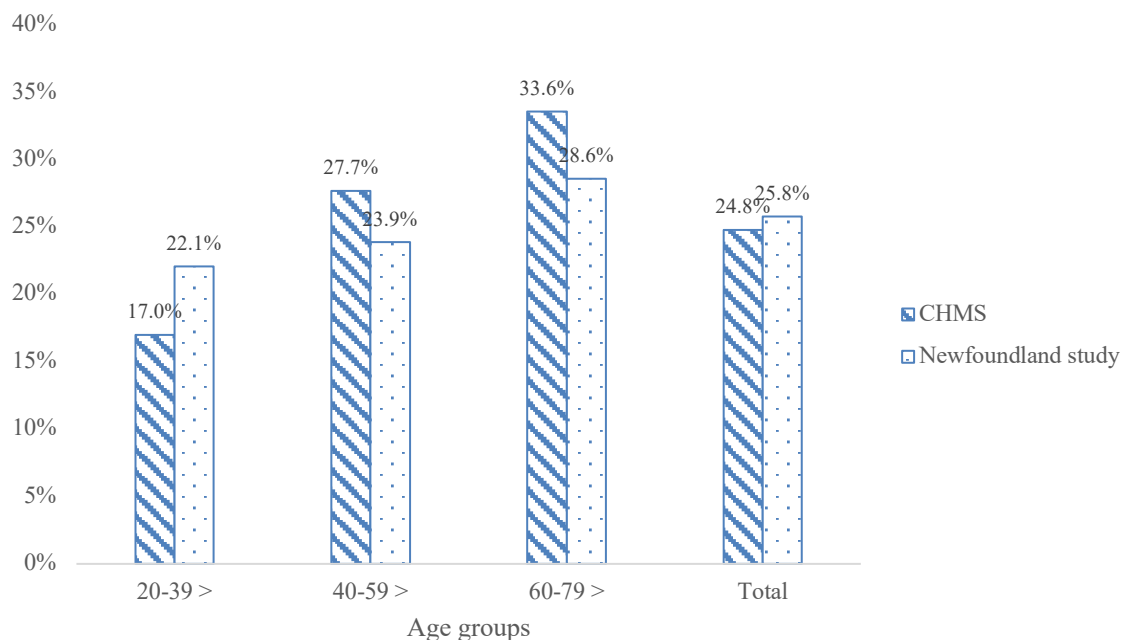


Figure 4-1. Percentage of adults with hypertriglyceridemia by age group. Comparison between NL study (n=3983) and CHMS (n=3508).

> P-value for two independent sample proportion test <0.01

Unhealthy level of lipids were significantly different by age group, using Chi Square Test (P-value <0.001), for each component of the lipid profile.

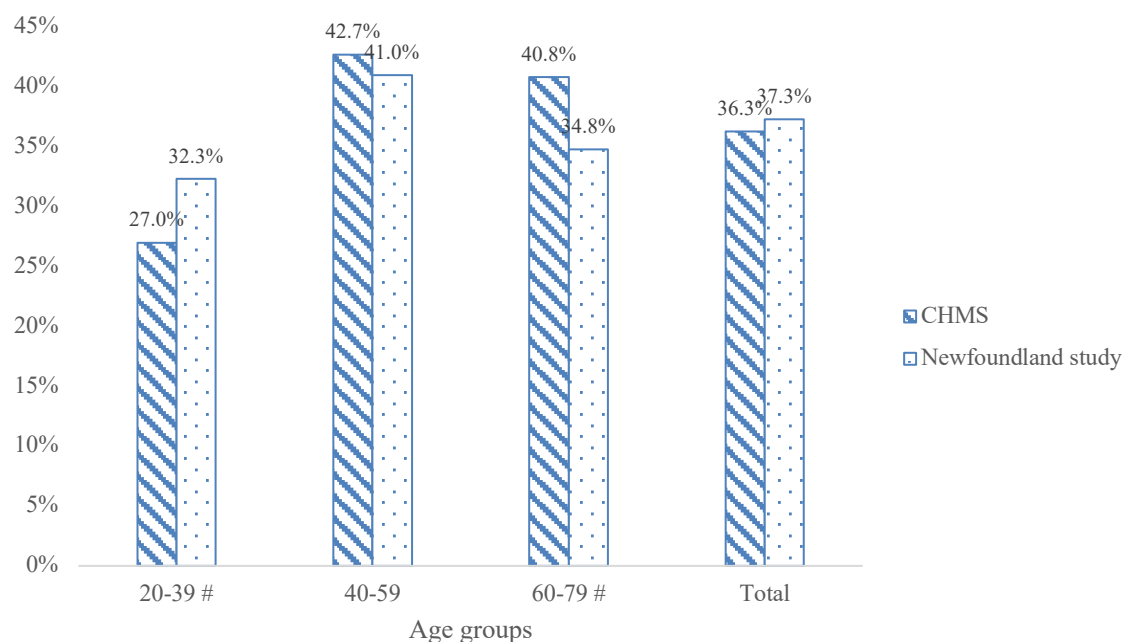


Figure 4-2. Percentage of adults with LDL dyslipidemia by age group. Comparison between NL study (n=3983) and CHMS (n=3508).

P-value for two independent sample proportion test <0.001

Unhealthy level of lipids were significantly different by age group, using Chi Square Test (P-value <0.001), for each component of the lipid profile.

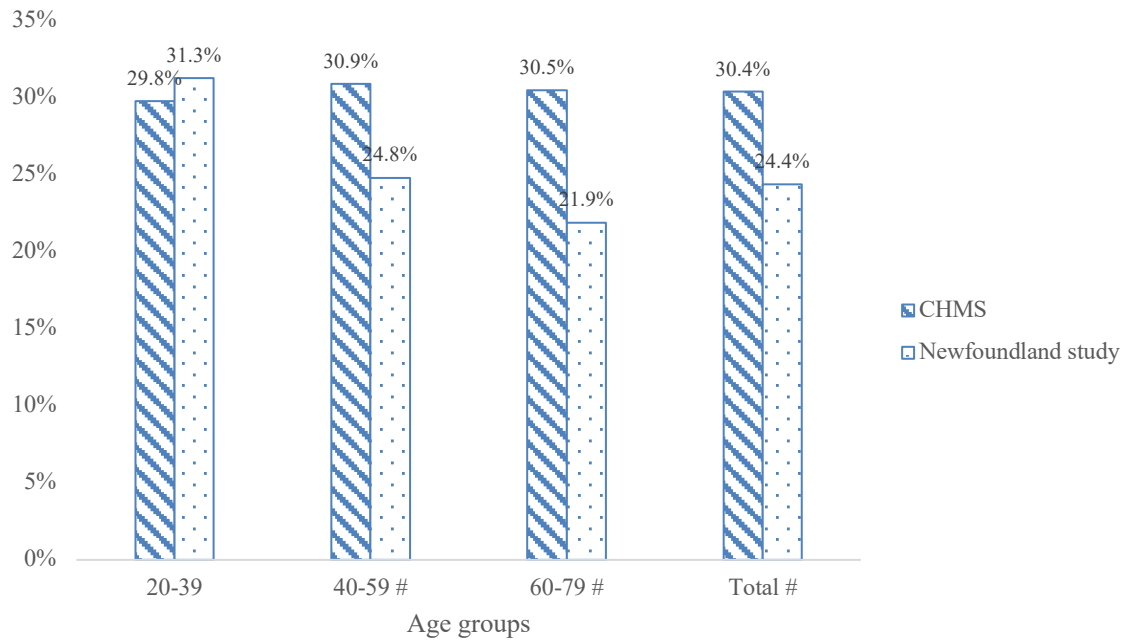


Figure 4-3. Percentage of adults with HDL dyslipidemia by age group. Comparison between NL study (n=3983) and CHMS (n=3508).

P-value for two independent sample proportion test <0.001

Unhealthy level of lipids were significantly different by age group, using Chi Square Test (P-value <0.001), for each component of the lipid profile.

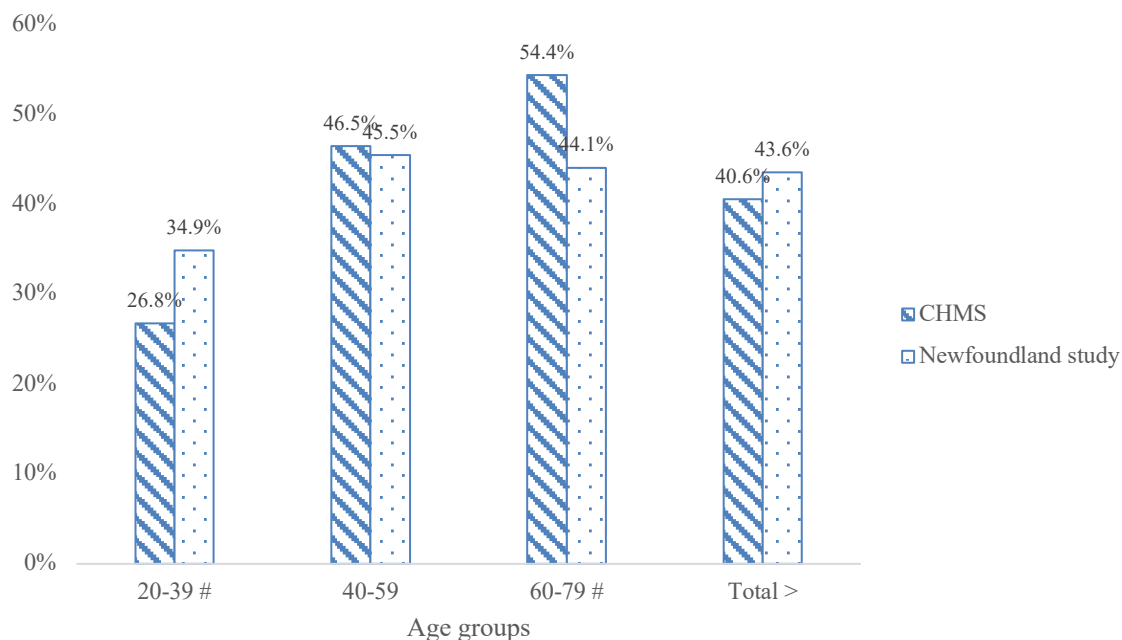


Figure 4-4. Percentage of adults with hypercholesterolemia by age group. Comparison between NL study (n=3983) and CHMS (n=3508).

> P-value for two independent sample proportion test <0.01

P-value for two independent sample proportion test <0.001

Unhealthy level of lipids were significantly different by age group, using Chi Square Test (P-value <0.001), for each component of the lipid profile.

4.5 DISCUSSION

Our study included a large population of individuals in NL and across Canada. We revealed no significant differences between the NL sample and the broader Canadian results on the mean values of the lipid profile components. This would appear to suggest that high rates of young adult mortality from CVD in NL are not due to differences in lipid profiles. There were, however, several differences in the prevalence of dyslipidemia. Specifically, the proportion of the older adults in the NL sample who hypercholesterolemia, hypertriglyceridemia, and LDL dyslipidemia is less than their age counterpart in CHMS. A larger proportion of the youngest patients (aged 20 to 39) in the NL sample had hypercholesterolemia, hypertriglyceridemia, and LDL dyslipidemia. These findings suggest that the lower proportions in elder adults may be a result of earlier cardiovascular death (the leading cause of death in NL [3-5]). A higher proportion of both men and women in the NL sample had HDL dyslipidemia. This result is particularly evident in men.

The NL lipid profile showed that in a sample of 3,983 patients from 4 primary care clinics in NL, 44% of the patients had high total cholesterol, 37% had high LDL, 24% had low HDL, and 26% had high TG [14, 15]. This information can help guide future research about risk factors for CVD and other diseases in NL and Canada. As EMR's become more readily used in medical practice, the method used here could be a complement to the CHMS as the best practice for examining discrepancies in lipid profiles between communities and geographic regions.

The association between lower lipid levels in people with diabetes, hypertension in this study should be interpreted with caution. It was not possible to differentiate between newly diagnosed and prevalent cases of diabetes and hypertension in this cross-sectional study using EMR. Furthermore, we did not have data on the duration of the management they received for their disease and lipid abnormality condition. Another limitation to consider in our multivariate analysis is that far more patients with diabetes will be on statins than other patients.

To compare the lipid profile pattern of NL residents to those of the rest of Canada, we used data from the CHMS which only aggregates data on sex and age. We had no knowledge of the medical histories or risk factors of these participants to allow further comparisons. Further, patients included in the CHMS were also likely healthier as this was a random sample, compared to CPCSSN which includes patients who are seeking care. Also, LDL and TG results from the CHMS were calculated using fasting sub-samples, not full samples. Age distribution of the two samples may also be a limitation to this study. The NL sample included less young adults (12%) compared to CHMS (33%).

Anecdotal evidence from family physicians in this province indicated that there may be a pattern of low HDL among men in this province. The results of this study provide more evidence for this theory and suggest that more research into lipid profiles, especially HDL, needs to be conducted in NL. This study may also shed light on risk factors for higher level of CVD mortality in younger adults in this province; however, these findings should be further investigated.

4.6 CONCLUSION

Analyses of the NL component of the CPSSN database did not show differences in mean level of lipid components compared to the rest of Canada; however, the NL sample was found to have a higher prevalence of low HDL among men and a higher prevalence of unhealthy levels of total cholesterol, TG and LDL in younger adults. It is difficult to know exactly what the lower HDL levels in younger men (and unhealthy levels of lipids in younger adults generally) means for CVD prevalence in the future as these individuals age, but it may be an indication that the high rate of CVD in NL is not likely to change anytime soon. The results of this study provide more evidence for our theory that a different pattern of serum lipids is present in NL residents and suggests more research into lipid profiles is needed, especially with regard to the effects of HDL on cardiovascular health in that population. The limitations of using EMR data for patients who received healthcare services in primary healthcare settings in St. John's, NL, relative to the more representative national CHMS should be considered when interpreting the results of this study. Furthermore, using EMR data may underestimate the true population prevalence of dyslipidemia because it requires the patient to seek out medical care and have their diagnosis recorded in the EMR and coded in the billing data.

4.7 REFERENCES:

1. Public Health Agency of Canada. Tracking heart disease and stroke in Canada. www.phac-aspc.gc.ca/publicat/2009/cvd-avc/pdf/cvd-avs-2009-eng.pdf. Updated 2009. Accessed 09/25, 2017.
2. Heart and Stroke Foundation of Canada. 2010 report - A perfect storm. http://www.healthyenvironmentforkids.ca/sites/healthyenvironmentforkids.ca/files/A_Perfect_Storm.pdf. Updated 2010. Accessed 09/25, 2017.
3. Filate WA, Helen LJ, Courtney CK, Jack VT. Regional variations in cardiovascular mortality in Canada. *Can J Cardiol*. 2003;19(11):1241-8.
4. Manuel DG, Leung M, Nguyen K, Tanuseputro P, Johansen H, Canadian Cardiovascular Outcomes Research Team. Burden of cardiovascular disease in Canada. *Can J Cardiol*. 2003;19(9):997-1004.
5. Statistics Canada. Mortality, summary list of causes. <http://www.statcan.gc.ca/pub/84f0209x/84f0209x2007000-eng.htm>. Updated 2012. Accessed 03/17, 2017.
6. Pope AM, Carr SM, Smith KN, Marshall HD. Mitogenomic and microsatellite variation in descendants of the founder population of Newfoundland: High genetic diversity in an historically isolated population. *Genome*. 2011;54(2):110-119.
7. Fodor JG, Balram BC, Heyden S. Lipid and lipoprotein levels of Newfoundland school children. *Ann Nutr Metab*. 1987;31(4):231-236.
8. Chockalingam A, Fodor JG. Cardiovascular risk factors in Newfoundland population and modification of their level through nonpharmacological intervention. *J Cardiovasc Pharmacol*. 1990;16 Suppl 8:S54-6.

9. MacLean DR, Petrasovits A, Nargundkar M, et al. Canadian heart health surveys: A profile of cardiovascular risk. survey methods and data analysis. Canadian heart health surveys research group. *CMAJ*. 1992;146(11):1969-1974.
10. Statistics Canada. Canadian health measures survey.
<http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=5071&lang=en&db=imdb&adm=8&dis=2>. Updated 2015. Accessed 09/28, 2017.
11. Statistics Canada. Canadian health measures survey.
http://www.statcan.gc.ca/access_acces/archive.action?loc=/pub/82-623-x/82-623-x2010002-eng.pdf&archive=1. Updated 2011. Accessed 09/28, 2017.
12. Statistics Canada. Canadian health measures survey. <http://www.statcan.gc.ca/daily-quotidien/100323/dq100323a-eng.htm>. Updated 2012. Accessed 09/28, 2017.
13. Birtwhistle R, Keshavjee K, Lambert-Lanning A, et al. Building a pan-Canadian primary care sentinel surveillance network: Initial development and moving forward. *J Am Board Fam Med*. 2009;22(4):412-422.
14. Statistics Canada. Population and dwelling counts, for Canada, census metropolitan areas, census agglomerations and census subdivisions (municipalities), 2011 and 2006 censuses. <http://www12.statcan.gc.ca/census-recensement/2011/dp-pd/hltfst/pd-pl/Table-Tableau.cfm?LANG=Eng&T=203&CMA=001&S=0&O=D&RPP=25>. Updated 2016. Accessed 09/28, 2017.
15. Statistics Canada. Heart health and cholesterol levels of Canadians, 2007 to 2009. <http://www.statcan.gc.ca/pub/82-625-x/2010001/article/11136-eng.pdf>. Updated 2010. Accessed 09/25, 2017.

16. Genest J, McPherson R, Frohlich J, et al. 2009 Canadian cardiovascular Society/Canadian guidelines for the diagnosis and treatment of dyslipidemia and prevention of cardiovascular disease in the adult - 2009 recommendations. *Can J Cardiol.* 2009;25(10):567-579.
17. Statistics Canada. Rural and small town Canada analysis bulletin.
<http://www.statcan.gc.ca/pub/21-006-x/21-006-x2001003-eng.pdf>. Updated 2001.
Accessed 09/28, 2017.
18. Public Health Agency of Canada. Report from the Canadian chronic disease surveillance system: Hypertension in Canada, 2010. http://www.phac-aspc.gc.ca/cd-mc/cvd-mcv/ccdss-snsmc-2010/pdf/CCDSS_HTN_Report_FINAL_EN_20100513.pdf. Updated 2010.
Accessed 02/07, 2017.
19. Mantel-Teeuwisse AK, Kloosterman JM, Maitland-van der Zee AH, Klungel OH, Porsius AJ, de Boer A. Drug-induced lipid changes: A review of the unintended effects of some commonly used drugs on serum lipid levels. *Drug Saf.* 2001;24(6):443-456.
20. Kasiske BL, Ma JZ, Kalil RS, Louis TA. Effects of antihypertensive therapy on serum lipids. *Ann Intern Med.* 1995;122(2):133-141.
21. Sharp RP, Sirajuddin R, Sharief IM. Impact of carvedilol on the serum lipid profile. *Ann Pharmacother.* 2008;42(4):564-571.
22. Williams WR, Schneider KA, Borhani NO, Schnaper HW, Slotkoff LM, Ellefson RD. The relationship between diuretics and serum cholesterol in hypertension detection and follow-up program participants. *Am J Prev Med.* 1986;2(5):248-255.

23. IBM corp. released 2013. IBM SPSS statistics for windows, version 22.0. Armonk, NY: IBM corp. <https://www.ibm.com/analytics/data-science/predictive-analytics/spss-statistical-software> Accessed 09/28, 2017.

Chapter 5: Using Electronic Medical Records to Identify Patients with Dyslipidemia in Primary Care Settings: International Classification of Disease Code Matters from One Region to a National Database

A version of this chapter was published in the Journal of Biomedical Informatics Insights, 2017; 9.

5.1 ABSTRACT

The primary objective of this study was to assess the validity of ICD codes for identifying patients with dyslipidemia in EMR data. Secondly, we sought to develop multiple testing algorithms to best identify patients with dyslipidemia. This is a cross-sectional study using the EMRs of patients receiving primary care at family medicine clinics in St. John's, NL, Canada. The data were retrieved from the CPCSSN database. ICD codes were first compared with laboratory lipid data as an independent criterion standard, and next with a "comprehensive criterion standard," defined as any existence of abnormal lipid test, lipid-lowering medication record, or dyslipidemia ICD codes. The ability of ICD coding alone or combined with other components was evaluated against the two criterion standards using ROC analysis, sensitivity, specificity, NPV, PPV and Kappa agreement. A total of 4,382 patients were studied. The ICD codes led to a poor outcome when compared with the serum lipid levels (sensitivity, 27%; specificity, 76%; PPV, 71%; NPV, 33%; Kappa, 0.02; AUC, 51%) or with the comprehensive criterion standard (sensitivity, 32%; NPV, 25%; Kappa, 0.15; AUC, 66%). The addition of laboratory lipid levels to ICD coding marginally improved the algorithm (sensitivity,

94%; NPV, 79%; Kappa, 0.85; AUC, 97%). The use of the ICD coding, either alone or in combination with laboratory data or lipid-lowering medication records, was not an accurate indicator in identifying dyslipidemia.

5.2 INTRODUCTION

Dyslipidemia is one of the most modifiable risk factors for CVD, an important chronic condition which imposes a substantial burden of morbidity and mortality and is the leading cause of death worldwide [1]. As a result, dyslipidemia has been widely studied for projecting CVD population incidence, identifying CVD high-risk groups and evaluating prevention strategies for reducing individual and population risks. Accurate identification of dyslipidemia in the population is crucial to enhance the ability to perform epidemiologic studies including health systems planning, resource allocation, and pharmacoepidemiologic investigations to promote preventive and acute care programs related to CVDs.

Medico-administrative data, recorded according to the ICD coding system, have increasingly been used in large-scale studies in recent years due to higher accessibility and lower costs compared with population-based surveys. These data allow for the passive surveillance of the disease, and are available at lower costs compared to active surveillance, particularly in Canada where a centralized government-based structure in health care exists. As the reliability of the findings from such studies depends on the accuracy of medico-administrative data, studies have attempted to assess the reliability of such coding systems. Although the outcome of these studies has varied according to the

type of data used and the disease under study, unreliability of the ICD codes for the purpose of diagnosis of medical conditions is frequently reported [2-4].

The majority of such studies have been performed on CVDs rather than their risk factors such as dyslipidemia and diabetes. The outcome of these studies have questioned the sensitivity and specificity of medico-administrative record data in identifying the trends of stroke and CVDs [5, 6], although stroke coding has been found to be useful for high-level comparisons, particularly when compared with other diseases [5]. ICD codes have been particularly shown to have restricted potentials for patients with dyslipidemia. This has been shown by the results of few previous studies available using secondary data for lipid research. An algorithm developed by an American study reported that 62.3% of patients with dyslipidemia were not identified by the ICD codes [4]. Another study in a large US medical insurance claims database found that only 15% of laboratory-defined patients had a dyslipidemia diagnosis [7]. In addition, some studies suggest more than one record of the ICD coding during 2 or 3 years to identify patients with a particular disease condition [8-10]. One disadvantage of ICD code for CVDs is the inability to ascertain severity, which is the most important prognostic variable in the surveillance of CVDs. Coding for CVD risk factors as an alternative, however, has the potential to tackle these limitations and enrich the utility of medico-administrative data for surveillance of CVDs. They have been rarely examined using medico-administrative data, and a handful of such studies on dyslipidemia have limitations from being performed using databases that contain no record of other potential identifying markers of dyslipidemia besides the ICD coding, such as the history of lipid-lowering medication use or laboratory lipid levels [7].

The recent emergence of EMRs, however, seems to have eliminated this barrier. Patients' records from a growing number of health providers are being collected in electronic format, which not only provides access to medico-administrative data (e.g., ICD codes) but they also contain information on medical histories, comorbidities, laboratory test results, and medication use [11-13]. The regular management of dyslipidemia is conducted using lipid lowering medications and routine laboratory testing. The structured format of an EMR would, therefore, be ideal for evaluating the accuracy of medico-administrative records compared with other diagnostic criteria. This study examines the degree to which the ICD code alone, or in combination with data on lipid-lowering medications or laboratory lipid levels, can predict a diagnosis of dyslipidemia relative to laboratory data or a more elaborate criterion standard. This investigation is conducted using the multi-disease record surveillance system within the CPCSSN, which contains the ICD codes and lipid-lowering medication records by primary care physicians as well as a link to laboratory data for every record. This strategy is particularly important because not all of the existing EMRs have the entire components of criterion standard algorithms to allow for comparison.

5.3 METHODS

5.3.1 Study Design

This cross-sectional study was designed using the secondary analysis of data from EMRs of primary care clinics in St. John's, NL, Canada. Records of patients with complete lipid profiles undertaken during January 1, 2009 to December 31, 2010 were included.

5.3.2 Study Database

The multi-disease record surveillance system within the CPCSSN is commonly used for chronic disease surveillance in primary care and for conducting primary care research [14-17]. This database contains the EMR records of family physicians which are abstracted quarterly and uploaded to a de-identified system to regional and central pan-Canadian databases. An electronic chart abstraction was performed using the EMRs of clinics in St. John's which form part of the NL component of the CPCSSN [11]. The data for this study comes from three different sections of the EMR:

1. ICD coding for disease diagnosis which is an AutoFill section of the EMR and is completed when the physician selects a disease diagnosis;
2. Laboratory results, which are electronically linked to the Laboratory Information System database. These data are completed at the laboratory and are transferred to the EMR, independent of ICD coding and medication data;
3. Medication prescriptions, which are entered into the EMR by physicians at every visit according to the medication prescribed during the visit.

5.3.3 Study Population

The study population consists of subjects from the NL component of the CPCSSN database aged 20 years or older. Among the total 15,865 patients who received healthcare services, specifically those who saw a physician participating in CPCSSN, during the study timeframe (January 1, 2009-December 31, 2010), 4,382 patients were identified as having had a complete lipid profile taken. Pregnant women were excluded from the analysis.

5.3.4 Algorithm Development and Evaluation

The algorithm validation/testing was performed in the following steps:

1. To determine the performance of ICD coding in comparison with laboratory lipid measurements, as an “independent criterion standard,” ICD coding was compared with the lipid levels from laboratory data.
2. In the second step, a combination of the three criteria (laboratory lipid levels, ICD codes, and lipid-lowering medication use) was used to develop a “comprehensive criterion standard” algorithm to identify any record of patients with dyslipidemia in our database, as follows:
 - a. Any ICD of dyslipidemia recorded (ICD-9-CM 272, disorders of lipid metabolism;
 - b. Any laboratory serum measurements of lipid levels deviating from the cut-offs defined by the Canadian guidelines for the diagnosis and management of dyslipidemias (Table 5.1) [18];
 - c. Any record of using lipid-modifying medications during the study period.

Table 5.1. Healthy Levels of Serum Lipids for Canadian Adults [18]

Lipid Component	Normal Levels
Total cholesterol	< 5.2 mmol/L
TG	< 1.7 mmol/L
LDL	< 3.4 mmol/L
HDL	> 1.0 mmol/L Men > 1.3 mmol/L Women

HDL, high-density lipoprotein; LDL, low-density lipoprotein; TG, triglycerides

Use of lipid-modifying agents (HMG-CoA reductase inhibitors, fibrates, bile acid sequestrants, nicotinic acid, and other agents) was identified using the text record of the medication name and/or Anatomical Therapeutic Chemical codes [19]. Every group was assumed to have dyslipidemia independent of each other. For instance, the patients with normal lipid levels, but with a history of lipid modification use, were categorized as having dyslipidemia because the medication therapy is expected to alter the lipid levels.

Given that the local clinicians had determined that these three criteria would likely detect the significant majority of patients with dyslipidemia in the EMRs, we deemed that the existence of any one or several of these three criteria in an individual would be a criterion standard diagnosis of dyslipidemia. Furthermore, it is common in population screening studies to have results from one or more tests investigating the same condition, none of which can be considered the “criterion standard” alone [20]. In addition, the eMERGE network, a consortium of 5 US institutions linked to secure encrypted EMR data that are designed with the aim of identifying disease phenotypes from EMR, suggest the use of the above three criteria to detect the phenotype of LDL dyslipidemia from EMRs [21,22].

The performance of ICD coding against this comprehensive criterion standard was then examined. Table 5.2 provides a detailed description of the three indicators, as well as the “criterion standard.”

3. The combinations of ICD coding with medication use or laboratory lipid data were compared against the “comprehensive criterion standard.”
4. The above analysis was repeated using the national CPCSSN data between 2010

and 2012 to assess the replicability of the findings

5. In the end, the association of ICD coding with other factors associated with dyslipidemia, including age, sex, diabetes, hypertension, medication use, smoking, and BMI, was examined to determine the factors with the most influence on the ICD coding. We assumed that individuals with different demographics and comorbidities may have variable ICD coding accuracy due to the difference in their management.

Table 5.2. Number of Patients with Dyslipidemia and Associated Prevalence Categorized by Algorithm

	Definition	No. of Cases	Apparent Prevalence (%)
Situation A An abnormal lipid level is reported in laboratory data.	The most recent lipid profile (total cholesterol, HDL, LDL, TG) on an individual showed one component of the lipid profile was not in the normal range as recommended by the Canadian lipid guidelines: total cholesterol >5.2 mmol/L, HDL <1.0 mmol/L, LDL >3.4 mmol/L, and TG >1.7 mmol/L (Statistics Canada, 2011; http://www.statcan.gc.ca/pub/82-625-x/2012001/article/11732-eng.htm)	3,035	69.0
Situation B The individual is on a lipid-lowering drug.	Any record of using a lipid-modifying agent including statins, fibrates, bile acid sequestrants, nicotinic acid and derivatives, and other lipid-modifying agents during the study period or an Anatomical Therapeutic Chemical Classification System (ATC) code C10 for these lipid-modifying agents, (WHO, 2012 - within two years before the date the lipid tests were done; http://www.who.int/classifications/atcddd/en/)	1,556	35.4
Situation C The individual has a diagnosis of abnormal lipids.	There is a diagnosis of a “disorder of lipid metabolism” (ICD code 272) according to ICD code 272 in the EMR; (http://icd9cm.chrisendres.com/index.php?action=child&recordid=2055)	1,147	26.1
Comprehensive criterion standard Any one or more of A, B, and C above.	Patients were deemed to have dyslipidemia if they fitted into either one or more of the Situations above: (A) had one component of the lipid profile not in the normal range recommended by Canadian lipid guidelines; (B) there was record of using a lipid-modifying agent; (C) had an ICD code 272 diagnosis on record	3,573	81.2

HDL, high-density lipoprotein; LDL, low-density lipoprotein; TG, triglycerides

5.3.5 Statistical Analysis

An analysis, using 2 x 2 table formats, was conducted to evaluate the variation in the diagnosis of dyslipidemia. Sensitivity, specificity, PPV, NPV, Kappa agreement, and AUC were calculated for every algorithm in comparison with the “criterion standard”.

Sensitivity was defined as the proportion of patients identified by the testing algorithms who had dyslipidemia according to the “criterion standard”. *Specificity* was defined as the proportion of patients excluded by the testing algorithms who did not have dyslipidemia according to the “criterion standard”. *PPV* was defined as the proportion of patients with dyslipidemia identified by the testing algorithms that were also confirmed by the “criterion standard”. *NPV* was defined similarly for patients who did not have dyslipidemia according to the testing algorithms (Table 5.3). The Kappa agreement was calculated between every testing algorithms and the “criterion standard”. The kappa values of 0 to 0.20, 0.21 to 0.40, 0.41 to 0.60, 0.61 to 0.80, 0.81 to 0.90 and 0.91 to 1.0 indicate poor, slight, fair, good, very good and excellent agreements, respectively [23,24]. A ROC curve for each algorithm was measured against the “criterion standard”. ROC curves were obtained by calculating the sensitivity and specificity of the test and plotting the sensitivity against 1-specificity. AUC of the ROC is a reflection of how reliable the test is in distinguishing between patients with disease and those without disease [25]. The AUC's greater than 0.9 are considered to have high accuracy, whereas an AUC in the range of 0.7 to 0.9 indicates moderate accuracy, 0.5 to 0.7 indicates low accuracy, and 0.5 a chance result [26]. Prevalence was estimated according to the number of patients with dyslipidemia identified by each definition. A logistic regression analysis was performed to determine which factors influenced ICD coding for dyslipidemia. Significance of

effects was evaluated at $\alpha = 0.05$. All of the analyses were conducted using Stata SE 11.2 [27].

Table 5.3. Definitions for sensitivity, specificity, NPV, and PPV

		Criterion standard	
		Dyslipidemia	Healthy Lipid
ICD code 272	Dyslipidemia	A (true-positive)	B (false-positive)
	Healthy Lipid	C (false-negative)	D (true-negative)

Sensitivity: $A/(A+C) \times 100$; specificity: $D/(D+B) \times 100$; positive predictive value:

$A/(A+B) \times 100$; negative predictive value: $D/(D+C) \times 100$

5.3.6 Ethics

The Human Research Ethics Authority, Memorial University of Newfoundland, reviewed and approved the study protocol, reference number 11.090. All the data were de-identified prior to the analysis.

5.4 RESULTS

The EMRs from a total of 4,382 patients (mean age, 58.1 ± 14.8 years, 58.8% females) were included in the study. The population had a BMI of 31.1 ± 15.8 , 42.3% of whom were present/former smokers. The prevalence of hypertension and diabetes was 33.5% ($n=1468$) and 15.2% ($n=666$), respectively. Among this population, 3,573 patients

had dyslipidemia during the study period according to the “comprehensive criterion standard” definition (prevalence of 81.2%, n=3,573). As shown in Table 5.2, among all patients, 69.0% (n=3,035) were diagnosed with dyslipidemia according to laboratory results (independent criterion standard), 26.1% (n=1,147) had an ICD coding for dyslipidemia, and 35.4% (n=1,556) had used one or more lipid-lowering medication. The overlap of these three components is shown as a Venn diagram in Figure 5-1.

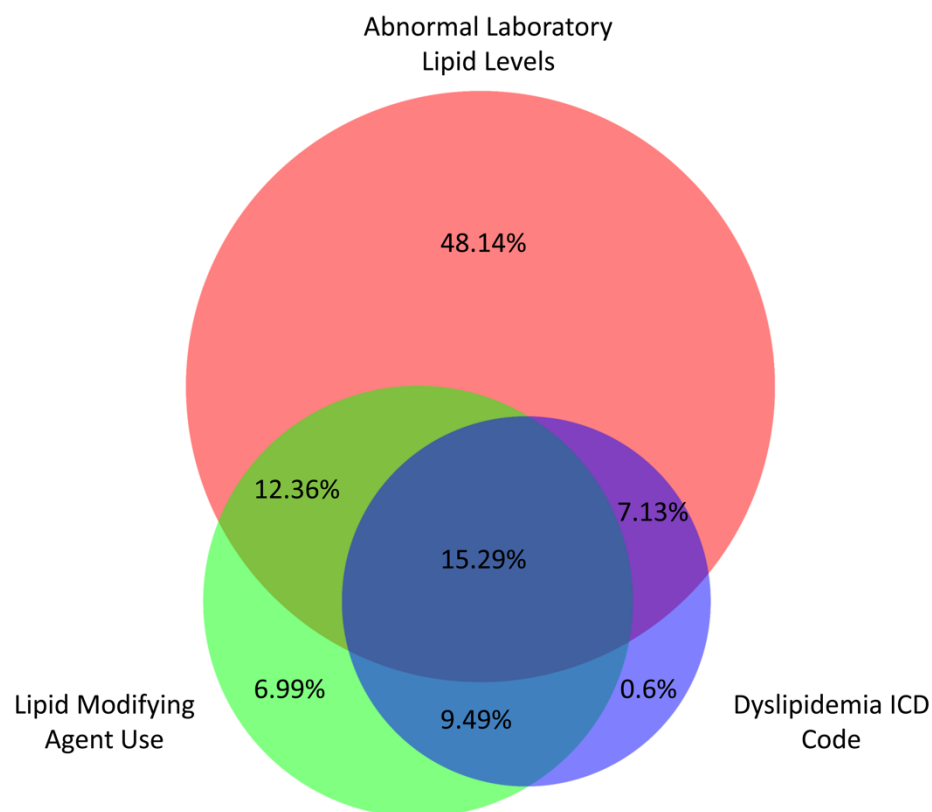


Figure 5-1. Venn diagram of the three components of the criterion standard algorithm (n=4382).

The ICD codes resulted in a poor outcome when compared with the independent criterion standard (serum lipid levels). This analysis led to a sensitivity of 27.0%,

specificity of 76.7%, PPV of 71.1%, NPV of 33.1%, a Kappa agreement of 0.02 and an AUC of 0.51.

In the second attempt, Situation C was compared with the comprehensive criterion standard as shown in Table 5.4. Situation C led to the lowest sensitivity (32.1%), NPV (25.4%), Kappa agreement (0.151), and AUC (66.1%) compared with the “comprehensive criterion standard.”

Table 5.4. Sensitivity, Specificity, and Predictive Values of all Combinations of Situations A, B, and C as Compared with Comprehensive Criterion Standard

Situation	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Kappa Value	AUC (95% CI)
A	85	100	100	60	0.68	0.92 (0.92-0.93)
B	43	100	100	29	0.23	0.72 (0.71-0.72)
C	32	100	100	25	0.15	0.66 (0.65-0.67)
A&B	99	100	100	98	0.99	1.00 (0.99-1.00)
A&C	94	100	100	79	0.85	0.97 (0.96-0.97)
B&C	51	100	100	32	0.28	0.76 (0.75-0.76)

Positive Predictive Value, PPV; Negative Predictive Value, NPV; Area under the receiver operating characteristic curve, AUC; Confidence Interval, CI.

Situation B&C, also yielded a poor result. This algorithm resulted in a low sensitivity (51.2%), NPV (32.2%), Kappa agreement (0.283), and AUC (75.6%) (Table 5.4). Situation A (Table 5.4) had the highest sensitivity (84.9%), NPV (60.6%), Kappa agreement (0.680), and AUC (92.5%) compared with Situation C or Situation B on their own. Situation A&B further increased the sensitivity (99.6%), NPV (98.1%), Kappa agreement (0.988), and AUC (99.8%) (Table 5.4).

To replicate our results, we assessed the repeatability of our findings using the Canada-wide records of 2010-2012 in a similar approach. This analysis also showed the lowest sensitivity (32.1%), NPV (26.0%), Kappa agreement (0.15), and AUC (0.66) for Situation C compared with the “comprehensive criterion standard.” Situation B&C also yielded a low sensitivity (51.2%), NPV (32.2%), Kappa agreement (0.28), and AUC (0.76).

Given that the use of ICD coding is not reliable in identifying dyslipidemia in EMRs, an additional analysis was conducted using logistic regression to explore which demographic factors and co-morbidities may influence the ICD coding for dyslipidemia (Table 5.5). Results from this analysis showed that patients prescribed lipid-lowering medication were very likely (OR, 9.75; 95% CI, 6.82-13.95; $p < 0.001$) to have the ICD codes for dyslipidemia.

Table 5.5. Factors Associated with ICD Coding for Dyslipidemia

Variable	Odds Ratio (95% Confidence Interval)
ICD Code	
Sex (Male)	1.318 (0.943-1.843)
Aged 41-64 [^]	2.921 (1.060-8.050)*
Aged ≥65 [^]	4.276 (1.521-12.021)*
Hypertension	1.502 (1.062-2.125)*
Diabetes	1.299 (0.847-1.990)
Former/Current Smoker	1.040 (0.876-1.233)
BMI ≥30	1.044 (0.734-1.484)
Lipid Lowering Medication User	9.754 (6.821-13.947) [#]

BMI, body mass index; ICD, International Classification of Disease.

[^]Reference age category 20-40

*p<0.05

[#]p<0.001

5.5 DISCUSSION

This study has demonstrated that using the ICD coding alone is an unreliable indicator of dyslipidemia. The ICD coding data represented a substantial underestimation of dyslipidemia cases. The use of ICD codes in combination with data from laboratory results or lipid-lowering medication added only marginal value to the respective algorithms. In addition, the ICD coding data alone yielded the most false-negatives. The ICD codes were also unreliable when compared with serum lipid levels alone as an independent criterion standard.

Although the ICD codes are reported to be able to accurately identify patients with

many medical conditions in administrative health data such as ischemic heart disease [28], diabetes mellitus [29], and preeclampsia [30], their potential for patients with dyslipidemia is restricted. Our results are consistent with the few previous studies available using secondary data for lipid research. In support of our notion regarding the inaccuracy and unreliability of using the ICD coding data for lipid research, we learned that an American study created an algorithm for detecting dyslipidemia and diabetes. The algorithm identified 58.4% of patients with hyperlipidemia, 62.3% of whom were not recorded as having dyslipidemia in accordance with the ICD codes [4]. Another study in a large US medical insurance claims database found that only 15% of laboratory-defined patients had a dyslipidemia diagnosis [11]. In Alberta, Canada, Kokotailo and Hill [31] showed that although the medico-administrative billing system is a good indicator of stroke and some of its risk factors including diabetes mellitus and hypertension, the identification of hyperlipidemia is not confidently made where the sensitivity was reported to be 57%. The exact reason for incomplete coding of dyslipidemia is unclear. Kokotailo and Hill [31] considered “a lack of perceived importance by the physician, or a lack of time to code everything”, as the putative reason.

The use of advanced technologies in disease coding may be a solution to this problem and to improve the accuracy of ICD codes. Further, the number of digits that can be captured and having the most up-to-date version of ICD is also important. Wockenfuss et al. investigated the reliability of the ICD-10 in primary care. They determined that three- and four-digit ICD-10 was not reliable in primary care. It was reliable only at the chapter level [32]. Natural language processing, a range of computational techniques for analyzing written or oral texts for the purpose of achieving human-like language

processing, has been applied to the EMRs and have shown to improve the accuracy of case definition for inflammatory bowel disease [33], venous thromboembolic disease [34], and cancer [35]. Multimodal fusion/interaction are multiple modes of interaction with a system which provides several distinct tools for input and output of data. This technique has been implemented in different aspects of medical diagnosis including the processing of brain imaging [36] and magnetic resonance imaging (MRI) [37] data, as well as discriminative learning for Alzheimer's disease diagnosis [38]. This, however, has rarely been implemented in disease coding and EMR processing. The use of this method might have a potential for improving the disease coding in medical administrative data.

Consideration ought to be given to possible limitations when interpreting and applying these data. The possibility of information bias and data inaccuracy, despite the fact that the direct link between Laboratory Information System and EMRs should decrease the probability of data entry errors. Also, these results are based on ICD version 9.0. Newer versions of ICD, including ICD 10.0 and ICD 11.0, have been released. It is notable that the case definition for dyslipidemia does not change considerably between these versions, and thus, the findings of this study may be able to be applied to newer versions of ICD.

Limitations of this article include the representativeness of the study as this study focused only on the data from EMR clinics in St. John's, NL, Canada, to assess the validity of the ICD coding in identifying patients with dyslipidemia. In addition, our data only apply to primary care, and it may not be extended to hospital-based and specialized care where more severe and acute cases of CVDs and dyslipidemia exist.

5.6 CONCLUSION

Using secondary data to identify patients diagnosed with dyslipidemia could involve information on laboratory values, lipid-lowering medication data, or diagnostic data. Often, a given EMR will have only one of these pieces of information. Results from laboratory data may only have levels of lipids, pharmacy data may only have prescription records, and provincial billing databases may only have diagnostic data. Databases that contain all three of these (lipid levels, medications, and diagnoses) can be used to understand how either one or any two of these pieces of information can predict whether dyslipidemia exists in an individual. The CPCSSN database contains all three of these types of information.

Although the ICD codes have typically been used for the diagnosis of many medical conditions in both research and practice, our research suggests that they are not an accurate indicator of patients with dyslipidemia. Therefore, caution ought to be taken into account when using the databases established according to the ICD codes for research involving dyslipidemia.

5.7 REFERENCES

1. World Health Organization. Global status report on noncommunicable diseases 2010. Geneva, Switzerland.

2. Birman-Deych E, Waterman AD, Yan Y, Nilasena DS, Radford MJ, Gage BF. Accuracy of ICD-9-CM codes for identifying cardiovascular and stroke risk factors. *Med Care*. 2005;43(5):480-485.
3. O'Malley KJ, Cook KF, Price MD, Wildes KR, Hurdle JF, Ashton CM. Measuring diagnoses: ICD code accuracy. *Health Serv Res*. 2005;40(5 Pt 2):1620-1639.
4. Kandula S, Zeng-Treitler Q, Chen L, Salomon WL, Bray BE. A bootstrapping algorithm to improve cohort identification using structured data. *J Biomed Inform*. 2011;44 Suppl 1:S63-8.
5. Tirschwell DL, Longstreth WT, Jr. Validating administrative data in stroke research. *Stroke*. 2002;33(10):2465-2470.
6. Mayo NE, Goldberg MS, Levy AR, Danys I, Korner-Bitensky N. Changing rates of stroke in the province of Quebec, Canada: 1981-1988. *Stroke*. 1991;22(5):590-595.
7. Li J, Motsko SP, Goehring EL, Jr, Vendiola R, Maneno M, Jones JK. Longitudinal study on pediatric dyslipidemia in population-based claims database. *Pharmacoepidemiol Drug Saf*. 2010;19(1):90-98.
8. Juurlink D, Preya C, Croxford R, Chong A, Austin P, Tu J, Laupacis A. Canadian institute for health information discharge abstract database: A validation study. Toronto, ON: Institute for clinical evaluative sciences. 2006.
9. Lix L, Yogendran M, Burchill C, Metge C, McKeen N, Moore D, Bond R. Defining and validating chronic diseases: An administrative data approach. Winnipeg, MB: Manitoba centre for health policy. 2006.

10. Tu K, Mitiku T, Guo H, Lee DS, Tu JV. Myocardial infarction and the validation of physician billing and hospitalization data using electronic medical records. *Chronic Dis Can.* 2010;30(4):141-146.
11. Birtwhistle R, Keshavjee K, Lambert-Lanning A, et al. Building a pan-Canadian primary care sentinel surveillance network: Initial development and moving forward. *J Am Board Fam Med.* 2009;22(4):412-422.
12. Crawford AG, Cote C, Couto J, et al. Prevalence of obesity, type II diabetes mellitus, hyperlipidemia, and hypertension in the united states: Findings from the GE centricity electronic medical record database. *Popul Health Manag.* 2010;13(3):151-161.
13. Young WB, Ryu H. Secondary data for policy studies: Benefits and challenges. *Policy, Politics, & Nursing Practice.* 2000;1(4):302-307.
14. Aref-Eshghi E, Leung J, Godwin M, et al. Low density lipoprotein cholesterol control status among Canadians at risk for cardiovascular disease: Findings from the Canadian primary care sentinel surveillance network database. *Lipids Health Dis.* 2015;14:60-015-0056-8.
15. Nicholson K, Stewart M, Thind A. Examining the symptom of fatigue in primary care: A comparative study using electronic medical records. *J Innov Health Inform.* 2015;22(1):235-243.
16. Biro S, Barber D, Williamson T, Morkem R, Khan S, Janssen I. Prevalence of toddler, child and adolescent overweight and obesity derived from primary care electronic medical records: An observational study. *CMAJ Open.* 2016;4(3):E538-E544.

17. Singer A, Yakubovich S, Kroeker AL, Dufault B, Duarte R, Katz A. Data quality of electronic medical records in Manitoba: Do problem lists accurately reflect chronic disease billing diagnoses? *J Am Med Inform Assoc*. 2016;23(6):1107-1112.
18. Genest J, McPherson R, Frohlich J, et al. 2009 Canadian cardiovascular Society/Canadian guidelines for the diagnosis and treatment of dyslipidemia and prevention of cardiovascular disease in the adult - 2009 recommendations. *Can J Cardiol*. 2009;25(10):567-579.
19. World Health Organization. The Anatomical Therapeutic Chemical classification system. <http://www.who.int/classifications/atcddd/en/>. Updated 2003. Accessed 04/01, 2018.
20. Joseph L, Gyorkos TW, Coupal L. Bayesian estimation of disease prevalence and the parameters of diagnostic tests in the absence of a gold standard. *Am J Epidemiol*. 1995;141(3):263-272.
21. McCarty CA, Chisholm RL, Chute CG, et al. The eMERGE network: A consortium of biorepositories linked to electronic medical records data for conducting genomic studies. *BMC Med Genomics*. 2011;4:13-8794-4-13.
22. Rasmussen-Torvik LJ, Pacheco JA, Wilke RA, et al. High density GWAS for LDL cholesterol in African Americans using electronic medical records reveals a strong protective variant in APOE. *Clin Transl Sci*. 2012;5(5):394-399.
23. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. *J Clin Epidemiol*. 1993;46(5):423-429.
24. Szklo M NF. *Epidemiology: Beyond the basics*. 2nd edition. Jones and Bartlett Publishers; 2007.

25. Vining DJ, Gladish GW. Receiver operating characteristic curves: A basic understanding. *Radiographics*. 1992;12(6):1147-1154.
26. Fischer JE, Bachmann LM, Jaeschke R. A readers' guide to the interpretation of diagnostic test properties: Clinical example of sepsis. *Intensive Care Med*. 2003;29(7):1043-1051.
27. Stata SE 11. <http://www.stata.com>. Updated 2015. Accessed 07/13, 2016.
28. Cheng CL, Kao YH, Lin SJ, Lee CH, Lai ML. Validation of the national health insurance research database with ischemic stroke cases in Taiwan. *Pharmacoepidemiol Drug Saf*. 2011;20(3):236-242.
29. Chen G, Khan N, Walker R, Quan H. Validating ICD coding algorithms for diabetes mellitus from administrative data. *Diabetes Res Clin Pract*. 2010;89(2):189-195.
30. Geller SE, Ahmed S, Brown ML, Cox SM, Rosenberg D, Kilpatrick SJ. International classification of diseases-9th revision coding for preeclampsia: How accurate is it? *Am J Obstet Gynecol*. 2004;190(6):1629-33; discussion 1633-4.
31. Kokotailo RA, Hill MD. Coding of stroke and stroke risk factors using international classification of diseases, revisions 9 and 10. *Stroke*. 2005;36(8):1776-1781.
32. Wockenfuss R, Frese T, Herrmann K, Claussnitzer M, Sandholzer H. Three- and four-digit ICD-10 is not a reliable classification system in primary care. *Scand J Prim Health Care*. 2009;27(3):131-136.
33. Ananthakrishnan AN, Cai T, Savova G, et al. Improving case definition of crohn's disease and ulcerative colitis in electronic medical records using natural language processing: A novel informatics approach. *Inflamm Bowel Dis*. 2013;19(7):1411-1420.

34. McPeck Hinz ER, Bastarache L, Denny JC. A natural language processing algorithm to define a venous thromboembolism phenotype. *AMIA Annu Symp Proc.* 2013;2013:975-983.
35. Strauss JA, Chao CR, Kwan ML, Ahmed SA, Schottinger JE, Quinn VP. Identifying primary and recurrent cancers using a SAS-based natural language processing algorithm. *J Am Med Inform Assoc.* 2013;20(2):349-355.
36. Sui J, Adali T, Yu Q, Chen J, Calhoun VD. A review of multivariate methods for multimodal fusion of brain imaging data. *J Neurosci Methods.* 2012;204(1):68-81.
37. Meng X, Jiang R, Lin D, et al. Predicting individualized clinical measures by a generalized prediction framework and multimodal fusion of MRI data. *Neuroimage.* 2017;145(Pt B):218-229.
38. Lei B, Chen S, Ni D, Wang T. Discriminative learning for alzheimer's disease diagnosis via canonical correlation analysis and multimodal fusion. *Front Aging Neurosci.* 2016;8:77.

Chapter 6: Summary

This thesis focused on describing lipid profiles and the prevalence of dyslipidemia in NL. A major impetus for this current research project is the prevalence and pattern of dyslipidemia in NL and the lipid profiles, in general, in people from this province. Dyslipidemia is an accepted risk factor for CVD. NL has a higher level of CVD mortality than any other Canadian province [1,2]. According to Statistics Canada, approximately 40% of Canadians have elevated cholesterol [3], while as many as 10 million Canadian adults have a cholesterol level higher than the recommended target [4]. Several studies have found a high prevalence for hypercholesterolemia in NL [5-7]. These studies also suggest that there might be a different pattern of dyslipidemia in NL compared to other Canadian provinces. There are factors that may explain this difference in lipid profiles; one is the homogeneity of both genetics and culture in NL that is not present in most other provinces.

The project had several key findings and outcomes. First was the finding that a significant number of patients in primary care settings in NL have evidence of single and mixed dyslipidemia. Results from the NL subset of CPCSSN showed that nearly one of every five patients in primary care settings in NL have mixed dyslipidemia. This cohort of patients is at high risk for developing CVDs [8-10].

Once we described the prevalence of single and mixed dyslipidemia in a primary care setting in NL, we subsequently investigated whether there were lipid profile differences in NL compared to Canadian lipid profile data from the CHMS. Our findings did not show differences in mean level of lipid components compared

to the rest of Canada; however, the NL sample was found to have a higher prevalence of low HDL among men and a higher prevalence of unhealthy levels of total cholesterol, TG and LDL in younger adults. It is difficult to know exactly what the lower HDL levels in younger men (and unhealthy levels of lipids in younger adults generally) means for CVD prevalence in the future as these individuals age, but it may be an indication that the high rate of CVD in NL is not likely to change anytime soon. The results of this study provided evidence for our hypothesis that a different pattern of serum lipids is present in NL residents and suggests more research into lipid profiles is needed, especially with regard to the effects of HDL on cardiovascular health in that population.

Secondary analysis of existing medico-administrative data is being increasingly utilized in medical research and provides valuable sources of information [11,12]. Given that the sensitivity and specificity of using medico-administrative data to identify dyslipidemia [12], and CVD [13,14], have been questioned, we sought to assess the validity of ICD codes for identifying patients with dyslipidemia. Our findings suggested that while ICD codes have typically been used in both the research and practice of many medical conditions, they are not an accurate indicator of patients with dyslipidemia. Given these inaccurate findings we then sought to develop multiple testing algorithms to best identify patients with dyslipidemia. Ultimately, combining laboratory lipid results together with lipid-lowering medication data resulted in the best sensitivity, NPV, Kappa agreement, and AUC.

EMRs are used extensively throughout North America. They are a longitudinal electronic record of patient health information and allow access to detailed patient information, assist in chronic disease management, include patient medications and past medical history, and facilitate disease coding for billing and disease demographics [15]. Although EMRs are an excellent tool for research as they can provide a vast amount of clinical information, challenges do exist in interpreting and utilizing that data. Limitations such as the level or lack of quality control over data, the possibility of having missing items or missing records, and timelines of the data collected need to be considered.

Translating the knowledge gained from this project can help policymakers and healthcare providers better understand the magnitude of dyslipidemia as an important risk factor for CVD in NL. Furthermore, it will provide information to support the development of policies to decrease the high prevalence of major chronic diseases, to improve healthcare, and introduce quality improvement initiatives to the health care system that will contribute to population health from adults to seniors.

6.1 REFERENCES

1. Filate WA, Helen LJ, Courtney CK, Jack VT. Regional variations in cardiovascular mortality in Canada. *Can J Cardiol*. 2003;19(11):1241-8.
2. Statistics Canada. Mortality, summary list of causes.
<http://www.statcan.gc.ca/pub/84f0209x/84f0209x2007000-eng.htm>. Updated 2007.
Accessed 09/25, 2017.

3. Statistics Canada. Canadian health measures survey.
http://www.statcan.gc.ca/access_acces/archive.action?loc=/pub/82-623-x/82-623-x2010002-eng.pdf&archive=1. Updated 2011. Accessed 09/28, 2017.
4. Heart and Stroke Foundation of Canada.
<http://www.heartandstroke.com/site/c.ikIQLcMWJtE/b.3483991/k.34A8/Statistics.htm#bloodcholesterol>. Updated 2011. Accessed 09/29, 2017.
5. Fodor JG, Balram BC, Heyden S. Lipid and lipoprotein levels of Newfoundland school children. *Ann Nutr Metab*. 1987;31(4):231-236.
6. Chockalingam A, J GF. Cardiovascular risk factors in Newfoundland population and modification of their level through nonpharmacological intervention. *J Cardiovasc Pharmacol*. 1990;16 Suppl 8:S54-6.
7. Fodor JG, Chockalingam A. Modification of cardiovascular risk factors in rural population of Newfoundland through nutritional counselling. *Cor Vasa*. 1991;33(2):107-13.
8. Gordon T, Castelli WP, Hjortland MC, Kannel WB, Dawber TR. High density lipoprotein as a protective factor against coronary heart disease. The framingham study. *Am J Med*. 1977;62(5):707-714.
9. Gordon DJ, Probstfield JL, Garrison RJ, et al. High-density lipoprotein cholesterol and cardiovascular disease. Four prospective American studies. *Circulation*. 1989;79(1):8-15.
10. Sarwar N, Danesh J, Eiriksdottir G, et al. Triglycerides and the risk of coronary heart disease: 10,158 incident cases among 262,525 participants in 29 western prospective studies. *Circulation*. 2007;115(4):450-458.

11. Birtwhistle R, Keshavjee K, Lambert-Lanning A, et al. Building a pan-Canadian primary care sentinel surveillance network: Initial development and moving forward. *J Am Board Fam Med*. 2009;22(4):412-422.
12. Crawford AG, Cote C, Couto J, et al. Comparison of GE centricity electronic medical record database and national ambulatory medical care survey findings on the prevalence of major conditions in the United States. *Popul Health Manag*. 2010;13(3):139-150.
13. Tirschwell DL, Longstreth WT, Jr. Validating administrative data in stroke research. *Stroke*. 2002;33(10):2465-2470.
14. Mayo NE, Goldberg MS, Levy AR, Danys I, Korner-Bitensky N. Changing rates of stroke in the province of Quebec, Canada: 1981-1988. *Stroke*. 1991;22(5):590-595.
15. Menachemi N, Collum TH. Benefits and drawbacks of electronic health record systems. *Risk Manag Healthc Policy*. 2011;4:47-55.

Chapter 7: Overall Thesis References (Alphabetical)

1. Aguilar-Salinas CA, Gomez-Perez FJ, Rull J, Villalpando S, Barquera S, Rojas R. Prevalence of dyslipidemias in the Mexican national health and nutrition survey 2006. *Salud Publica Mex.* 2010;52 Suppl 1:S44-53.
2. Alaghebandan R. *Assessing administrative databases for surveillance of depressive disorders in Newfoundland and Labrador.* Memorial University of Newfoundland, Newfoundland; 2010.
3. Ananthakrishnan AN, Cai T, Savova G, et al. Improving case definition of crohn's disease and ulcerative colitis in electronic medical records using natural language processing: A novel informatics approach. *Inflamm Bowel Dis.* 2013;19(7):1411-1420.
4. Anderson TJ, Gregoire J, Pearson GJ, et al. 2016 Canadian Cardiovascular Society Guidelines for the Management of Dyslipidemia for the Prevention of Cardiovascular Disease in the Adult. *Can J Cardiol.* 2016;32(11):1263-1282.
5. Arca M, Gaspardone A. Atorvastatin efficacy in the primary and secondary prevention of cardiovascular events. *Drugs.* 2007;67 Suppl 1:29-42.
6. Aref-Eshghi E, Leung J, Godwin M, et al. Low density lipoprotein cholesterol control status among Canadians at risk for cardiovascular disease: Findings from the Canadian primary care sentinel surveillance network database. *Lipids Health Dis.* 2015;14:60-015-0056-8.
7. Asghari S, Aref-Eshghi E, Godwin M, Duke P, Williamson T, Mahdavian M. Single and mixed dyslipidaemia in Canadian primary care settings: Findings from the

Canadian primary care sentinel surveillance network database. *BMJ Open*. 2015;5(12):e007954-2015-007954.

8. Benjamin EJ, Blaha MJ, Chiuve SE, et al. Heart disease and stroke statistics-2017 update: A report from the American heart association. *Circulation*. 2017;135(10):e146-e603.
9. Birman-Deych E, Waterman AD, Yan Y, Nilasena DS, Radford MJ, Gage BF. Accuracy of ICD-9-CM codes for identifying cardiovascular and stroke risk factors. *Med Care*. 2005;43(5):480-485.
10. Biro S, Barber D, Williamson T, Morkem R, Khan S, Janssen I. Prevalence of toddler, child and adolescent overweight and obesity derived from primary care electronic medical records: An observational study. *CMAJ Open*. 2016;4(3):E538-E544.
11. Birtwhistle R, Keshavjee K, Lambert-Lanning A, et al. Building a pan-Canadian primary care sentinel surveillance network: Initial development and moving forward. *J Am Board Fam Med*. 2009;22(4):412-422.
12. Biskupiak JE, Kim J, Phatak H, Wu D. Prevalence of high-risk cardiovascular conditions and the status of hypertension management among hypertensive adults 65 years and older in the United States: Analysis of a primary care electronic medical records database. *J Clin Hypertens*. 2010;12(12):935-944.
13. Bruckert E, Baccara-Dinet M, Eschwege E. Low HDL-cholesterol is common in European type 2 diabetic patients receiving treatment for dyslipidaemia: Data from a pan-European survey. *Diabet Med*. 2007;24(4):388-391.

14. Bruckert E, Baccara-Dinet M, McCoy F, Chapman J. High prevalence of low HDL-cholesterol in a pan-European survey of 8545 dyslipidaemic patients. *Curr Med Res Opin.* 2005;21(12):1927-1934.
15. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. *J Clin Epidemiol.* 1993;46(5):423-429.
16. Canada Health Infoway. EMR, EHR, and PHR – why all the confusion? <http://infowayconnects.infoway-inforoute.ca/blog/electronic-health-records/374-emr-ehr-and-phr---why-all-the-confusion/#axzz1LCQ7XyAm>. Updated 2011. Accessed 09/28, 2017.
17. Canada Health Infoway. Laboratory information systems. <https://www.infoway-inforoute.ca/about-infoway/approach/investment-programs/laboratory-information-systems>. Updated 2011. Accessed 09/28, 2017.
18. Cannon CP. Mixed dyslipidemia, metabolic syndrome, diabetes mellitus, and cardiovascular disease: Clinical implications. *Am J Cardiol.* 2008;102(12A):5L-9L.
19. Carleton RA, Dwyer J, Finberg L, et al. Report of the expert panel on population strategies for blood cholesterol reduction. A statement from the national cholesterol education program, national heart, lung, and blood institute, national institutes of health. *Circulation.* 1991;83(6):2154-2232.
20. Centers for disease control and prevention, national center for health statistics. Multiple cause of death 1999-2015 on CDC WONDER online. <https://wonder.cdc.gov/mcd-icd10.html>. Updated 2016. Accessed 10/01, 2017.
21. Chen G, Khan N, Walker R, Quan H. Validating ICD coding algorithms for diabetes mellitus from administrative data. *Diabetes Res Clin Pract.* 2010;89(2):189-195.

22. Cheng CL, Kao YH, Lin SJ, Lee CH, Lai ML. Validation of the national health insurance research database with ischemic stroke cases in Taiwan. *Pharmacoepidemiol Drug Saf.* 2011;20(3):236-242.
23. Chockalingam A, J GF. Cardiovascular risk factors in Newfoundland population and modification of their level through nonpharmacological intervention. *J Cardiovasc Pharmacol.* 1990;16 Suppl 8:S54-6.
24. Colquhoun D, Chirovsky D, Sazonov V, Cui YA, Ambegaonkar B. Prevalence of mixed dyslipidemia among Australian patients undergoing lipid-modifying therapy. *Exp Clin Cardiol.* 2013;18(1):e32-6.
25. CPCSSN disease definitions. <http://cpcssn.ca/research-resources/case-definitions/>. Updated 2016. Accessed 04/01, 2018.
26. Crawford AG, Cote C, Couto J, et al. Prevalence of obesity, type II diabetes mellitus, hyperlipidemia, and hypertension in the United States: Findings from the GE centricity electronic medical record database. *Popul Health Manag.* 2010;13(3):151-161.
27. Crawford AG, Cote C, Couto J, et al. Comparison of GE centricity electronic medical record database and national ambulatory medical care survey findings on the prevalence of major conditions in the United States. *Popul Health Manag.* 2010;13(3):139-150.
28. Devine S, West S, Andrews E, et al. The identification of pregnancies within the general practice research database. *Pharmacoepidemiol Drug Saf.* 2010;19(1):45-50.
29. Eaton CB. Hyperlipidemia. *Prim Care.* 2005;32(4):1027-55, viii.

30. Esteban-Vasallo MD, Dominguez-Berjon MF, Astray-Mochales J, et al. Epidemiological usefulness of population-based electronic clinical records in primary care: Estimation of the prevalence of chronic diseases. *Fam Pract.* 2009;26(6):445-454.
31. Esteghamati A, Meysamie A, Khalilzadeh O, et al. Third national surveillance of risk factors of non-communicable diseases (SuRFNCD-2007) in Iran: Methods and results on prevalence of diabetes, hypertension, obesity, central obesity, and dyslipidemia. *BMC Public Health.* 2009;9:167.
32. Fesharakinia A, Zarban A, Sharifzadeh GR. Lipid profiles and prevalence of dyslipidemia in schoolchildren in south Khorasan province, eastern Iran. *Arch Iran Med.* 2008;11(6):598-601.
33. Filate WA, Helen LJ, Courtney CK, Jack VT. Regional variations in cardiovascular mortality in Canada. *Can J Cardiol.* 2003;19(11):1241-8.
34. Fischer JE, Bachmann LM, Jaeschke R. A readers' guide to the interpretation of diagnostic test properties: Clinical example of sepsis. *Intensive Care Med.* 2003;29(7):1043-1051.
35. Fodor JG, Chockalingam A. Modification of cardiovascular risk factors in rural population of Newfoundland through nutritional counselling. *Cor Vasa.* 1991;33(2):107-13.
36. Fodor JG, Balram BC, Heyden S. Lipid and lipoprotein levels of Newfoundland school children. *Ann Nutr Metab.* 1987;31(4):231-236.
37. Franssen R, Monajemi H, Stroes ESG, Kastelein JJP. Obesity and dyslipidemia. *Endocrinology & Metabolism Clinics of North America.* 2008;37(3):623-633.

38. Fuentes R, Uusitalo T, Puska P, Tuomilehto J, Nissinen A. Blood cholesterol level and prevalence of hypercholesterolaemia in developing countries: A review of population-based studies carried out from 1979 to 2002. *Eur J Cardiovasc Prev Rehabil.* 2003;10(6):411-419.
39. Geller SE, Ahmed S, Brown ML, Cox SM, Rosenberg D, Kilpatrick SJ. International classification of diseases-9th revision coding for preeclampsia: How accurate is it? *Am J Obstet Gynecol.* 2004;190(6):1629-33; discussion 1633-4.
40. Genest J, McPherson R, Frohlich J, et al. 2009 Canadian cardiovascular Society/Canadian guidelines for the diagnosis and treatment of dyslipidemia and prevention of cardiovascular disease in the adult - 2009 recommendations. *Can J Cardiol.* 2009;25(10):567-579.
41. Ghoddusi K, Ameli J, Kachuee H, Pourfarziani V, Saadat A, Karami Q. Association of diabetes mellitus and dyslipidaemias in the Tehran population. *East Mediterr Health J.* 2008;14(3):647-653.
42. Gitt AK, Junger C, Smolka W, Bestehorn K. Prevalence and overlap of different lipid abnormalities in statin-treated patients at high cardiovascular risk in clinical practice in Germany. *Clin Res Cardiol.* 2010;99(11):723-733.
43. Goodman SG, Langer A, Bastien NR, et al. Prevalence of dyslipidemia in statin-treated patients in Canada: Results of the DYSlipidemia international study (DYSIS). *Can J Cardiol.* 2010;26(9):e330-5.
44. Gordon DJ, Probstfield JL, Garrison RJ, et al. High-density lipoprotein cholesterol and cardiovascular disease. four prospective American studies. *Circulation.* 1989;79(1):8-15.

45. Gordon T, Castelli WP, Hjortland MC, Kannel WB, Dawber TR. High density lipoprotein as a protective factor against coronary heart disease. the framingham study. *Am J Med.* 1977;62(5):707-714.
46. Grover SA, Coupal L, Kaouache M, Lowensteyn I. Preventing cardiovascular disease among Canadians: What are the potential benefits of treating hypertension or dyslipidemia? *Can J Cardiol.* 2007;23(6):467-473.
47. Grundy SM, Cleeman JI, Merz CN, et al. Implications of recent clinical trials for the national cholesterol education program adult treatment panel III guidelines. *Circulation.* 2004;110(2):227-239.
48. Gu Q, Paulose-Ram R, Burt V, Kit B. Prescription cholesterol-lowering medication use in adults aged 40 and over: United states, 2003–2012. 2014.
49. Hadaegh F, Harati H, Ghanbarian A, Azizi F. Prevalence of coronary heart disease among Tehran adults: Tehran lipid and glucose study. *East Mediterr Health J.* 2009;15(1):157-166.
50. Halcox J, Misra A. Type 2 diabetes mellitus, metabolic syndrome, and mixed dyslipidemia: How similar, how different, and how to treat? *Metab Syndr Relat Disord.* 2015;13(1):1-21.
51. Heart and Stroke Foundation of Canada.
<http://www.heartandstroke.com/site/c.ikIQLcMWJtE/b.3483991/k.34A8/Statistics.htm#bloodcholesterol>. Updated 2011. Accessed 09/29, 2017.
52. Heart and Stroke Foundation of Canada. 2010 report - A perfect storm.
http://www.healthyenvironmentforkids.ca/sites/healthyenvironmentforkids.ca/files/A_Perfect_Storm.pdf. Updated 2010. Accessed 09/25, 2017.

53. Henkin RE, Harolds JA. Health information technology and the electronic medical record. *Clin Nucl Med*. 2010;35(10):788-789.
54. Heron M. Deaths: Leading causes for 2014. 2016;65(5).
55. Holt TA, Stables D, Hippisley-Cox J, O'Hanlon S, Majeed A. Identifying undiagnosed diabetes: Cross-sectional survey of 3.6 million patients' electronic records. *Br J Gen Pract*. 2008;58(548):192-196.
56. 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):512-516.
57. IBM corp. released 2013. IBM SPSS statistics for windows, version 22.0. Armonk, NY: IBM corp. <https://www.ibm.com/analytics/data-science/predictive-analytics/spss-statistical-software>. 2017.
58. Janus ED, Tideman PA, Dunbar JA, et al. Dyslipidaemia in rural Australia: Prevalence, awareness, and adherence to treatment guidelines in the greater green triangle risk factor study. *Med J Aust*. 2010;192(3):127-132.
59. Jones PH, Nair R, Thakker KM. Prevalence of dyslipidemia and lipid goal attainment in statin-treated subjects from 3 data sources: A retrospective analysis. *J Am Heart Assoc*. 2012;1(6):e001800.
60. Joseph L, Gyorkos TW, Coupal L. Bayesian estimation of disease prevalence and the parameters of diagnostic tests in the absence of a gold standard. *Am J Epidemiol*. 1995;141(3):263-272.

61. Juurlink D, Preya C, Croxford R, Chong A, Austin P, Tu J, Laupacis A. Canadian institute for health information discharge abstract database: A validation study. Toronto, ON: Institute for clinical evaluative sciences. 2006.
62. Kandula S, Zeng-Treitler Q, Chen L, Salomon WL, Bray BE. A bootstrapping algorithm to improve cohort identification using structured data. *J Biomed Inform.* 2011;44 Suppl 1:S63-8.
63. Kasiske BL, Ma JZ, Kalil RS, Louis TA. Effects of antihypertensive therapy on serum lipids. *Ann Intern Med.* 1995;122(2):133-141.
64. Kokotailo RA, Hill MD. Coding of stroke and stroke risk factors using international classification of diseases, revisions 9 and 10. *Stroke.* 2005;36(8):1776-1781.
65. Krentz AJ. Lipoprotein abnormalities and their consequences for patients with type 2 diabetes. *Diabetes Obes Metab.* 2003;5 Suppl 1:S19-27.
66. Laforest L, Ambegaonkar BM, Souchet T, Sazonov V, Van Ganse E. Mixed dyslipidemias in primary care patients in France. *Vasc Health Risk Manag.* 2012;8:247-254.
67. Lardizabal JA, Deedwania P. Lipid-lowering therapy with statins for the primary and secondary prevention of cardiovascular disease. *Cardiol Clin.* 2011;29(1):87-103.
68. Lazarus R, Klompas M, Campion FX, et al. Electronic support for public health: Validated case finding and reporting for notifiable diseases using electronic medical data. *J Am Med Inform Assoc.* 2009;16(1):18-24.
69. Lei B, Chen S, Ni D, Wang T. Discriminative learning for alzheimer's disease diagnosis via canonical correlation analysis and multimodal fusion. *Front Aging Neurosci.* 2016;8:77.

70. Li J, Motsko SP, Goehring EL,Jr, Tave A, Pezzullo JC, Jones JK. Prevalence of pediatric dyslipidemia: Comparison of a population-based claims database to national surveys. *Pharmacoepidemiol Drug Saf.* 2010;19(10):1031-1040.
71. Li J, Motsko SP, Goehring EL,Jr, Vendiola R, Maneno M, Jones JK. Longitudinal study on pediatric dyslipidemia in population-based claims database. *Pharmacoepidemiol Drug Saf.* 2010;19(1):90-98. doi: 10.1002/pds.1877 [doi].
72. Li ZY, Xu GB, Xia TA. Prevalence rate of metabolic syndrome and dyslipidemia in a large professional population in Beijing. *Atherosclerosis.* 2006;184(1):188-192.
73. Lix L, Yogendran M, Burchill C, Metge C, McKeen N, Moore D, Bond R. Defining and validating chronic diseases: An administrative data approach. Winnipeg, MB: Manitoba centre for health policy. 2006.
74. Lopez-Jimenez F, Batsis JA, Roger VL, Brekke L, Ting HH, Somers VK. Trends in 10-year predicted risk of cardiovascular disease in the United States, 1976 to 2004. *Circ Cardiovasc Qual Outcomes.* 2009;2(5):443-450.
75. MacLean DR, Petrasovits A, Nargundkar M, et al. Canadian heart health surveys: A profile of cardiovascular risk. survey methods and data analysis. Canadian heart health surveys research group. *CMAJ.* 1992;146(11):1969-1974.
76. Mantel-Teeuwisse AK, Kloosterman JM, Maitland-van der Zee AH, Klungel OH, Porsius AJ, de Boer A. Drug-induced lipid changes: A review of the unintended effects of some commonly used drugs on serum lipid levels. *Drug Saf.* 2001;24(6):443-456.

77. Manuel DG, Leung M, Nguyen K, Tanuseputro P, Johansen H, Canadian Cardiovascular Outcomes Research Team. Burden of cardiovascular disease in Canada. *Can J Cardiol*. 2003;19(9):997-1004.
78. Mayo NE, Goldberg MS, Levy AR, Danys I, Korner-Bitensky N. Changing rates of stroke in the province of Quebec, Canada: 1981-1988. *Stroke*. 1991;22(5):590-595.
79. McAdam-Marx C, Ye X, Sung JC, Brixner DI, Kahler KH. Results of a retrospective, observational pilot study using electronic medical records to assess the prevalence and characteristics of patients with resistant hypertension in an ambulatory care setting. *Clin Ther*. 2009;31(5):1116-1123.
80. McCarty CA, Chisholm RL, Chute CG, et al. The eMERGE network: A consortium of biorepositories linked to electronic medical records data for conducting genomic studies. *BMC Med Genomics*. 2011;4:13-8794-4-13.
81. McPeck Hinz ER, Bastarache L, Denny JC. A natural language processing algorithm to define a venous thromboembolism phenotype. *AMIA Annu Symp Proc*. 2013;2013:975-983.
82. Menachemi N, Collum TH. Benefits and drawbacks of electronic health record systems. *Risk Manag Healthc Policy*. 2011;4:47-55.
83. Meng X, Jiang R, Lin D, et al. Predicting individualized clinical measures by a generalized prediction framework and multimodal fusion of MRI data. *Neuroimage*. 2017;145(Pt B):218-229.
84. Minard JP, Turcotte SE, Lougheed MD. Asthma electronic medical records in primary care: An integrative review. *J Asthma*. 2010;47(8):895-912.

85. Mitiku TF, Tu K. Using data from electronic medical records: Theory versus practice. *Healthc Q*. 2008;11(4):23-25.
86. Mohammadbeigi A, Farahani H, Moshiri E, et al. Prevalence of metabolic syndrome and associations with lipid profiles in Iranian men: A population-based screening program. *World J Mens Health*. 2018;36(1):50-56.
87. National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). Third report of the national cholesterol education program (NCEP) expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (adult treatment panel III) final report. *Circulation*. 2002;106(25):3143-3421.
88. National Clinical Guideline Centre (UK). Lipid Modification: Cardiovascular Risk Assessment and the Modification of Blood Lipids for the Primary and Secondary Prevention of Cardiovascular Disease. 2014.
89. Nicholson K, Stewart M, Thind A. Examining the symptom of fatigue in primary care: A comparative study using electronic medical records. *J Innov Health Inform*. 2015;22(1):235-243.
90. Nishida Y, Takahashi Y, Nakayama T, Soma M, Kitamura N, Asai S. Effect of candesartan monotherapy on lipid metabolism in patients with hypertension: A retrospective longitudinal survey using data from electronic medical records. *Cardiovasc Diabetol*. 2010;9:38.
91. Orozco-Beltran D, Gil-Guillen VF, Redon J, et al. Lipid profile, cardiovascular disease and mortality in a Mediterranean high-risk population: The ESCARVAL-RISK study. *PLoS One*. 2017;12(10):e0186196.

92. O'Malley KJ, Cook KF, Price MD, Wildes KR, Hurdle JF, Ashton CM. Measuring diagnoses: ICD code accuracy. *Health Serv Res.* 2005;40(5 Pt 2):1620-1639.
93. Padwal R, Hackam D, Khan N, Tobe S. Primary prevention of CVD: Modification of diet in people with hypertension. *BMJ Clin Evid.* 2016;2016:0214.
94. Petrella RJ, Merikle E, Jones J. Prevalence and treatment of dyslipidemia in Canadian primary care: A retrospective cohort analysis. *Clin Ther.* 2007;29(4):742-750.
95. Pettersson B, Ambegaonkar B, Sazonov V, Martinell M, Stalhammar J, Wandell P. Prevalence of lipid abnormalities before and after introduction of lipid modifying therapy among Swedish patients with dyslipidemia (PRIMULA). *BMC Public Health.* 2010;10:737.
96. Pope AM, Carr SM, Smith KN, Marshall HD. Mitogenomic and microsatellite variation in descendants of the founder population of Newfoundland: High genetic diversity in an historically isolated population. *Genome.* 2011;54(2):110-119.
97. Protti D. Comparison of information technology in general practice in 10 countries. *Healthc Q.* 2007;10(2):107-116.
98. Public Health Agency of Canada. Report from the Canadian chronic disease surveillance system: Hypertension in Canada, 2010. http://www.phac-aspc.gc.ca/cd-mc/cvd-mcv/ccdss-snsmc-2010/pdf/CCDSS_HTN_Report_FINAL_EN_20100513.pdf. Updated 2010. Accessed 02/07, 2017.
99. Public Health Agency of Canada. Tracking heart disease and stroke in Canada. www.phac-aspc.gc.ca/publicat/2009/cvd-avc/pdf/cvd-avs-2009-eng.pdf. Updated 2009. Accessed 09/25, 2017.

100. Qi L, Ding X, Tang W, Li Q, Mao D, Wang Y. Prevalence and risk factors associated with dyslipidemia in Chongqing, China. *Int J Environ Res Public Health*. 2015;12(10):13455-13465.
101. Rasmussen-Torvik LJ, Pacheco JA, Wilke RA, et al. High density GWAS for LDL cholesterol in African Americans using electronic medical records reveals a strong protective variant in APOE. *Clin Transl Sci*. 2012;5(5):394-399.
102. Robitaille C, Dai S, Waters C, et al. Diagnosed hypertension in Canada: Incidence, prevalence and associated mortality. *CMAJ*. 2012;184(1):E49-56.
103. Sarwar N, Danesh J, Eiriksdottir G, et al. Triglycerides and the risk of coronary heart disease: 10,158 incident cases among 262,525 participants in 29 western prospective studies. *Circulation*. 2007;115(4):450-458.
104. Sharp RP, Sirajuddin R, Sharief IM. Impact of carvedilol on the serum lipid profile. *Ann Pharmacother*. 2008;42(4):564-571.
105. Shea JL, Loredó-Osti JC, Sun G. Association of RBP4 gene variants and serum HDL cholesterol levels in the Newfoundland population. *Obesity (Silver Spring)*. 2010;18(7):1393-1397.
106. Singer A, Yakubovich S, Kroeker AL, Dufault B, Duarte R, Katz A. Data quality of electronic medical records in Manitoba: Do problem lists accurately reflect chronic disease billing diagnoses? *J Am Med Inform Assoc*. 2016;23(6):1107-1112.
107. Smith DG. Epidemiology of dyslipidemia and economic burden on the healthcare system. *Am J Manag Care*. 2007;13 Suppl 3:S68-71.
108. Smith ER. The Canadian heart health strategy and action plan. *Can J Cardiol*. 2009;25(8):451-452.

109. Soran H, Dent R, Durrington P. Evidence-based goals in LDL-C reduction. *Clin Res Cardiol.* 2017;106(4):237-248.
110. Stata SE 11. <http://www.stata.com>. Updated 2015. Accessed 07/13, 2013.
111. Statistics Canada. Population and dwelling counts, for Canada, census metropolitan areas, census agglomerations and census subdivisions (municipalities), 2011 and 2006 censuses. <http://www12.statcan.gc.ca/census-recensement/2011/dp-pd/hlt-fst/pd-pl/Table-Tableau.cfm?LANG=Eng&T=203&CMA=001&S=0&O=D&RPP=25>. Updated 2016. Accessed 09/28, 2017.
112. Statistics Canada. Canadian health measures survey. <http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=5071&lang=en&db=imdb&adm=8&dis=2>. Updated 2015. Accessed 09/28, 2017.
113. Statistics Canada. Mortality, summary list of causes. <http://www.statcan.gc.ca/pub/84f0209x/84f0209x2007000-eng.htm>. Updated 2012. Accessed 03/17, 2017.
114. Statistics Canada. Canadian health measures survey. <http://www.statcan.gc.ca/daily-quotidien/100323/dq100323a-eng.htm>. Updated 2012. Accessed 09/28, 2017.
115. Statistics Canada. Canadian health measures survey. http://www.statcan.gc.ca/access_acces/archive.action?loc=/pub/82-623-x/82-623-x2010002-eng.pdf&archive=1. Updated 2011. Accessed 09/28, 2017.

116. Statistics Canada. Heart health and cholesterol levels of Canadians, 2007 to 2009. <http://www.statcan.gc.ca/pub/82-625-x/2010001/article/11136-eng.pdf>. Updated 2010. Accessed 09/25, 2012.
117. Statistics Canada. Rural and small town Canada analysis bulletin. <http://www.statcan.gc.ca/pub/21-006-x/21-006-x2001003-eng.pdf>. Updated 2001. Accessed 09/28, 2017.
118. Strauss JA, Chao CR, Kwan ML, Ahmed SA, Schottinger JE, Quinn VP. Identifying primary and recurrent cancers using a SAS-based natural language processing algorithm. *J Am Med Inform Assoc*. 2013;20(2):349-355.
119. Sui J, Adali T, Yu Q, Chen J, Calhoun VD. A review of multivariate methods for multimodal fusion of brain imaging data. *J Neurosci Methods*. 2012;204(1):68-81.
120. Szklo M NF. *Epidemiology: Beyond the basics*. 2nd edition ed. Jones and Bartlett Publishers; 2007.
121. Tirschwell DL, Longstreth WT, Jr. Validating administrative data in stroke research. *Stroke*. 2002;33(10):2465-2470.
122. Total and high-density lipoprotein cholesterol in adults: United states, 2011-2014. <https://www.cdc.gov/nchs/data/databriefs/db226.htm>. Updated 2015. Accessed 10/01, 2017.
123. Tu K, Mitiku T, Guo H, Lee DS, Tu JV. Myocardial infarction and the validation of physician billing and hospitalization data using electronic medical records. *Chronic Dis Can*. 2010;30(4):141-146.
124. Tu K, Mitiku T, Lee DS, Guo H, Tu JV. Validation of physician billing and hospitalization data to identify patients with ischemic heart disease using data from

- the electronic medical record administrative data linked database (EMRALD). *Can J Cardiol.* 2010;26(7):e225-8.
125. Tulenko TN, Sumner AE. The physiology of lipoproteins. *J Nucl Cardiol.* 2002;9(6):638-649.
 126. Verges B. New insight into the pathophysiology of lipid abnormalities in type 2 diabetes. *Diabetes Metab.* 2005;31(5):429-439.
 127. Vining DJ, Gladish GW. Receiver operating characteristic curves: A basic understanding. *Radiographics.* 1992;12(6):1147-1154.
 128. Wang S, Xu L, Jonas JB, You QS, Wang YX, Yang H. Prevalence and associated factors of dyslipidemia in the adult Chinese population. *PLoS One.* 2011;6(3):e17326.
 129. Williams WR, Schneider KA, Borhani NO, Schnaper HW, Slotkoff LM, Ellefson RD. The relationship between diuretics and serum cholesterol in hypertension detection and follow-up program participants. *Am J Prev Med.* 1986;2(5):248-255.
 130. Wockenfuss R, Frese T, Herrmann K, Claussnitzer M, Sandholzer H. Three- and four-digit ICD-10 is not a reliable classification system in primary care. *Scand J Prim Health Care.* 2009;27(3):131-136.
 131. World Health Organization. The Anatomical Therapeutic Chemical classification system. <http://www.who.int/classifications/atcddd/en/>. Updated 2003. Accessed 04/01, 2018.
 132. World Health Organization. Global status report on noncommunicable diseases 2010. Geneva, Switzerland. 2011.
 133. World Health Organization. Estimated deaths per 100,000 population by cause, and member state. 2008.

134. Wu JY, Duan XY, Li L, et al. Dyslipidemia in shanghai, China. *Prev Med.* 2010;51(5):412-415.
135. Young WB, Ryu H. Secondary data for policy studies: Benefits and challenges. *Policy, Politics, & Nursing Practice.* 2000;1(4):302-307.
136. Yusuf S, Hawken S, Ounpuu S, et al. Obesity and the risk of myocardial infarction in 27,000 participants from 52 countries: A case-control study. *Lancet.* 2005;366(9497):1640-1649.
137. Zhang L, Qin LQ, Liu AP, Wang PY. Prevalence of risk factors for cardiovascular disease and their associations with diet and physical activity in suburban Beijing, China. *J Epidemiol.* 2010;20(3):237-243.
138. Zhang Q, Rajagopalan S, Mavros P, et al. Differences in baseline characteristics between patients prescribed sitagliptin versus exenatide based on a US electronic medical record database. *Adv Ther.* 2010;27(4):223-232.